ISBN 978-969-8858-25-4

PROCEEDINGS

Vol. 34

15th Islamic Countries Conference on Statistical Sciences (ICCS-15)

"Statistics for Social Justice with Healthy Living"

Dedicated to

Late Prof. Dr. Munir Ahmad

(Founding President ISOSS)

December 21-24, 2019

at

Lahore Institute of Science & Technology (LIST)

Lahore (Pakistan)

JOINTLY ORGANIZED BY

ISOSS

Islamic Countries Society of Statistical Sciences
44-A, Civic Centre, Sabzazar, Multan Road, Lahore, Pakistan
Email: secretary@isoss.net
URL: www.isoss.net

Lahore Institute of Science & Technology
7-Madina Park, Bedian Road, Lahore Cantt, Pakistan

Published by: ISOSS, Lahore, Pakistan.

"All papers published in the

PROCEEDINGS

were accepted after formal peer review

by the experts in the relevant field.

Editors:

- **Professor Dr. Muhammad Hanif**
  National College of Business
  Administration and Economics
  Pakistan

- **Professor Dr. Shahjahan Khan**
  University of Southern Queensland
  Australia

CONTENTS

FOREWORD

The Islamic Countries Society of Statistical Sciences (ISOSS) is pleased to publish the Proceedings of the 15th Islamic Countries Conference on Statistical Sciences (ICCS-15). This adds a new crown to the long history and tradition of promoting creative research, scholarly work and research management of ISOSS.

The theme of the conference "Statistics for Social Justice and Healthy Living" was very timely and highly appropriate in the context of the contemporary world. The conference attracted delegates and presenters from many countries from the continents of Australia, Europe, USA, and Asia. Papers were presented from diverse areas of statistics and related disciplines. These were a good mixture of theoretical and applied work.

High quality selected research articles from among the papers that were presented in the conference were accepted via peer review by the experts of the relevant field for publication in this Proceedings. On behalf of ISOSS, we are thankful to all contributing authors and reviewers for their scholarly work and professional support.

As per earlier announcement of late Dr. Munir Ahmad, the ISOSS decided to introduce the best paper award in ISOSS conferences and awarded the inaugural "Dr. Munir Ahmad Best Paper Award" in ICCS-15. All the presented papers were evaluated by the members of the Award Committee and the winner was announced in the Business Session/Closing Ceremony. The Award was given to ***Dr. Ahmad Farooqi*** from Wayne State University, School of Medicine, USA for his paper entitled ***"Simulation and Comparison of Theil Sen Regression, Theil Sen Siegel Regression, and Quantile Regression with Outliers in Small Samples"***.

This was the first ISOSS conference after the demise of it's founder Professor Dr. Munir Ahmad on 29 June, 2019. This was a challenging time for ISOSS. There is no one there to feel the shoes of Professor Ahmad. However, many of his esteemed colleagues from all over the world are highly committed to continue the mission that he initiated. The ICCS-15 held in Lahore during 21-24 December, 2019 is an expression of determination of ISOSS family to move forward.

Our gratefulness is to Mr. Muhammad Iftikhar and his dedicated team for managing the preparation and publication of the Proceedings with the highest professional standard.

We sincerely believe that the Proceedings will be very useful for the research community.

Finally, we thank everyone involved in the organization of ICCS-15, specially the key people in various committees and sponsors.

Editors:

- **Professor Dr. Muhammad Hanif**
  National College of Business
  Administration and Economics, Pakistan

- **Professor Dr. Shahjahan Khan**
  University of Southern Queensland, Australia

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 1-14

SIMULATION AND COMPARISON OF THEIL SEN REGRESSION, THEIL SEN SIEGEL REGRESSION, AND QUANTILE REGRESSION WITH OUTLIERS

Ahmad Farooqi and **Ronald Thomas**
CRCM, Department of Pediatrics, Wayne State University School of Medicine
Detroit, MI, 48201, United States of America
Email: afarooq@med.wayne.edu; rthomas@med.wayne.edu

ABSTRACT

Researchers in medical, social and behavioral, sciences usually interested to know the relationship between a response variable Y_i and one or more independent predictors X_i either for the purpose of estimation and prediction. Ordinary Least Square Regression is a parametric approach used to study this kind of relationship. One of the disadvantages of Ordinary Least Square is it does not fit well in the presence of outliers. Quantile Regression, Theil-Sen Regression, and the modified Theil-Sen Siegel Regression are non-parametric approaches are very useful to study the relationship between a response variable Y_i and one or more independent predictors X_i and are more robust methods to outliers as compare to Ordinary Least Square Regression. Several comparisons are made among Ordinary Least Square Regression, Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression, but no direct comparison is yet made between Quantile Regression, Theil Sen Regression and Theil Sen Siegel Regression in the presence of outliers. In order to investigate this, Monte Carlo simulation study was employed and observations were generated using normal distribution with different small sample sizes in the presence of 10%, 20%, 30% and 50% outliers of the sample sizes. A comparison based on, graphs, Standard Errors (S.E), Confidence Interval, Median Absolute Error (MEDAE), and Relative Median Absolute Error (RMEDAE) are made with all three non-parametric regression procedures to evaluate the model fitting. The results of the study showed, when there are outliers in X and Y direction Theil Sen Siegel Regression should be the most suitable regression procedure sequentially followed by Quantile Regression and Theil Sen Regression.

KEYWORDS

Theil Sen Regression, Theil Sen Siegel Regression, Quantile Regression, Outliers, Small Samples.

1. INTRODUCTION

Researchers in medical, social and behavioral, sciences are always interested to study the relationship between a response variable Y_i and one or more independent predictors X_i, either for the purpose of estimation and prediction. However, a weakness invariably occurring and often found in studies of medical and educational research are the presence of data outliers in the response variable or both in response variable and predictor variable. The presence of outliers can increase error variance and reduce the power of statistical tests. Unfortunately, outliers can often times be impossible to prevent even when data has been carefully collected from respectable sources. Since outliers affect analysis and interpretation of statistical outcomes, a greater understanding of statistical procedures for handling the presence of outliers in medical and educational datasets is needed. Specifically, identifying which procedures operate best in identifying and handling data corruption from outliers would be a welcome contribution to the field for scholars, data analysts, medical and educational researchers. Institutions can benefit from the results of this research by providing recommendations useful in practice and substantive interest, leading to more accurate and reliable conclusions and decision making. Ordinary Least Square (OLS) Regression is a parametric approach used to study the relationship between a response variable Y_i with at least one predictor X_i, also called an independent variable by describing the mean of response variable for each value of the given predictors, using a function called the conditional mean of the response variable. This relationship can be created by developing a statistical model with certain unknown population parameters called regression coefficients. The parameters are then estimated by the method of lease square and the fitted model is used to get an approximate idea of the trend for prediction and forecasting of the data. Although OLS estimators, enjoys the properties of best linear unbiased and minimum variance among all the unbiased linear or nonlinear estimators, yet there are certain disadvantages of OLS. It provides only a partial view of the relationship between response and independent variable(s) through a conditional mean point of the response variable. There may be interest in studying the relationship at different points in the conditional distribution of response variable. Also, it does not fit well when there are some regular outliers present in the response variable, or if the data were sampled from a non-normal distribution. Therefore, the OLS regression estimator is not robust (e.g., Hampel et al., 1987; Huber & Ronchetti, 2009; Maronna et al. 2006; Staudte & Sheather, 1990; Wilcox, 2012a, b). There are various alternatives to the OLS modeling has roots that can be traced to the mid of the 18th century, one of the alternative approaches can be referred to as median regression, where focus of the modeling is the median instead of mean. It is to me noted that median is the special case of quantile, which can be used to model the non-central position of a distribution. Quantile Regression is another very flexible approach that can be used to study the relationship between a response variable Y_i with at least one predictor X_i at different points in the conditional distribution of response variable, using the conditional median or other quantile functions, where the median is the 50^{th} percentile; and the quartiles are the 25^{th}, 50^{th} and 75^{th} percentiles. Similarly, the deciles are the 10^{th}, 20^{th} and so on until the 90^{th} percentile of the empirical distribution can be used to study the response variable. Theil (1950) first proposed another robust linear regression method where there are one response and one predictor variable and is robust to outliers in the response variable. In this method, the slope of the regression line is estimated as the median of all pairwise slopes between each pair of points in the dataset. Sen (1968) extended this estimator to handle ties. The Theil-Sen estimator (TSE) is robust with a high breakdown point 29.3%, has a bounded influence function, and possesses a high asymptotic efficiency. A modified and preferred method is Siegel (1982), where the repeated median is used, the repeated median algorithm is a robustified U-statistic (Hoeffing, 1948), in which nested medians replaces the single median and has a 50% breakdown point. Regular outliers and non-normality are problems in statistics, an outlier is a value that is far from the general distribution of the other observed values, and can often perturb the results of a statistical analysis (Greenacre and Ayhan, 2015). As a result of these outliers there may be a breakdown in the model at the ith point produce a location shift and the variance exceed the error variance at the other data points, also there may be a large random disturbance that can be produced by chance.

1.1 Objectives of the Study

Several comparisons are made between Ordinary Least Square Regression, Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression but no direct comparison is yet made between Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression in the presence of outliers. So, the main objective of this study is to make comparison between Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression in the presence of outliers.

1.2 Limitations of the Study

- There are several other non-parametric regression methods, but our focus is on Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression.
- This study focused on 50^{th} percentile i.e. median in Quantile Regression only.
- This study focused on simple linear regression with small sample only.
- This study focused on outliers in X and Y direction only.

2. LITERATURE REVIEW

2.1 Quantile Regression

Quantile Regression is another very flexible approach developed by Koenker and Bassett (1978), that can be used an alternative of ordinary lease square regression and allows researcher to study the relationship between a response variable Y_i with at least one predictor X_i at different points in the conditional distribution of response variable Y_i , at several points using the conditional median function $Q_q(Y_i|X_i)$, or other quantile function where median is the 50^{th} percentile and is the best-known quantile, similarly, the other quartiles, e.g. 25^{th}, 30^{th}, 75^{th} , and so on 95^{th} percentiles or simply a qth quantile q, of the empirical distribution $F(Y)$ can be defined. It is to be noted that quantile and percentiles are synonymous i.e. the 0.90 quantile is the 90^{th} percentile. Using Quantile Regression investigator/researcher can see a more comprehensive picture of the effect of the predictors on the response variable.

2.2 Quantile Regression Model

The q^{th} quantile, $0 < q < 1$ split the response variable Y_i into proportion q below and $1 - q$ above such that $F(Y_q) = q$ and $Y_q = F^{-1}(q)$ for the median $q = 0.5$. Quantile Regression uses an asymmetric weighting system of data points and therefore, all data points are weighted based on their distance from the researcher-specified quantile for that estimation. Consequently, Quantile Regression is not synonymous with fitting a separate OLS regression line at each quantile (Petscher & Logan, 2014; Petscher et al., 2013). In Quantile Regression the fitted model is

$$Y_i = \beta_{q0} + \beta_{q1}X_i + \varepsilon_{qi}, i = 1, 2, \ldots n,$$

where $\varepsilon_{qi}{}'s$ are random errors, β_{q0} and β_{q1} are the unknown parameters associated with q^{th} quantile, and $0 < q < 1$. Recall in the Ordinary Least Square Regression model, the conditional mean is

$$E(Y_i|X_i) = \beta_0 + \beta_1 X_i.$$

In contrast, for the corresponding Quantile Regression model, the q^{th} conditional quantile given X_i is specified as,

$$Q_q(Y_i|X_i) = \beta_{q0} + \beta_{q1} X_i.$$

Thus, the q^{th} quantile is determined by the quantile specific parameters β_{q0} and β_{q1}, with a specific predictor value of X_i. Like, Ordinary Least Square Regression the $E(\varepsilon_i) = 0$, in Quantile Regression $Q_q(\varepsilon_{qi}) = 0$. It is to be noted that for different values of the quantile q of interest, the error terms ε_{qi} for fixed i are related. By extending the idea of several equations can be expressed at different quantiles. For example, if the Quantile Regression model specifies the 9^{th} quantiles, the 9 different models yields 9 Quantile Regression coefficients for X_i, one at each of the 9 conditional quantiles, i.e. $\beta_{0.10}, \beta_{0.20}, \ldots, \beta_{0.90}$. In ordinary lease square regression, the least squares (LS) method tries to minimize $\sum_{i=1}^{n} \varepsilon_i{}^2$, the sum of squared error from the fitted straight line to the observed outcome variable Y_i whereas, in Quantile Regression absolute sum of error from the fitted q^{th} line to the observed outcome variable Y_i, is tried to minimize, i.e. $\sum_{i=1}^{n}|\varepsilon_{qi}|$ is to minimize.

2.3 Estimation of Parameters in Quantile Regression with Single Independent Predictor

Suppose, in a Quantile Regression $\hat{\beta}_{q0}$, and $\hat{\beta}_{q1}$ are the estimates of the corresponding unknown parameters β_{q0} and β_{q1} respectively. A method of the absolute sum of error is used to estimate the parameters by minimizing the sum of absolute errors. The attempt is to minimize $\sum_{i=1}^{n}|\varepsilon_{qi}|$. The Quantile Regression minimizes the $\sum_{i=1}^{n} q|\varepsilon_{qi}| + \sum_{i=1}^{n}(1-q)|\varepsilon_{qi}|$, a sum that gives the asymmetric penalties $q|\varepsilon_{qi}|$ for under prediction and $(1-q)|\varepsilon_{qi}|$ for over prediction. For example, in a median regression, if $q = 0.5$ then the quantity $\sum_{i=1}^{n}|\varepsilon_{qi}|$ will collapse to a median regression. In ordered to find the quartile regression coefficients, a criterion function is defined for q^{th} Quantile Regression estimator $\hat{\beta}_{q1}$ that minimizes $Q(\beta_q)$ with objective function along with penalty q when response variable is higher than the predicted values i.e. $Y_{i \geq X_{i\beta}}$ and penalty $1-q$ when response variable is higher than the predicted values i.e. $Y_{i < X_{i\beta}}$.

$$Q(\beta_q) = \sum_{i=1}^{n} q|\varepsilon_{qi}| + \sum_{i=1}^{n}(1-q)|\varepsilon_{qi}|$$

$$Q(\beta_q) = \sum_{i:Y_{i \geq X_{i\beta}}} q|Y_i - \beta_{q0} - \beta_{q1}X_i| + \sum_{i:Y_{i<X_{i\beta}}}(1-q)|Y_i - \beta_{q0} - \beta_{q1}X_i|,$$

where $\varepsilon_{qi} = Y_i - \beta_{q0} - \beta_{q1}X_i$ equivalently, it can be written as

$$Q(\beta_q) = \sum_{i=1}^{n} \left[I_{\{Y_{i \geq X_{i\beta}}\}} q \left| Y_i - \beta_{q0} - \beta_{q1} X_i \right| + I_{\{Y_{i<X_{i\beta}}\}} (1-q) \left| Y_i - \beta_{q0} - \beta_{q1} X_i \right| \right],$$

where $0 < q < 1$ and I is an indicator function. In contrast to ordinary lease square regression or maximum likelihood, the Quantile Regression computational implementation uses linear programming method to find the regression coefficients.

2.4 Theil Sen Siegel Regression Model

Theil (1950) regression is another completely nonparametric approach to linear regression with one predictor and one response variable. The Theil estimator provides a robust estimator for linear regression and outliers in the response variable. When the estimator is a line, then the Ordinary Least Square estimate corresponds with the mean, and is not robust estimate. A single point can easily affect the slope of the line. The Theil estimator is a robust version of a linear regression. It simply computes all the lines between each pair of points and uses the median of the slopes of these lines. This procedure is sometimes called Theil–Sen procedure. A modified, and more robust, method is named after Siegel. The method yields a slope and intercepts for the fit line, and a p-value for the slope can be determined as well. Typically, no measure analogous to r-squared is reported. Theil-Sen single (1950) median method computes slopes of lines crossing all possible pairs of points, when x coordinates differ. After calculating these $n\,(n-1)/2$ slopes (these values are true only if X_i is distinct), the median of them is taken as slope estimator. Next, the intercepts of n lines, crossing each point and having calculated slope are calculated. The median from them is intercept estimator. Sen (1968) extended this estimator to handle ties and obtained unbiasedness and asymptotic normality of the estimator for absolutely continuous error distribution and a no identical covariate. A variability of the Theil-Sen estimator due to Siegel (1982) determines, for each sample point, the median m_i of the slopes of lines through that point, and then determines the overall estimator as the median of these medians. These repeated medians are sometime more complicated. For each point, the slopes between it and the others are calculated (resulting $(n-1)$ slopes) and the median is taken. This results in n medians and median from these medians is slope estimator. Intercept is calculated in similar way. The breakdown point of Theil-Sen method is about 29% and, Siegel extended it to 50%, so these regression methods are very robust. Additionally, if the errors are normally distributed and no outliers are present, the estimators are very similar to classic least squares.

2.5 Estimation of Parameters Kendall–Theil Sen Siegel Regression with Single Independent Predictor

In Theil Sen Regression, suppose a sample of n pair of observation $(X_i,\ Y_i)$, $i = 1, 2, \ldots n$ were taken. Suppose a straight line given below is used as a best fit model to the given set of data.

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i,\ i = 1, 2, \ldots n$$

where, Y_i is the response variable for each data point (i), X_i is the predictor variable for each data point (i), ε_i is the residual in prediction of Y for each data point (i), $\hat{\beta}_1$ is the estimated regression coefficient, $\hat{\beta}_0$ is the estimated Y intercept, and n is the number of

XY data points in the sample. In Kendall–Theil Sen Regression, the regression coefficient β_1 can be estimated by the median of all pairwise slopes between each pair of points in the given data set (Theil, 1950; Sen, 1968; Helsel and Hirsch, 2002). Each regression coefficient passing through (X_i, Y_i) and (X_j, Y_j), the data point can be estimated by the

$$\hat{\beta}_1 = Median_j \left\{ \frac{(Y_j - Y_i)}{(X_j - X_i)}, \text{for } i = 1, 2, \ldots (n-1), X_j \neq X_i \text{and } j = 2, \ldots n \right\}$$

The number of possible regression coefficients between data pairs can be calculated by $N_p = \frac{n(n-1)}{2}$. All possible estimated b_{ij} are sorted and ranked by ascending order. Sorting is a computationally intensive process because each slope estimate in the array of slopes must be compared to other values and put in the proper order. If N_p is an odd number, the median slope is selected as the middle value of the array otherwise, the median is calculated as the arithmetic average of the two center points. The Y-intercept of the line can be estimate by the equation used by Conover (1980) as

$$\hat{\beta}_0 = \{\tilde{Y} - \hat{\beta}_1 \tilde{X}\}$$

where, $\hat{\beta}_0$ is the estimated Y-intercept, $\tilde{Y}$ is the median of the response variable, $\hat{\beta}_1$ is the estimated slope, and $\tilde{X}$ is the median of the predictor variable. The error tem ε_i are the random errors, should be independently and normally distributed, i.e. $\varepsilon_i \sim N(0,\sigma^2)$. However, In Kendall-Theil regression model these assumptions associated with error term are not bounded to fulfill. If N_p is an odd number, the median regression coefficient is selected as the middle value of the array, otherwise, median regression coefficient is selected by taking the mean of the two middle values of an array. Hence the estimated Theil Sen Regression line is

$$\hat{Y}_{TS} = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Siegel (1982) considered repeated medians. For each observation (X_i, Y_i), the regression coefficients between it and the others $(n-1)$ are calculated and the median is taken. This results in n medians and median from these medians is a regression coefficients estimator. A robust estimator $\tilde{\beta}_1$ of the slop can be estimated by taking the median of these repeated medians for each point (X_i, Y_i) i.e.

$$\tilde{\beta}_1 = Median_i \left\{ \hat{\beta}_1 = \frac{(Y_j - Y_i)}{(X_j - X_i)} : X_i \neq X_j, 1 \leq i < j \leq n \right\}.$$

Similarly, the y-intercept can be estimated by the medians of all possible least square estimates

$$\tilde{\beta}_0 = Median_i \left\{ \hat{\beta}_0 = \frac{(Y_j X_i - Y_i X_j)}{(X_j - X_i)} : X_i \neq X_j, 1 \leq i < j \leq n \right\}.$$

Hence the estimated Theil Sen Siegel Regression line is

$$\hat{Y}_{TSS} = \tilde{\beta}_0 + \tilde{\beta}_1 X_i$$

3. METHODOLOGY

Monte Carlo simulation technique will be used to generate from randomly sampled observation with replacement from a normal distribution and estimate of regression coefficients, standard errors, median absolute deviation, p-values, and confidence intervals are calculated, based on Quantile Regression, Theil Sen Regression and Theil Sen Siegel Regression. A comparison will be made using these three regression techniques. The author wrote several essential codes in R in order to compare regression coefficient, confidence intervals, test of significance and to generate different figures.

3.1 Procedure

Observations for the Monte Carlo simulations will be randomly generated with replacement from the normal distribution using statistical software R. Similarly, observations for the Monte Carlo simulations will be randomly generated with different small sample sizes in the presence of 10%, 20%, 30% and 50% outliers. The idea for outliers is to find inner and outer fences of the data using Grubbs' test (Grubbs 1969). We use outer fence for major our outliers, with lower Bound=Q1–(3*(Q3–Q1)) and Upper Bound=Q3+(3*(Q3–Q1)), such that a data point that falls outside the inner fence called a major outlier. Values of regression coefficients, confidence intervals and test of significance will be obtained and tested by fitting four regression models to the simulated data. Parametric values of regression coefficients were sat at certain values to generate response variable Y_i . After each sample has been generated, the regression coefficients, standard errors, and confidence intervals are constructed based on Quantile Regression and the Theil Sen Siegel Regression. A comparison based on Median Absolute Error (MEDAE) and Relative Median Absolute Error (RMEDAE) of the three regression methods to evaluate the model fit. A negative value of Relative Median Absolute Error (RMEDAE) refer to a proportional increase in MEDAE obtained by other regression model, on the other hand positive value of Relative Median Absolute Error (RMEDAE) refer to a proportional decrease in MEDAE obtained by other regression model (Syed et al., 2016). This procedure is repeated some hundred thousand times for different sample sizes.

4. RESULTS

A visual as well as numerical comparison was made using these three regression procedures. For visual comparison scatter plots with fitted regression lines using all three regression procedures were used. For numerical compression, standard errors, median absolute deviation, confidence intervals, median absolute error (MEDAE), and relative median absolute error were used.

4.1 Regression Procedure under the Normality Assumption with Outliers in both X and Y Direction

If the errors (e_i) are independent and normally, then a random sample of size n was generated from a bivariate normal distribution with mean (0, 0) and variances equal to 1, and a correlation coefficient equal to 0.80. Outliers of 10%, 30% 50% and 100% of n was generated in both X and Y variables from a bivariate normal distribution with means (3, 6) with variances 0.1*variance of the above bivariate normal distribution, i.e. the variances (0.1, 0.1).

Table 1
Results from the Three Regression Procedures with Outliers of 10%, 20%, 30% and 50% of n=10, in both X and Y Direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80

Outliers	Regression	$\hat{\beta}_0$	$\hat{\beta}_1$	S.E($\hat{\beta}_1$)	P-value	95% CI for β_1	MEDAE
10%	**QR**	0.22	1.06	0.437	0.038	(0.64, 1.48)	0.408
	TS	0.11	0.79	0.247	0.011	(0.23, 1.35)	0.163
	TSS	0.08	0.74	0.260	0.019	(0.15, 1.33)	0.204
20%	**QR**	0.28	0.87	0.632	0.197	(-0.55, 1.19)	0.298
	TS	0.28	0.86	0.332	0.026	(0.13, 1.61)	0.297
	TSS	0.28	0.87	0.330	0.025	(0.14, 1.61)	0.298
30%	**QR**	-0.16	2.05	0.621	<0.01	(1.25, 2.86)	0.941
	TS	-0.16	1.42	0.303	<0.001	(0.76, 2.09)	0.644
	TSS	-0.38	1.11	0.401	0.018	(0.23, 1.99)	0.523
50%	**QR**	-1.07	2.24	0.425	<0.001	(1.65, 2.83)	0.554
	TS	0.05	1.86	0.218	<0.001	(1.39, 2.33)	1.010
	TSS	-0.67	1.87	0.223	<0.001	(1.39, 2.35)	1.016

Table 2
Results of Relative Median Absolute Error of the Four Regression Procedures with Outliers of 10%, 20%, 30% and 50% of n=10, in both X and Y Direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80

Outliers	Relative Median Absolute Error	Value
10%	QR vs TS	0.599
	QR vs TSS	0.498
	TS vs TSS	-0.235
20%	QR vs TS	0.003
	QR vs TSS	0.000
	TS vs TSS	-0.003
30%	QR vs TS	-0.015
	QR vs TSS	0.176
	TS vs TSS	0.188
50%	QR vs TS	-0.822
	QR vs TSS	-0.835
	TS vs TSS	-0.006

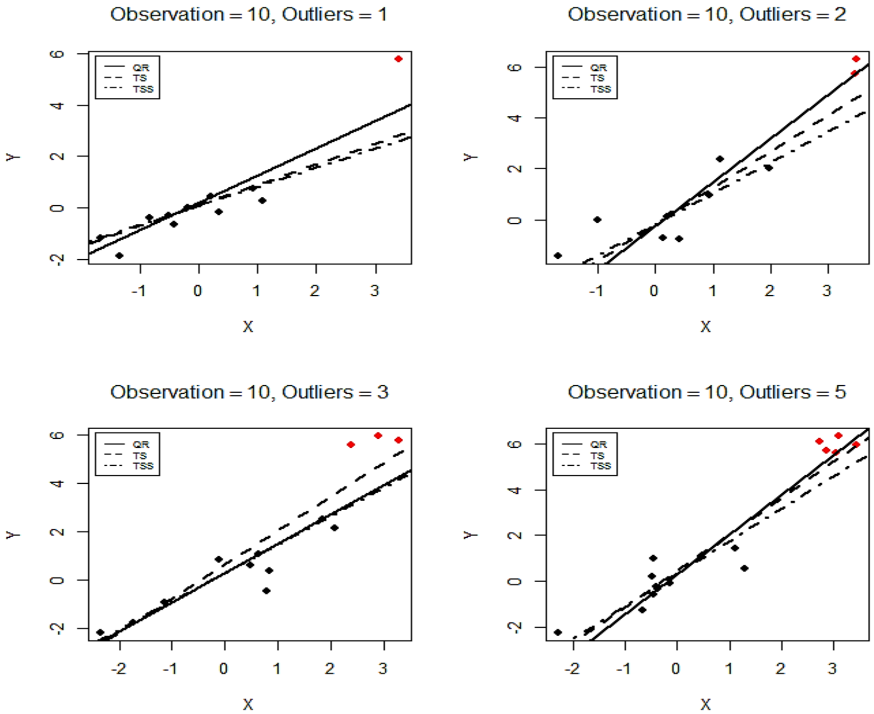


Figure 1: Four Regression Lines are shown in each Plot with n=10 and Outliers of 10%, 20%, 30% and 50% of n=10 in both X and Y Direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80

Table 3
Results from the Three Regression Procedures with Outliers of 10%, 20%, 30% and 50% of n=20, in both X and Y Direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80

Outliers	Regression	$\hat{\beta}_0$	$\hat{\beta}_1$	S.E($\hat{\beta}_1$)	P-value	95% CI for β_1	MEDAE
10%	QR	-0.29	1.03	0.256	<0.001	(0.86, 1.20)	0.406
	TS	-0.16	1.04	0.158	<0.001	(0.71, 1.37)	0.391
	TSS	-0.31	1.01	0.165	<0.001	(0.67, 1.36)	0.425
20%	QR	-0.24	1.80	0.401	<0.001	(1.48, 2.12)	0.542
	TS	-0.25	1.66	0.194	<0.001	(1.26, 2.06)	0.509
	TSS	-0.14	1.64	0.192	<0.001	(1.24, 2.04)	0.546
30%	QR	0.29	1.76	0.212	<0.001	(1.58, 1.95)	0.479
	TS	0.26	1.52	0.112	<0.001	(1.29, 1.75)	0.693
	TSS	0.16	1.44	0.116	<0.001	(1.20, 1.68)	0.645
50%	QR	0.53	1.54	0.233	<0.001	(1.18, 1.90)	0.951
	TS	0.56	1.50	0.148	<0.001	(1.20, 1.81)	0.946
	TSS	0.35	1.23	0.172	<0.001	(0.88, 1.58)	0.977

Table 4
Results of Relative Median Absolute Error of the four Regression Procedures with Outliers of 10%, 20%, 30% and 50% of n=20, in both X and Y Direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80

Outliers	Relative Median Absolute Error	Value
10%	QR vs TS	0.036
	QR vs TSS	-0.048
	TS vs TSS	-0.086
20%	QR vs TS	0.059
	QR vs TSS	0.007
	TS vs TSS	-0.071
30%	QR vs TS	-0.446
	QR vs TSS	-0.335
	TS vs TSS	0.070
50%	QR vs TS	0.005
	QR vs TSS	-0.027
	TS vs TSS	-0.032

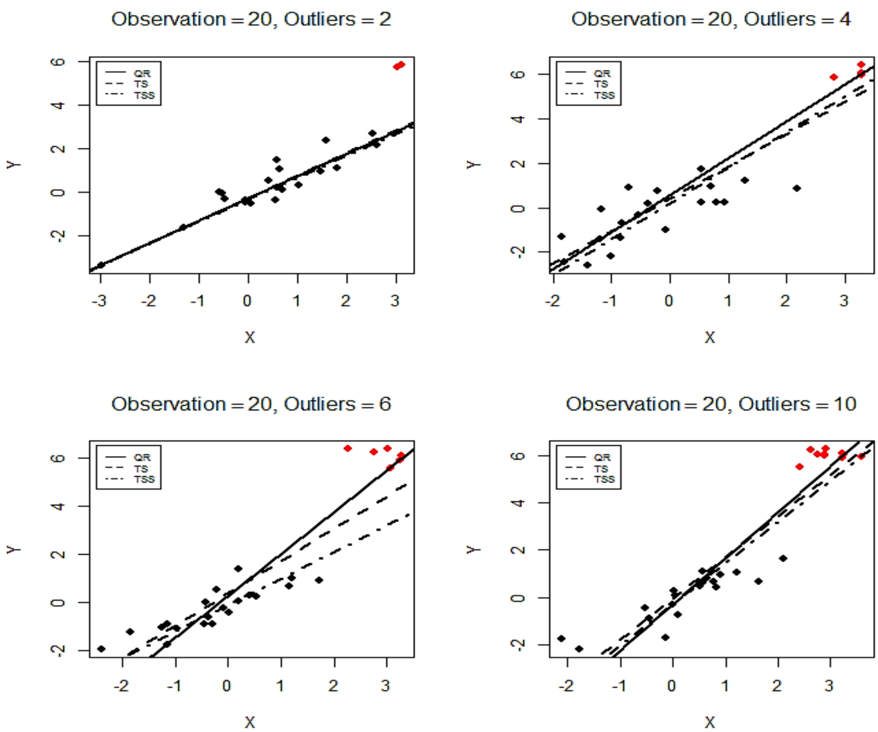


Figure 2: Four regression lines are shown in each plot with n=20 and outliers of 10%, 20%, 30% and 50% of n=10 in both X and Y direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80:

Table 5
Results from the three regression procedures with outliers of 10%, 20%, 30% and 50% of n=30, in both X and Y direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80:

Outliers	Regression	$\hat{\beta}_0$	$\hat{\beta}_1$	S.E($\hat{\beta}_1$)	P-value	95% CI for β_1	MEDAE
10%	QR	0.27	1.07	0.448	0.023	(0.66, 1.49)	0.601
	TS	0.35	0.95	0.157	<0.001	(0.62, 1.27)	0.544
	TSS	0.07	0.93	0.167	<0.001	(0.59, 1.28)	0.543
20%	QR	0.24	1.26	0.345	<0.001	(0.95, 1.58)	0.501
	TS	0.23	1.13	0.144	<0.001	(0.84, 1.42)	0.472
	TSS	0.06	1.08	0.156	<0.001	(0.77, 1.40)	0.462
30%	QR	0.31	1.79	0.197	<0.001	(1.58, 2.01)	0.637
	TS	0.38	1.59	0.110	<0.001	(1.37, 1.82)	0.764
	TSS	0.09	1.52	0.115	<0.001	(1.28, 1.75)	0.823
50%	QR	0.46	1.74	0.139	<0.001	(1.41, 2.06)	1.000
	TS	0.52	1.62	0.102	<0.001	(1.41, 1.83)	0.751
	TSS	0.52	1.68	0.102	<0.001	(1.47, 1.88)	0.626

Table 6
Results of Relative Median Absolute Error of the four regression procedures with outliers of 10%, 20%, 30% and 50% of n=30, in both X and Y direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80:

Outliers	Relative Median Absolute Error	Value
10%	QR vs TS	0.094
	QR vs TSS	0.110
	TS vs TSS	0.018
20%	QR vs TS	0.057
	QR vs TSS	0.077
	TS vs TSS	0.022
30%	QR vs TS	-0.200
	QR vs TSS	-0.292
	TS vs TSS	-0.077
50%	QR vs TS	-0.113
	QR vs TSS	0.072
	TS vs TSS	0.166

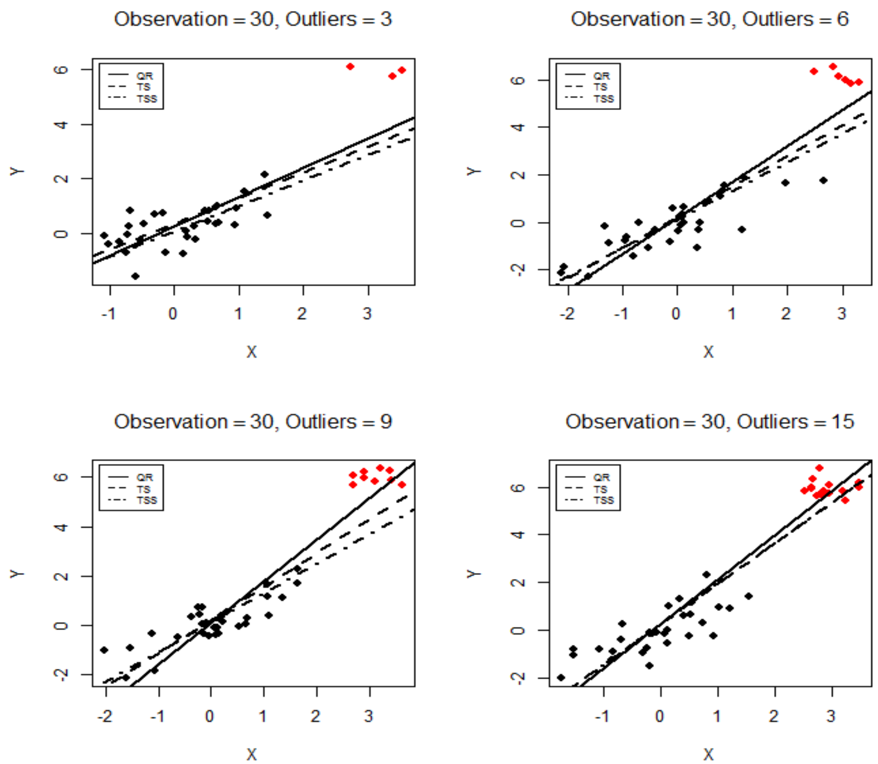


Figure 3: Four Regression Lines are shown in each Plot with n=30 and Outliers of 10%, 20%, 30% and 50% of n=30 in both X and Y Direction with Nsim=1000, X~ Normal (n, 0, 1), Y~ Normal (n, 0, 1), and Cor (X, Y) = 0.80

5. DISCUSSION

As stated, Monte Carlo techniques were used to estimate the regression coefficients, standard errors, p-values, confidence intervals and median absolute deviation based Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression procedures. Visual as well as numerical comparisons were made using these three regression procedures. For visual comparison scatter plots with fitted regression lines using all three regression procedures were used. For numerical compression, standard errors, confidence intervals, median absolute error (MEDAE), and Relative Median Absolute Error were used. The results from the simulation study, compiled were indicated in the Tables 1 to 6 and based on n=10, 20, and 30 with 1,000 simulations. The performance of Quantile Regression, Theil Sen Regression, and Theil Sen Siegel Regression lines are also presented for visual comparison with scatter plots and fitted regression lines. If the errors (e_i) were independent and normally, then a random sample of size n was generated from a bivariate normal distribution with mean (0, 0) and variances equal to 1, and a correlation coefficient equal to 0.80. Outliers of 10%, 30% 50% and 100% of n were introduced in both X and Y variables from a bivariate normal distribution with means (2, 6) and variances 0.1*variance of the above bivariate normal distribution, i.e. the variances (0.1, 0.1). When $n = 10$, it can be seen from Table 1, all the slop estimates were positive and significant with an overall slight increase in standard errors. The regression coefficient $\hat{\beta}_1$ has an overall minimum MEDAE in TSS as compared to QR and TS regressions. The relative performance with median absolute deviation in Table 2 is consistent with Table 1. The four regression lines with 10%, 20%, 30% and 50% outliers of $n = 10$ can be seen at each plot in Figure 1. It can be seen, TSS were more robust regression lines as compared to QR and TS regressions. In Table 2, when $n = 20$, the performance of all the regression procedures were same with 10% increase of outliers whereas, an overall performance of TSS was better as compare to QR and TS regressions which can also be seen from Figure 2. Finally, when $n = 30$, it can be seen from Table 5, all the slop estimates were remains positive and significant with an overall slight increase in standard errors. The regression coefficient $\hat{\beta}_1$ has an overall minimum MEDAE in TSS as compared to QR and TS regressions.

6. CONCLUSION AND RECOMMENDATIONS

When the regression model is being used with outliers in both X and Y directions, it can be seen TSS has an overall small standard error as compared to QR and TS. All figures illustrated in regression procedures with outliers in both X and Y directions that TSS was more robust regression procedure compared to QR and TS. Therefore, it is recommended that, under the normality assumption with outliers in both X and Y direction TSS should be the most suitable followed by QR and TS. For future study, all three regression techniques can be compared with the presence of outliers in X and Y directions separately. Similarly, all three regression techniques can be compared with the presence of outliers in large samples. It is also recommended to study the performance of these techniques with multiple linear regressions.

REFERENCES

1. Chaudhary, S.M. and Kamal, S. (2000). *Introduction to Statistical Theory* (Part II): Lahore, Punjab: Ilmi Katab Khana.
2. Efron, B. (1987). Better bootstrap confidence intervals. *Journal of the American Statistical Association*, 82, 171-185.
3. Efron, B. and Tibshirani R.J. (1993). *An Introduction to the Bootstrap*, London: Chapman & Hall.
4. Edgeworth, F.Y. (1887). On Observations Relating to Several Quantities, *Hermathena*, 6, 279-285.
5. Grubbs, F. (1969). Procedures for detecting outlying observations in samples. *Technimetrics*, 11(1), 1-21.
6. Hampel, F.R., Ronchetti, E.M., Rousseeuw, P.J. and Stahel, W.A. (1986). *Robust Statistics*. New York: Wiley.
7. Helsel D.R. and Hirsch, R.M. (2002). *Statistical methods in water resources: Hydrologic analysis and interpretation: Techniques of Water-Resources Investigations of the U.S. Geological Survey*, Chap. A3, Book 4, 510 p.
8. Hao, L. and Naiman, D.Q. (2007). *Regression procedure*. Thousand Oaks: Sage Publications Inc.

9. Hoeffding, W. (1048). A class of statistics with asymptotically normal distribution. *Ann. Math. Statist.*, 19, 293-325.
10. Hollander, M. and Wolfe, D. (1999). *Nonparametric statistical methods*: New York: Wiley.
11. Huber, P.J. and Ronchetti, E. (2009). *Robust Statistics*, 2nd Ed. New York: Wiley.
12. Hussain, S.S. and Sprent, P. (1983). Nonparametric regression. *Journal of the Royal Statistical Society*, Series A, 146, 182-191.
13. Koenker, R. and Bassett, G.W. (1978). Regression Quantiles. *Econometrica*, 46, 33-50.
14. Koenker R. and Bassett GJ. (1982). Robust tests for heteroscedasticity based on regression quantiles. *Econometrica*, 50, 43-61.
15. Koenker, R. and Bassett, G.W. (1982b). Test for linear hypothesis and L1 estimation: *Econometrica*, 50, 1577-1583.
16. Koenker, R. and Machado, J.A.F. (1999). A goodness of fit and related inference processes for regression procedure. *Journal of the American Statistical Association*, 94, 1296-1310.
17. Maronna, R.A., Martin, D.R. and Yohai, V.J. (2006). *Robust Statistics: Theory and Methods*. New York: Wiley.
18. Greenacre, M. and Ayhan, O. (2015). Identifying Inliers. *Barcelona GSE Working Paper Series* Working Papers 763.
19. Mosteller, F. and Tukey, J.W. (1977). *Data Analysis and Regression: A Second Course in Statistics*. Reading, MA: Addison-Wesley Publishing Company.
20. Nelson, D. (1998). *Dictionary of Mathematics*: London, England: Penguin.
21. Petscher, Y., Logan, J.A. and Zhou, C. (2013). Extending conditional means modeling: An introduction to quantile regression. In *Applied quantitative analysis in education and the social sciences* (pp. 15-45). Routledge.
22. Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall’s tau. *J. Amer. Statist. Assoc.*, 63, 1379-1389.
23. Siegel, A.F. (1982). Robust Regression Using Repeated Medians. *Biometrika*, 69, 242-244.
24. Shah, S.H., Rehman, A., Rashid, T., Karim, J. and Shah, S. (2016). A comparative study of ordinary least squares regression and Theil-Sen regression through simulation in the presence of outliers. *J. Sci. Technol.*, V, 137-142.
25. Theil, H. (1950). A rank invariant method for linear and polynomial regression analysis: *Nederl. Akad. Wetensch. Proc. Ser. A*, 53, 386-392 (Part I), 521-525 (Part II), 1397-1412 (Part III).
26. Wilcox, R.R. (1998). A note on the Theil-Sen regression estimator when the regressor is random and the error term is heteroscedastic. *Biometrical Journal*, 3, 261-268.
27. Wilcox, R.R. (2012b). *Introduction to Robust Estimation and Hypothesis Testing*, 3rd Edition. San Diego, CA: Academic Press.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 15-26

CONSTRUCTION OF EWMA CONTROL CHART USING MEASURES OF DISPERSION AS ESTIMATE

Liaquat Ahmad[1], **Muhammad Aslam**[2] and **Qurrat-ul-Ain Naeem**[1]

[1] Department of Statistics and Computer Science
University of Veterinary and Animal Sciences, Lahore, Pakistan
Email: liaquatahmad@uvas.edu.pk, anniequrratulann@gmail.com

[2] Department of Statistics, Faculty of Science
King Abdulaziz University, Jeddah, Saudi Arabia
Email: aslam_ravian@hotmail.com

ABSTRACT

In this research paper Exponentially Weighted Moving Average (EWMA) control chart using Measures of dispersion has been presented. EWMA technique has widely been accepted because of its capability for quick detecting of small variations of process shifts, which are expectedly lie within certain limits. We proposed an EWMA control chart using measures of dispersion as an estimate for quick and early monitoring of small variations in five blood sample parameters of human beings data. It has been observed that EWMA scheme is efficient for monitoring small variations of blood sample parameters. The comparison figures of Measures of dispersions (Variance and Standard deviation) and EWMA by using R-Crane software have been shown for all five blood parameters.

KEYWORDS

Blood samples, EWMA Chart, Measures of dispersion, Small variations.

1. INTRODUCTION

Blood tests are very important regarding the health of patients to check the factors that are involving leading to abnormalities and diagnosis of disease so that possible treatment should have taken. Complete blood count of cells of human body is an important test in medical research. It is very common test mostly physicians prefer this test to detect variety of reasons that are responsible for disturbing the immune system based on a patient's symptoms. This test may also be a part of routine check-up. A complete blood examination of cells includes count of red blood cells, white blood cells, count of platelets, infection in blood. In this research our focal point is to locate the variations that effect the health of patients and comparison with Measures of dispersion and shown which measure is good to get quick or early detection or small shifts of variations by using R-crane software. In this research five blood parameters of human beings have been collected from laboratory are Hemoglobin (HB), Platelets (PLT), WBCs (LTC), Glucose (Blood sugar), ESR. Blood culture is a common and essential procedure utilized in research in advance medical science for human beings, birds or animals and plays an important role in reducing the risk factors. This system was developed to assist researchers with blood sample collection while reducing risk, pain and distress. A blood culture is a procedure in which a blood sample has been taken to test the outside invaders like bacteria, yeast, infection and other microorganisms in blood in modern medicine. The basic purpose of blood culture or blood sampling is to diagnose a factor for the treatment of a patient. The results of the blood test can help physician determine which specific disease, bacteria or yeast is responsible for causing the blood infection and results have been used how best to overcome it. The factors have been diagnosed that are causing the infection in bloodstream of patient and interfere with patient body's normal defense system and prevent their immune system from working perfectly. Blood culture or screening is very important for any kind of treatment. In another words we can say that the blood is a window into the health of the body. An early diagnosis means you can take a start for treatment and take steps for healthy life. Blood screening is become much easier after the inventions of blood screening devices or machines that is called blood analyzer device, it counts and identifies the shape or size of blood cells. Annual blood screening is the most important step that can take to prevent life-threatening disease. In recent years, diabetes/ blood sugar has become one of the most leading causes of deaths worldwide. From last few decades diabetes has become a chronic disease. According to the World Health Organization (WHO), around 1.6 million people worldwide died due to diabetes in 2016. It is estimated that 425 million people are living with diabetes all over the world. In 2014, 8.5% of adults aged 18 years and older had diabetes. In 2015, diabetes was the direct cause of 1.6 million deaths and in 2012 high blood glucose was the cause of another 2.2 million deaths. In recent years, diabetes has become one of the leading causes of deaths worldwide. Through blood sugar it has become easier to check level of sugar anywhere anytime. Control charts are graphical tools used to study process variation or monitoring process stability or control of a process. Control charts are used for time-series or historical data. Point representing a statistic for example (range, variance, standard deviation, mean deviation, quartile deviation) measures a quality characteristic in samples taken from the process at different times. The mean of this statistic using all the samples is then calculated. A center line is drawn for value of the mean of the statistic corresponding to the in-control state. Two other horizontal lines called upper control limit (UCL) and lower control limit (LCL) of the statistic (data) are located on the chart. These control limits are chosen so that if the process is in control, nearly all of the samples falls in between the UCL& LCL. If maximum points plot within the limits then the process assumed to be in control. However if the points plot outside the control limits then the process will be out of control (Montgomery, 2009). The control charts, also known as Shewhart charts were introduced by Walter A. Shewhart in 1920. It is the most widely used chart for monitoring the mean of a process (Castagliola, Achouri, Taleb, Celano, & Psarakis, 2013). Shewhart chart is simpler in comparison to other charts (L. Ahmad, Sabqa, Khan., Aslam M. 2018), (L. Ahmad, Aslam, & Jun, 2014). Roberts (1959) introduced the Exponentially Weighted Moving Average (EWMA) control scheme. Using simulation to evaluate its properties, he showed that the EWMA is useful for detecting small shifts in the mean of a process. The EWMA is often applied to a time-ordered sequence of random variables. It computes a weighted average of the sequence by applying weights that decrease geometrically with the age of the observations. For the EWMA control technique, the decision depends on the EWMA statistic, which is an exponentially weighted average of all prior data, including the most recent measurement. The EWMA control chart has been studied by many researchers including (L. Ahmad, Aslam, Khan,

& Jun, 2017; L. Ahmad, Rafiq, I., Aldosari, M. S., and Aslam, M. , 2017; Aslam, Ali Raza, Azam, Ahmad, & Jun, 2019; Aslam, Azam, & Jun, 2014a, 2014b; Aslam, Balamurali, Jun, & Meer, 2016; Aslam, Khan, Aldosari, & Jun, 2016; Aslam, Saghir, Ahmad, Jun, & Hussain, 2017a, 2017b; Khan, Aslam, & Jun, 2016; Saghir, Ahmad, Aslam, & Jun, 2018). Only recently the EWMA control scheme has been exploited and its properties evaluated analytically (S. Crowder & Hamilton, 1992; S. V. Crowder, 1989; Hunter, 1986; Lucas & Saccucci, 1990). The rest of the paper is organized as: the methodology of the proposed chart has been discussed in Section 2. In Section 3 results and discussion have been given. Conclusion is given in the last Section.

2. METHODOLOGY

The EWMA chart is famous because of another characteristic. As Montgomery (2009) states: "It is almost a perfectly non parametric (distribution free) procedure". Given a series of observations and a fixed weight, the first element of EWMA is computed as:

$$(1 - \text{weight}) \, * \, \text{Previous EWMA} \, + \, (\text{weight} \, * \, \text{current observation})$$
$$\mathrm{w_i} = \lambda \mathrm{x_i} + (1-\lambda)\mathrm{w_{i-1}}$$

The current observation is modified by "shifting forward "and repeating the calculation. This process is repeated over the entire series creating the EWMA statistic. It weights λ where $0 \leq \lambda \leq 1$ and the observations in geometrically decreasing order so that the most recent observation contributes highly while the oldest observation contribute very little.

This research paper is based upon EWMA control charts by using estimates of measures of dispersion. We have collected two months blood samples data of human beings from laboratory. After arranging the data EWMA control charts have been constructed by using R-Language. Data have been collected for two months (January and February) on daily basis for the year 2018, which is used to test five parameters of blood samples to construct EWMA control charts. The main task of this research paper was to study EWMA control charts capabilities to detect small shifts of the considered processes and compare with measures of dispersion to show which measure is good in early detection of abnormalities of blood parameters.

Construction of Dispersion Charts

Total 59 observations have been collected, 31 observations for the month of January, 2018 and 28 observations for the month of February, 2018 in which each sample or subgroups consist of five blood count of size $n = 5$ human body.

Let $\sigma_1^2, \sigma_2^2, \sigma_3^2 \ldots . \sigma_m^2$ be the variance of samples "m", the formula to calculate variance is as under:

$$\sigma^2 = \frac{\sum_{i=1}^{n}(xi - xbar)^2}{n-1}$$

and the average variance for samples "m" become,

$$\bar{\sigma}^2 = \frac{\sum \sigma^2{}_m}{m}$$

The control limits for variance

$$UCL = \bar{\sigma}^{2} * B4$$

where, $B4 = 1 + \frac{3}{C_4}\sqrt{1} - C_4$, ($C_4 = 0.9400$)

Central line is $\bar{\sigma}^2$

$$LCL = \bar{\sigma}^{2} * B3$$

where, $B3 = 1 - \frac{3}{C_4}\sqrt{1} - C_4$, ($C_4 = 0.9400$)

Standard Deviation Control Chart

Let $S_1, S_2, S_3, \dots, S_m$ be the Standard deviation of samples "m", the formula to calculate standard deviation is as under:

$$s = \frac{\sqrt{(\sum_{i=1}^{n}(xi - xbar)^2}}{n-1}$$

and the average standard deviation

$$\bar{s} = \frac{\sum_{i=1}^{n} s_i}{m}$$

The Control Limits for Variance

$$UCL = \bar{s} * B4$$

Central line is $\bar{s}$

$$LCL = \bar{s} * B3$$

$B3$ & $B4$ are constants, where $B3 = -0.0886$ & $B4 = 2.0886$

EWMA-Measures of Dispersion Chart

Here we will apply EWMA scheme by using estimates of measures of dispersion (Variance and Standard deviation).

EWMA-Variance Chart

We put variance in EWMA statistic to construct control chart as given above:

$$Y_t = \lambda s_t^2 + (1 - \lambda)\, Y_{t-1}$$

Control limits for EWMA-Variance Chart

$$UCL = \bar{s}^2 + k * \sigma * \sqrt{\frac{\lambda}{2-\lambda}}$$

Central line is $\bar{s}^2$

$$LCL = \bar{s}^2 - k * \sigma * \sqrt{\frac{\lambda}{2-\lambda}}$$

EWMA-Standard Deviation Control Chart

We put standard deviation in EWMA statistic to construct control chart as given above:

$$Y_t = \lambda s_t + (1 - \lambda)\, Y_{t-1}$$

Control Limits for Standard Deviation

$$UCL = \bar{s} + k * \sigma * \sqrt{\frac{\lambda}{2-\lambda}}$$

Central line is $\bar{s}$

$$UCL = \bar{s} - k * \sigma * \sqrt{\frac{\lambda}{2-\lambda}}$$

3. RESULTS AND DISCUSSIONS

In this section, the results and discussion have been described based on the control charts constructed by using the above mentioned methodology.

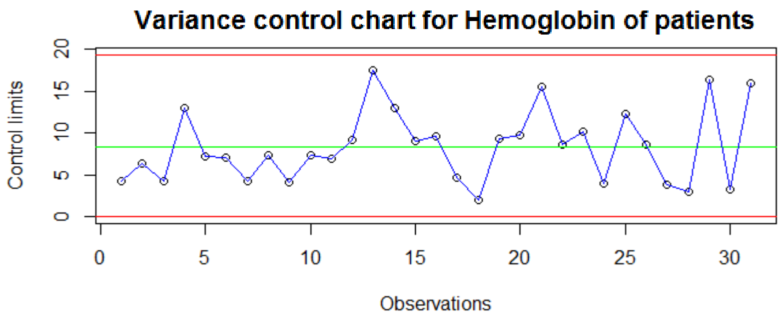


Figure 1: Variance Chart for Hemoglobin data for January, 2018

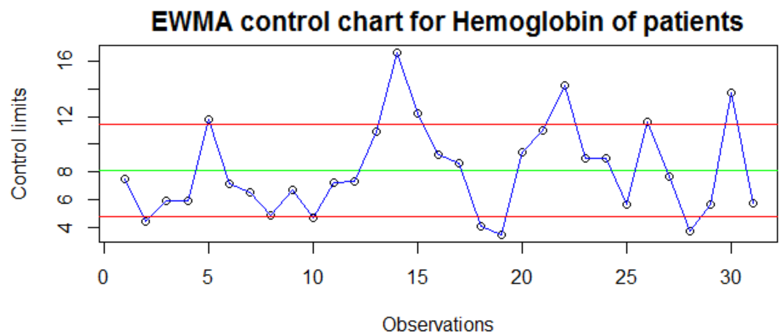


Figure 2: EWMA-variance Chart for Hemoglobin data for January, 2018

Figure 1 shows variance chart its UCL =19.38& LCL = 0, it shows no out of control values and Figure 2 shows EWMA-variance chart with parameters λ = 0.2 and k = 3 it's UCL = 11.43581 & LCL = 4.795806 detects the 9 out-of-control value.

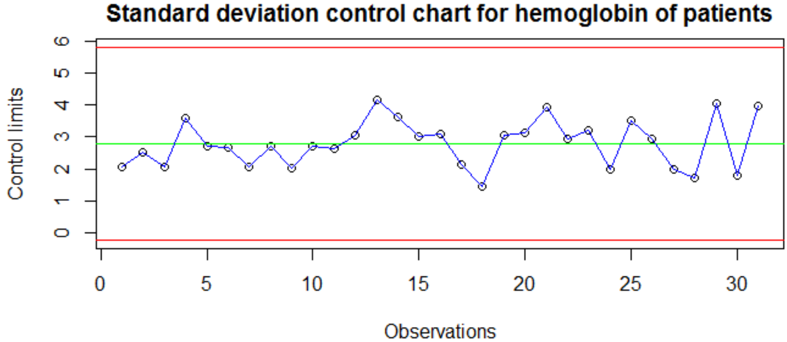


Figure 3: Standard deviation chart for Hemoglobin data for January, 2018

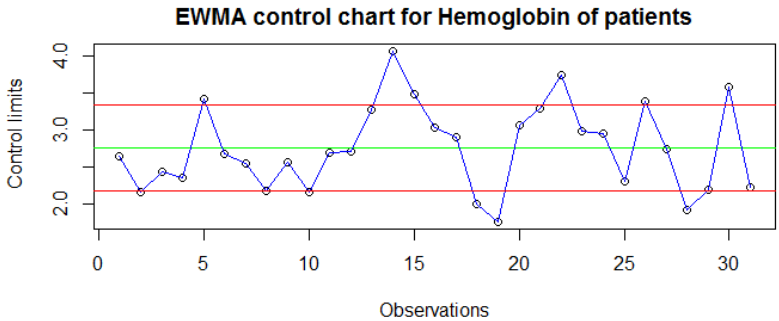


Figure 4: EWMA-standard deviation chart for Hemoglobin data for January, 2018

Figure 3 shows standard deviation chart it's UCL = 5.825173 & LCL = -0.2471083, it shows no out of control value and Figure 4 shows EWMA-standard deviation chart with parameters λ = 0.2 and k = 3 it's UCL = 3.337742 & LCL = 2.177742 detects the 9 out-of-control value.

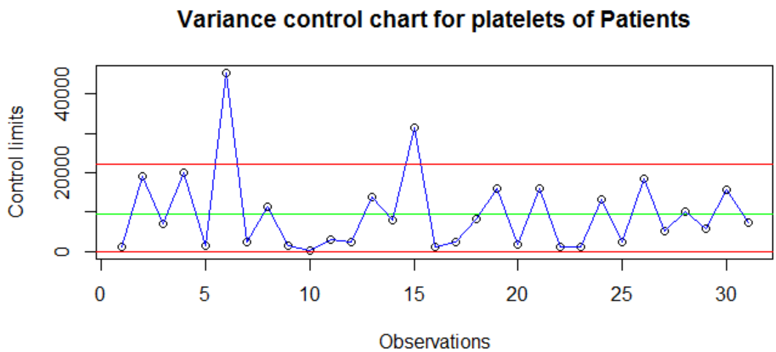


Figure 5: Variance chart for Platelets data for January, 2018

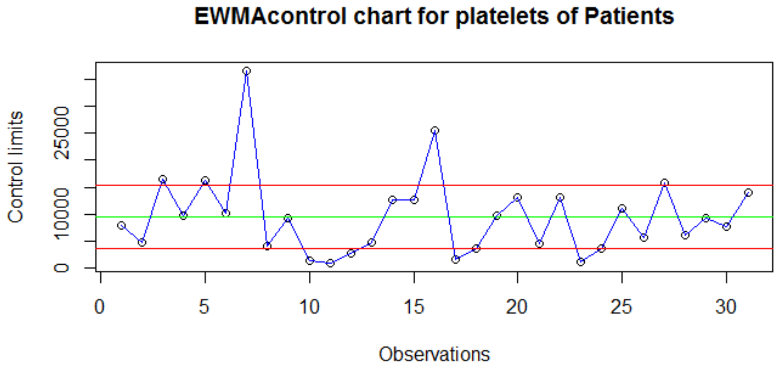


Figure 6: EWMA-variance for Platelets data for January, 2018

Figure 5 shows Variance chart its UCL = 22120.07 & LCL = 0, it shows 2 out of control values and figure 6 shows EWMA-variance chart in which it shows the posting of the observations based upon platelets of patients with parameters $\lambda = 0.2$ and $k = 3$ it's UCL = 15426.38 & LCL = 3665.893 detects the 7 out-of-control values.

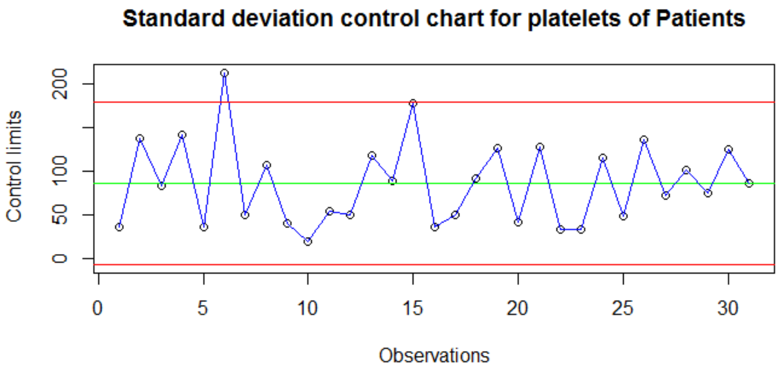


Figure 7: Standard deviation chart for Platelets data for January, 2018

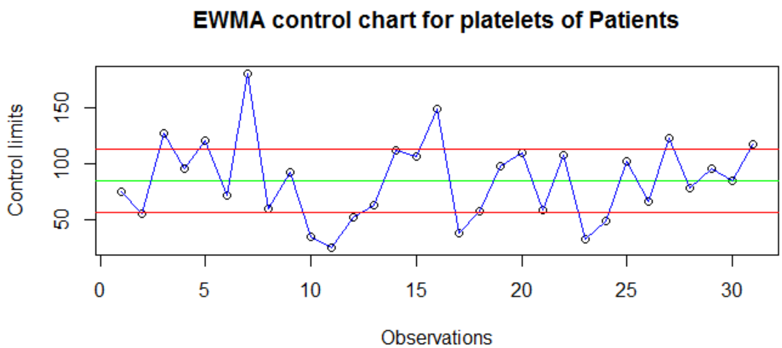


Figure 8: EWMA-standard deviation chart for Platelets data for January, 2018

Figure 7 shows Standard deviation chart its UCL =178.2674 & LCL = -7.5622, it shows 1 value out of control and Figure 8 shows EWMA- standard deviation chart in which it shows the posting of the observations based upon platelets of patients with parameters λ = 0.2 and k = 3 it's UCL = 113.2703 & LCL = 57.38213 detects the 10 out-of-control values.

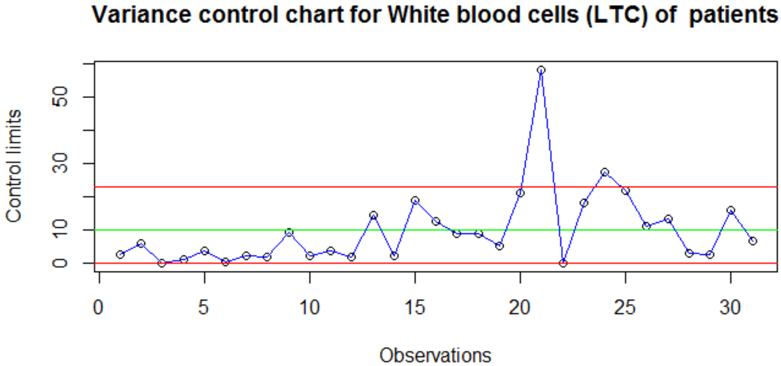


Figure 9: Variance chart for LTC data for January, 2018

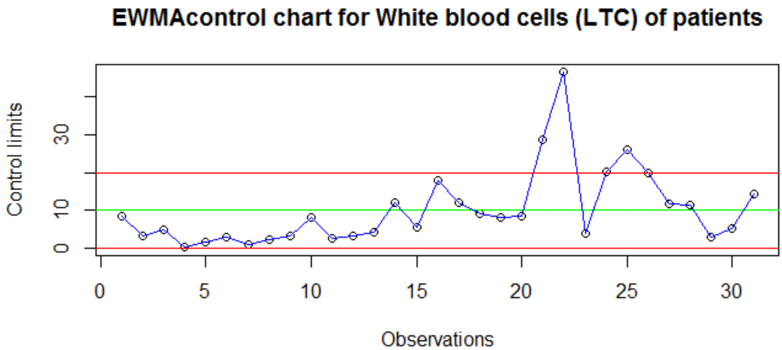


Figure 10: EWMA-Variance chart for LTC data for January, 2018

Figure 9 shows Variance chart its UCL = 23.16471 & LCL = 0, it shows 1 out of control value and figure 10 shows EWMA-variance Chart in which it shows the posting of the observations based upon LTC level of patients with parameters λ = 0.2 and k = 3 it's UCL = 19.98806 & LCL = 0.06806452 detects 3 out-of-control values.

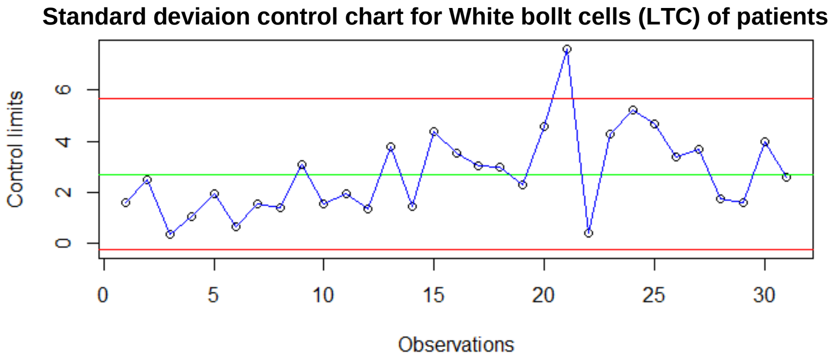


Figure 11: Standard deviation chart for LTC data for January, 2018

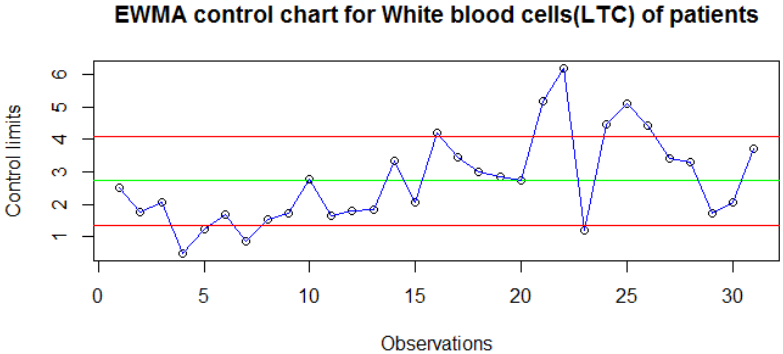


Figure 12: EWMA-Standard deviation chart for LTC data for January, 2018

Figure 11 shows Standard deviation chart its UCL = 5.680992 & LCL = -0.240992, it shows 1 out of control value and figure 12 shows EWMA-Standard deviation Chart in which it shows the posting of the observations based upon LTC level of patients with parameters $\lambda = 0.2$ and $k = 3$ it's UCL = 4.105806& LCL = 1.345806 detects 10 out-of-control values.

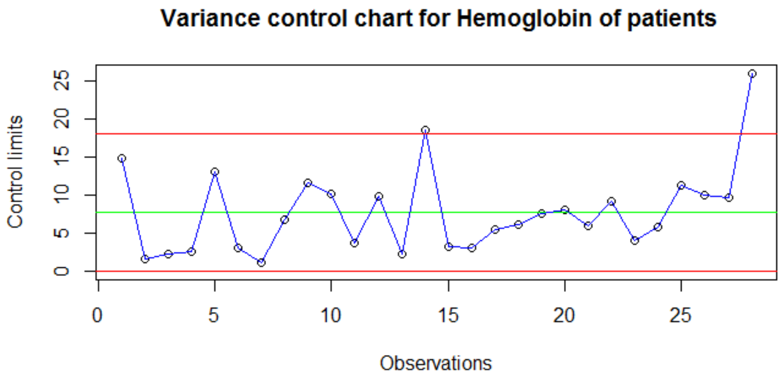


Figure 13: Variance chart for hemoglobin data for February, 2018

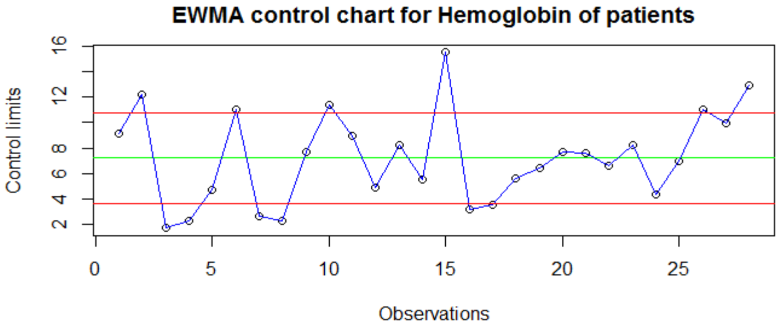


Figure 14: EWMA-variance chart for hemoglobin data for February, 2018

Figure 13 shows Variance chart its UCL= 18.04169 & LCL= 0, it shows 2 out of control values and figure 14 shows EWMA-variance Chart in which it shows the posting of the observations based upon Hemoglobin level of patients with parameters $\lambda = 0.2$ and $k = 3$ it's UCL = 10.80107 & LCL = 3.641071 detects 12 out-of-control value.

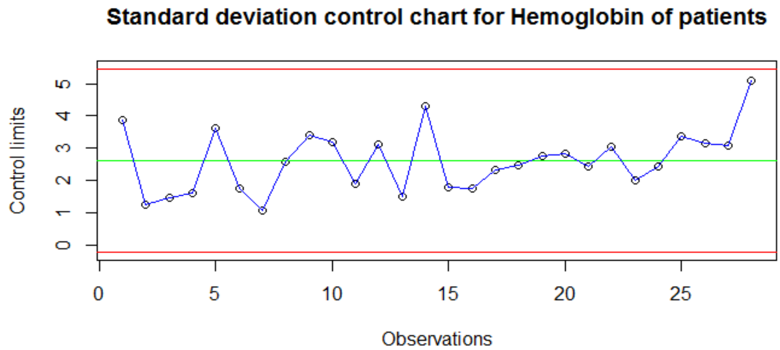


Figure 15: Standard deviation chart for hemoglobin data for February, 2018

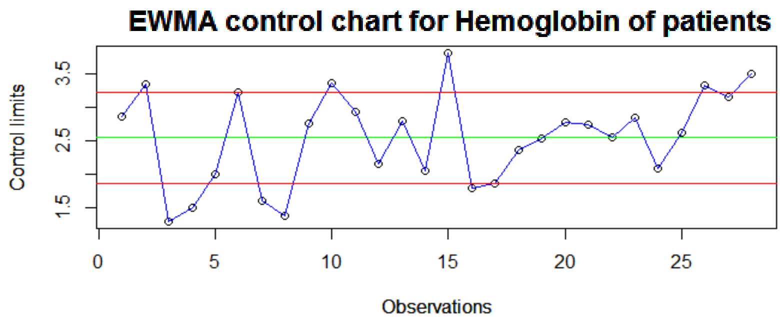


Figure 16: EWMA-standard deviation chart for hemoglobin data for February, 2018

Figure 15 shows Standard deviation chart its UCL = 5.465419& LCL= -0.2266136, it shows no out of control value and figure 16 shows EWMA-Standard dev chart in which it shows the posting of the observations based upon Hemoglobin level of patients with parameters $\lambda = 0.2$ and $k = 3$ it's UCL = 3.226786 & LCL = 1.866786 detects 12 out-of-control values.

CONCLUSIONS

Blood culture plays an important role in the detection of disease. Doctors referred it as CBC it stands for Count blood Cells. It's a count of cells in a body to determine the abnormalities present in cells. It has become very common test that is usually suggested by physicians. In this research we have collected blood parameters and construct EWMA Control charts by using measures of dispersion as an estimate to study which measure is efficient to show the significant variations. We have used R-language to plot data. Dispersion charts and EWMA-Dispersion charts have been constructed separately to show that simple measures of dispersion charts don't show out-of-control values but EWMA technique is efficient in early detection of out-of-control values. Standard deviation and Variance Control charts have shown us they are less efficient in detection of variation of process because mostly 1 or 2 values are out-of-control but it can be seen clearly that EWMA-Dispersion charts show us significant variation and detect more out-of-control values than simple dispersion charts. And it has been clearly shown in graphs that standard deviation and variance are good measures in detection of small variations for this type of data. So, in this research paper we have shown figures of control charts that give the evidence that EWMA scheme is more efficient that can be used for this type of health related data due to its quick detection of variations.

REFERENCES

1. Ahmad, L., Aslam, M. and Jun, C.H. (2014). Coal Quality Monitoring With Improved Control Charts. *European Journal of Scientific Research*, 125(2), 427-434.
2. Ahmad, L., Aslam, M., Khan, N. and Jun, C.H. (2017). *Double moving average control chart for exponential distributed life using EWMA.* Paper presented at the AIP Conference Proceedings.
3. Ahmad, L., Rafiq, I., Aldosari, M.S. and Aslam, M. (2017). Mixed Repetitive Sampling Plan using EWMA. *Advances and Applications in Statistics*, 51(3), 167-186. doi: http://dx.doi.org/10.17654/AS051030167
4. Ahmad, L., Sabqa, Khan., Aslam M. (2018). *Monitoring of Blood parameters of Dogs and Cats Using Control Chart Techniques.* Paper presented at the 11 International Statistics Days Conference (ISDC), Kefaluka Resort Hotel, Bordum, Turkey.
5. Aslam, M., Ali Raza, M., Azam, M., Ahmad, L. and Jun, C.H. (2019). Design of a sign chart using a new EWMA statistic. *Communications in Statistics-Theory and Methods*, 1-12.
6. Aslam, M., Azam, M. and Jun, C.H. (2014a). A new control chart for exponential distributed life using EWMA. *Transactions of the Institute of Measurement and Control*. doi: 10.1177/0142331214537293
7. Aslam, M., Azam, M. and Jun, C.H. (2014b). A new exponentially weighted moving average sign chart using repetitive sampling. *Journal of Process Control*, 24(7), 1149-1153.

8. Aslam, M., Balamurali, S., Jun, C.H. and Meer, A. (2016). Time-truncated attribute sampling plans using EWMA for Weibull and Burr type X distributions. *Communications in Statistics-Simulation and Computation*, 1-12.
9. Aslam, M., Khan, N., Aldosari, M.S. and Jun, C.H. (2016). Mixed Control Charts Using EWMA Statistics. *IEEE Access,* 4, 8286-8293.
10. Aslam, M., Saghir, A., Ahmad, L., Jun, C.H. and Hussain, J. (2017a). A control chart for COM-Poisson distribution using a modified EWMA statistic. *Journal of Statistical Computation and Simulation*, 87(18), 3491-3502.
11. Aslam, M., Saghir, A., Ahmad, L., Jun, C.H. and Hussain, J. (2017b). A control chart for COM-Poisson distribution using a modified EWMA statistic. *Journal of Statistical Computation and Simulation*, 1-12. doi: 10.1080/00949655.2017.1373114
12. Castagliola, P., Achouri, A., Taleb, H., Celano, G. and Psarakis, S. (2013). Monitoring the coefficient of variation using a variable sampling interval control chart. *Quality and Reliability Engineering International*, 29(8), 1135-1149.
13. Crowder, S. and Hamilton, M. (1992). An EWMA for monitoring a process standard deviation. *Journal of Quality Technology,* 24(1), 12-21.
14. Crowder, S.V. (1989). Design of Exponentially Weighted Moving Average Schemes. *Journal of Quality Technology,* 21(3), 155-162.
15. Hunter, J.S. (1986). The exponentially weighted moving average. *J. Quality Technol.,* 18(4), 203-210.
16. Khan, N., Aslam, M. and Jun, C.H. (2017). Design of a control chart using a modified EWMA statistic. *Quality and Reliability Engineering International*, 33(5), 1095-1104.
17. Lucas, J.M. and Saccucci, M.S. (1990). Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements. *Technometrics,* 32(1), 1-12.
18. Montgomery, C.D. (2009). *Introduction to Statistical Quality Control* (6th ed.). New York: John Wiley & Sons, Inc.
19. Roberts, S.W. (1959). Control Chart Tests Based on Geometric Moving Averages. *Technometrics,* 1(3), 239-250.
20. Saghir, A., Ahmad, L., Aslam, M. and Jun, C.H. (2018). A EWMA control chart based on an auxiliary variable and repetitive sampling for monitoring process location. *Communications in Statistics-Simulation and Computation*, 1-12. doi: doi.org/10. 1080/03610918.2018.1433837

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 27-48

INSTITUTIONS AND ECONOMIC GROWTH IN SELECTED DEVELOPING COUNTRIES: BAYESIAN ECONOMETRIC ANALYSIS

Qasim Shah, Muhammad Akbar Awan and **Mohammad Anwar**
Department of Mathematics and Statistics, International Islamic University, Islamabad, Pakistan.
Email: anwarmohammad.uos@gmail.com
qasimshah707@gmail.com
mmakbr@gmail.com

ABSTRACT

This study is investigated panel data regression models. In this paper, Classical and Bayesian approach is to examined the effects of institutions on economic growth of selected developing countries from (1990-2014). Under the classical and Bayesian methodology, the two panel data models is estimated which are common effect, fixed effect. For the Bayesian approach the prior information’s are used in this paper and normal gamma prior are used for the panel data models. The analysis was done through WinBUGS14 software. The estimated results of the study showed that panel data models are valid models in Bayesian methodology. The results of all independent variables in Bayesian approach is positively and significantly effect on dependent variable. Based on standard errors of all models, we must say that fixed effect model is the best model in Bayesian estimation of panel data models. Also, it was proved that the fixed effect model has the lowest value of standard error, as compared to other models.

KEYWORDS

Bayesian estimation, Gamma prior, Fixed effect, Common effect.

1. INTRODUCTION

For the period of the last twenty-five years, developing countries have faced very unusual growth performances. Differences among different countries are not justified due to a difference in the starting levels of GDP per capita. According to economic bulletin ECB (2015), the worth of domestic and political institutions and governance of government has a significant and positive effect on income growth per capita and deficiency of actual convergence is related to several factors, notably weak institutions, structural rigidities, weak productivity growth and insufficient policies to address asset price booms.

The economic institutions, they play a "key" role in this variation of development & growth between different nations. Besides, it also influences the use of new technologies in the process of production. Economic institutions are vital as they have an impact on the economic structure of an economy. Also, they help in the allocation of resources through production factor (Acemoglu & Robinson; 2004).

In describing the economic growth process, institutions play a very important role by disturbing the structure of investment in human and physical capital as well as technological advancement and innovations. It is assumed that security of property rights plays a noteworthy role for estimation of the long-term economic increase or growth. It is believed that insecurity of property rights plays an inverse relationship with economic growth. Mauro (1995) studied the association between economic growth and institutions and concluded that corruption has a strong negative effect on economic growth.

Political & economic growth also depends upon the institutions. Neoclassical assumption describes the fact that where there will be the availability of beneficial opportunities, growth will occur. Violence hinders the process of growth & development. It is one of those evils, which our societies are facing today & this is especially aggravated in the countries which is developing as people spread/promote violence for sake of wealth & other purposes. Institutions help to resolve the conflict between the social economic and political factors.

The differences in infrastructure are caused by human capital, education, and productivity and the cross-country difference is due to the difference in the institutional indicators among countries. Institutions and trade have statistically significant and positive effect on per capita growth while remittances hold back the economic growth in the case of sample countries. The country-specific institutions are positively and strongly affected by the growth rate of regional productivity.

Institutions are essential for any economy. Key role in the determination of differences in the cross- country income is played by its institutions have the effect on developmental & growth processes through numerous/various indicators. However, the conflict is there in the literature of economy among researchers. Some are the supporters of this view according to which economic growth depends upon the quality of the institutions while the researchers who oppose these views, they say that analysis of the institution is in the initial stage. Therefore, more work is required in this aspect. This study will focus on the investigation that either the institutions, help in the explanation of developing countries through panel data analysis.

Objectives

Estimation of panel data models by employing Classical and Bayesian technique.

Selection of the best model based on Classical and Bayesian methodology.

The structure of this study is as follows. Section 2 reviews the past studies on the economic growth (GPD) and their relationship with the capital stock, trade openness and the four institutions (economic institution, financial institution, social institution and political institution). Section 3 contains data sources and methodology employed and specification of model, moreover this section also indicates panel data models and their Bayesian derivations. Section 4 presents analyses and their interpretation, moreover diagnostic plots is also included. While section 5 presents the conclusion and discuss the policy implications.

2. LITERATURE REVIEW

Moral-Benito, E. (2012) took the panel data of 73 countries for the period of 1960 to 2000. In this paper they use the data of 35 variables including dependent variable (GDP).

They examined the Bayesian panel data models on the economic growth. Moreover, they used the informative prior for panel data model. They used some the cross-sectional analyses have employed Bayesian model averaging to remove the issue of uncertainty model. This paper is extended the approach of panel data model with country-specific fixed effect in the order to simultaneously address the uncertainty models and endogeneity problems. The results and findings suggest that in panel setting the most robust growth determinants are the price of investment goods, distance to major world cities and political rights.

There are numerous studies investigate the relationship between institutions on economic growth. Most of the studies used panel data estimation**.** For instance, North and Thomas (1973), North (1981) studied the function of institutions describing the difference in economic increase between different countries. These studies show that while calculating the performance of different countries, institutions play an extensive role. In describing the economic growth process, "institutions play a very important role by disturbing the structure of investment in human and physical capital as well as technological advancement and innovations. It is assumed that security of property rights plays a noteworthy role for estimation of the long-term economic increase or growth. Mauro (1995) studied the association between economic growth and institutions. He concluded that corruption has a strong negative effect on economic growth.

Barro (1998) took panel data of hundred countries observed from 1960 to 1995 and analyze determinants of investment and economic growth. He concluded that by holding institutions, fixed government policies and national population character, the early level of GDP per capita is inversely proportional to growth rate of GDP per capita. Hall and Jones (1999) took data from institutions from the international country risk guide dataset for one hundred twenty-seven countries. They examined that the changes in organization are caused by human capital, education, and productivity. 8 They also concluded that cross-country difference is due to the difference in the institutional indicators among countries.

Acemoglu et al. (2006) also determine the effect of institutions on the economic growth. They use "constraint on executive" as a substitute for property institutions and show that the property institutions have a great effect on the financial development, long-run economic growth, macroeconomic stability, and investment.

Drury et al., (2006) used panel data for more than a hundred countries from 1982 to 1997(taking data from ICRG). They study the relation of corruption with democracies and non-democracies and found that the corruption has an insignificant effect on economic growth in the case of democracies while in the case of no democracy's corruption has a significant effect and this significant effect is harmful to the economy.

Lee and Kim (2009) have studied the relationship between economic growth and institutions with the other control variables such as R and D and education. For both, developed and underdeveloped countries panel data was used. They used a fixed effect model for their study and apply GMM for estimation. They conclude that the institutions and R and D has significant effect on economic growth in the case of developed countries while in the case of under-developed or developing countries they found weak relationship of economic growth with R and D. They also find that primary education can affect positive and significant on economic growth in the case of developing countries.

Arbia at al., (2010) took panel data of European regions over the period from 1991 to 2004 for the statistical analysis of growth experiences. They examined the effects of country-specific institutions and regional productivity growth rate. By means of control charts and using spatial weight matrix, these effects have been designed. The study reveals that the country-specific institutions are positively and strongly affected by the growth rate of regional productivity.

Acemoglu and Robinson (2010) studied that why some countries grow slower than the others. They concluded that besides physical capital, human capital and technology are the determinants of economic growth, fundamental causes can also affect growth. Institutions show a significant role in influencing the economic growth process by affecting physical and human capital. In general, it is believed that the institutions especially property rights security play a significant role in economic increase or growth.

Valeriani and Peluso (2011) employing a panel data from 1950 to 2009 for one hundred eighty-one countries by using common effect model (CEM) and fixed effect model (FEM) to investigate the effect of institutions superiority on the economic growth at various steps of development. They tell that institutions and quality of government have a positive impact on economic growth. They moreover examined that institutions are more helpful for describing features of economic growth in developed countries instead of developing countries.

Chauffour (2011), use panel data for other than a hundred countries from 1975 to 2007 and find that "institutions play an extensive role in calculating the growth 10 performance of countries". Institutions play important role in influencing the economic growth process by affecting physical and human capital. In general, it is believed that the institutions especially property rights security play a significant role in economic increase or growth.

Massa (2011) takes a sample of one hundred one countries to examine the connection between financial institutions and economic growth. She uses panel data for the time period from 1986 to 2009. Data is collected from PRS (ICRG). She examines the significant and positive impact of investment on economic growth. His results show that eleven percent increase in the investment can cause one-point twenty-nine percent increase in economic growth in the case of low salary countries while eleven percent rise in the investment can cause point eight nine percent increase in economic growth in the case of high-income countries.

Vieira at al. (2012) used cross-sectional and panel data analysis for a set of developed and under-developed countries to investigate the part of institutions on per capita income and economic growth. Data was collected from PRS (ICRG). They used GMM for statistical analysis. Their economic growth models show that institutions like investment profile and law and order can affect significantly on the growth process.

Dias and Tebaldi (2012) can use panel data for cross countries from 1965 to 2005 and studied the relationship between human capital, institutions, and economic growth. They find that political institutions are not correlated with long-term 11 economic growth and productivity. They also show that human capital and physical capital determines economic growth in the long run.

Siddiqui and Ahmed (2012) studied that how institutions affect economic increase. Institutional indicators each covering eighty-four countries with a time period of five years has been used to obtain factors depend upon PCA. Factors depend upon on these indicators are divided as political rants, policy rents, and institutional and risk-decreasing technology. These factors are used in the model by employing a generalized method of moments-based estimation and panel ordinary least square estimation. They find that economic growth is positively affected by institutions. This study also finds that in underdeveloped countries the policy rent and institutional is more important than other 2 indices.

Saima Nawaz (2014) used panel data over the time period from 1981 to 2010 for fifty-six countries to study the effect of different institutions on the economic increase. Data is collected from PRS (international country risk guide). She calculates the fixed effect model and dynamic random effect model using system GMM. Statistical analysis of this study shows a positive relationship among economic growth and institutions. The effect of institutional indicators is positive in low profits countries as compared to high profits countries. Investment profile effect is less effective in developed countries as compared to developing countries for the growth process. The analysis shows that for the long-term economic increase institutions are very important. Institutions play a key role for development in developed countries as compared to developing countries.

Yıldırım and Gökalp (2016) analyzed the relationship between institutions and macroeconomic shows in term of developing countries. For this purpose, a time of 2000-2011 years they use 23 institution and 12 variables in the study, they examine the relationship between institutions and macroeconomic in some countries whereas 38 countries using panel data analysis. The investigated results tell us that the institution indicates as the trade barriers, the restriction of foreign investment have a positive effect on the economic growth of the countries. On the other side, according to that investigated results, variables as civil freedom, government expenditures, collective bargaining effect on the macroeconomic condition of developing countries.

Dutta and Williamson (2016) using a panel of 108 countries from 1971 to 2010. Their key motivation of the study is the aid effect on economic freedom conditional on quality of political institutions. By investigating aid influence on economic freedom provisional on the value of political institutions, they find that in the case of democracies aid can pick up economic freedom while in the case of autocracies aid can decrease economic freedom. Their results are used for making policies. They also conclude that those countries who need aid from other countries do not enjoy healthy political institutions.

3. DATA SOURCES AND METHODOLOGY

3.1 Data and Construction of Variable

Specifically, is to explore the relationship among various institutions and economic growth, thus panel data is used from the time span 1990 to 2014 (24 years) data for the case of 27 international development association countries (IDA), which are developing countries. Data of the 27 developing countries) were taken from the world development indicator (WDI). The data for GDP and Gross capital formation is collected from (WDI). International country risk guide (ICRG) data source is used to collect the data of different type's institutions. The ICRG was prepared to forecast political, economic, social and financial risks. The ICRG make available the data for 27 countries on annual basis. The ICRG used those kinds of data which belongs to banks, investors, and multinational importers and exporters' trades etc. The benefit of using ICRG data is to know the political, economic and financial risk that cans the investment and business and finally the economic growth of the country.

Table 3.1
Source of Data and Description of Variables

Variable	Notation	Description	Sources
GDP per Capita	Y_t	Total output of country divided by total number of people	World Development Indicator (IDA) (WDI-2016)
Physical Capital	K_t	Real Gross Capital Formation in US$ with base year (2010)	World Development Indicator (WDI-2016)
Economic Institution	E_t	1. GDP per Head. 2. Real GDP Growth. 3. Annual Inflation Rate. 4. Budget Balance as a Percentage of GDP. 5. Current Account as a Percentage.	International Country Risk Guide (ICRG)
Financial Institutions	F_t	1. Foreign Debt as a Percentage of GDP. 2. Foreign Debt Services as a Percentage of Exports of goods and services. 3. Current Account as a Percentage of Exports of Goods and Services. 4. Net International Liquidity as Month of Import. 5. Exchange Rate Stability.	International Country Risk Guide (ICRG)
Political Institutions	P_t	1. Government Stability. 2. Socioeconomic Variable. 3. Investment Profile. 4. Internal Conflict. 5. External Conflict. 6. Corruption. 7. Military in Politics 8. Law and order. 9. Democratic Accountability. 10. Bureaucracy Quality.	International Country Risk Guide (ICRG)
Social Institutions	S_t	1. Ethnic Tension. 2. Religious Tension	International Country Risk Guide (ICRG)

3.2 Model's Specification

Follow "Hall and Jones (1999)", we can estimate the influence of institutions on economic growth using the following statistical model. In this model we can use capital stock, trade openness, financial institutions, economic institutions, social institutions and political institutions as control variables. Hence, the model may be written as follows.

In this section we would specify our model

$$y_{it} = \beta_0 + \beta_1 k_{it} + \beta_2 T_{it} + \beta_3 F_{it} + \beta_4 E_{it} + \beta_5 S_{it} + \beta_6 P_{it} + \mu_{it}$$

where, Y_{it} is the dependent variable which represent GDP, βo is the intercept of the model,

k_{it} Capital stock
T_{it} Trade openness
F_{it} Financial Institutions
E_{it} Economic Institution
S_{it} Social institutions
P_{it} Political institutions
μ_{it} Error term of the model.

The model is further specifying on the following different methods.

3.3 Bayesian Statistical Modeling Approach

The basic knowledge of Bayesian theory is to use prior information and sample information to obtain the posterior knowledge precisely, which express the posterior density

$$f(\theta / y) = \frac{L(\theta / y)g(\theta)}{\int L(\theta / y)g(\theta)d\theta}$$

or

$$f(\theta / y) \propto L(\theta / y)g(\theta)$$

whereas, $g(\theta)$ is a prior probability density of parameter θ, and that defines what is known about the strange parameter prior to the analysis of sample x. And $L(\theta / y)$ is the likelihood function of sample x that is sampling distribution of the samples assumed chosen probability model and parameter.

Details about prior information Bayesian theory are discussed as follows:

a) **Prior Distribution**
In the Bayesian analysis, the necessary period is about to prior distribution. In overall non-informative prior, individual Bayes and empirical Bayesian are mostly proposed rules to fix a prior distribution. As an example of uniform distribution resulting via using the non-informative prior was selected in this study, i.e. defining the range of a parameter and assigning it uniformly distributed within the range.

b) **Likelihood Function and Posterior Density**
When we detected the data then the likelihood function is constructed. The likelihood function is the joint probability function of the data, but seen as like a function of the parameters, discussing the detected data as fixed quantities. Supposing that the data values, $y = (y_1, y_2, ..., y_n)$ are found independently, generally the likelihood function is written as

$$L(\theta / y) = f(y_1, y_2, ...y_n / \theta) = \prod_{i=1}^{n} f(y_i / \theta)$$

where, q_r is the return level related with the 1/r-year return period? In order to get the parameter values, Bayesian Markov Chain Monte Carlo (MCMC) simulation method will be used. Therefore, the likelihood functions for $X_1, X_2, ..., X_n$ is given by:

$$L(\theta / y) = L(\mu, \sigma, \xi : Y_1, ..., Y_n) = \prod_{i=1}^{n} f(y_i \mid \mu, \sigma, \xi)$$

where θ is the vector of parameters [i.e. $\theta = (\mu, \sigma, \xi)$].

Thus the density of posterior distribution is:

$$f(\theta / Y_1, ..., Y_n) \propto L(\theta / Y_1, ..., Y_n) \times g(\theta)$$

$$f(\mu, \sigma, \xi / Y_1, ..., Y_n) \propto L(\mu, \sigma, \xi / Y_1, ..., Y_n) \times g(\mu, \sigma, \xi)$$

c) **Posterior Predictive Distribution**
According to Coles (2003), Bayesian analysis is better due to the prediction of return level which is based on predictive distribution can be estimated easily.

Let y represents the future observation with probability density function:

$$h(x \mid Y_1, ..., Y_n) = \iiint f(x \mid \mu, \sigma, \xi) f(\theta \mid Y_1, ..., Y_n) d\mu d\sigma d\xi$$

$$h(x \mid Y_1, ..., Y_n) = \iiint f(x \mid \mu, \sigma, \xi) f(\mu, \sigma, \xi \mid Y_1, ..., Y_n) d\mu d\sigma d\xi$$

3.3.1 Bayesian Derivation of Common Effect Model

The model specified as follows

$$yit = \beta o + \beta_1 x_{1it} + \beta_2 x_{3it} + \beta_3 x_{3it} + \beta_4 x_{4it} + \beta_5 x_{5it} + \beta_6 x_{6it} + \varepsilon it \quad (1)$$

where I = 1, 2, 3….N, t = 1, 2…….N

We can write the above model in the matrix form and the final structure model we can write as follows.

$$y = x\beta + ui \qquad (2)$$

The likelihood function of the above model is follows

$$\mathrm{P}(y/\beta, \mathrm{H}) = \frac{H^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}} \{\exp[-\tfrac{H}{2}(y - x\beta)^T (y - x\beta)]\} \qquad (3)$$

$$[(y - x\beta)^T (y - x\beta)]$$

Add and subtract $x\beta$

$$[(y - x\hat{\beta} + x\hat{\beta} - x\beta)^T(y - x\hat{\beta} + x\hat{\beta} - x\beta)]$$

$$[(y - x\hat{\beta}) - x(\beta - \hat{\beta})]^T[(y - x\hat{\beta}) - x(\beta - \hat{\beta})]$$

$$[(y - x\hat{\beta})^T\big(y - x\hat{\beta}\big) + (\beta - \hat{\beta})^T x' x\big(\beta - \hat{\beta}\big)] \quad (4)$$

As well as the cross product expressions

$$[(\beta - \hat{\beta})^T x^T\big(y - x\hat{\beta}\big) = (\beta - \hat{\beta})^T(x^T y - x^T x(x^T x)^{-1}x^T y) = 0$$

$$(y - x\hat{\beta})^T\big(y - x\hat{\beta}\big) = \text{SSE}$$

$$(y - x\hat{\beta})^T\big(y - x\hat{\beta}\big) + (\beta - \hat{\beta})^T x^T x(\beta - \hat{\beta})$$

$$(y - x\hat{\beta})^T\big(y - x\hat{\beta}\big) = SSE + (\beta - \hat{\beta})^T x^T x\big(\beta - \hat{\beta}\big) \quad (5)$$

Put equation (5) in equation (3) we have

$$\text{P}(y/\beta, \text{H}) = \frac{H^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}}\{\exp[-\tfrac{H}{2}SSE + (\beta - \hat{\beta})^T x^T x\big(\beta - \hat{\beta}\big)]\}$$

Non Informative Prior for Linear Model

By the definition of the uniform prior

$$p(\beta) \propto c \text{ and } p(Y) = 1 \quad (6)$$

Using the likelihood function (3) and the prior distribution (6) then the posterior distribution of the model is

$$\text{P}(y/\beta, \text{H}) = \frac{H^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}}\{\exp[-\tfrac{H}{2}(y - x\beta)^T(y - x\beta)]\} \quad (7)$$

By using rules of OLS method

$$\hat{\beta} = (x^T x)x^T y \;\; \delta^2 = \frac{\big(y - x\hat{\beta}\big)^T\big(y - x\hat{\beta}\big)}{(n-k)}$$

By using the complete we write

$$= \frac{H^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}}\{\exp[-\tfrac{H}{2}SSE + (\beta - \hat{\beta})^T x^T x\big(\beta - \hat{\beta}\big)]\}$$

Therefore the joint posterior distribution function as follows

$$p(\beta, y/X, Y) \propto p(\beta, y/X, Y)p(\beta)p(y)$$

$$\propto \frac{H^{\frac{N}{2}}}{(2\pi)^{\frac{N}{2}}}\Big\{\exp\Big[-\tfrac{H}{2}SSE + \big(\beta - \hat{\beta}\big)^T x^T x\big(\beta - \hat{\beta}\big)\Big]\Big\} \quad (8)$$

By using the kernel density method we get the following equation

$$\propto H^{\frac{N}{2}}\, exp^{-\frac{H}{2}(SEE)} \times exp^{[-\frac{1}{2}(\beta-\widehat{\beta})^T H x^T x(\beta-\widehat{\beta})]}$$

$$\propto H^{a^*-1} exp^{-Hb^*} \times exp^{[-\frac{1}{2}(\beta-\widehat{\beta})^T Q^*(\beta-\widehat{\beta})]} \quad (9)$$

The posterior parameters are

$$a^* = \frac{N+2}{2},\ b^* = \frac{SSE}{2}, \hat{\beta} = (x^T x)x^T y\ ,\ Q^* = H(x^T x)$$

Normal Linear Regression Model through Independent Normal Gamma Prior

By using natural conjugate prior whereas $p(\,\beta/H)$ is a normal density and $p(H)$ is gamma density function. Here we use the same prior for the independence of β and H.

Specifically, we adopt $p(\,\beta/H) = p(\beta)p(H)$ with $p(\beta)$ existence normal distribution and $p(H)$ being the pdf of gamma distribution.

$$p(\beta) = \frac{1}{(2\pi)^{\frac{k}{2}}} |\underline{Q}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2}\left(\beta-\underline{\beta}\right)\underline{Q}^{-1}\left(\beta-\underline{\beta}\right)\right] \quad (10)$$

and

$$p(H) = C_G^{-1} H^{a-1} \exp(-Hb) \quad (11)$$

where C_G^{-1} is the integrating constant used for the gamma pdf? That is $\underline{\beta} = E(\beta/Y)$ is still the prior mean of β and $\underline{Q}$ is the variance and covariance matrix of β. Where var $(\beta/H) = H^{-1}\underline{Q}$. Where in $p(\beta)$ "a" is the scale parameter and "b" shape parameter. The parameter in $p(\beta)$ and $p(H)$ can be find as follows.

By the combining the equation (10) and equation (11) we get results of normal gamma prior

$$p(\beta, H) \propto \exp\left[-\frac{1}{2}\left(\beta-\underline{\beta}\right)\underline{Q}^{-1}\left(\beta-\underline{\beta}\right)\right][H^{a-1}\exp(-Hb)] \quad (12)$$

Posterior Distribution under Normal Gamma Prior

$$p(\beta, H/Y) \propto p(\beta, H)L(\beta, H) \quad (13)$$

$$\begin{aligned} p(\beta, H/Y) \propto \exp\left[-\frac{1}{2}\left(\beta-\underline{\beta}\right)\underline{Q}^{-1}\left(\beta-\underline{\beta}\right)\right][H^{a-1}\exp(-Hb)] \\ \times\left\{\exp\left[-\frac{H}{2}SSE + \left[\left(\beta-\hat{\beta}\right)^T x^T x\left(\beta-\hat{\beta}\right)\right]\right]\right\} \end{aligned} \quad (14)$$

$$\begin{aligned} = H^{a+\frac{1}{2}-1}\exp(^{-H\left(b-\frac{SSE}{2}\right)}) \exp\left[-\frac{1}{2}\left(\beta-\underline{\beta}\right)\underline{Q}^{-1}\left(\beta-\underline{\beta}\right)\right. \\ \left. + (\beta-\hat{\beta})^T H x^T x\left(\beta-\hat{\beta}\right)\right] \end{aligned}$$

$$\begin{aligned} = H^{a_1-1} exp^{(-Hb_1)} \exp\left[-\frac{1}{2}\left(\beta-\underline{\beta}\right)\underline{Q}^{-1}\left(\beta-\underline{\beta}\right)\right. \\ \left. + \left(\beta-\hat{\beta}\right)^T H x^T x\left(\beta-\hat{\beta}\right)\right] \end{aligned} \quad (15)$$

Now taking a part of equation 15

$$\left[\left(\beta-\underline{\beta}\right)\underline{Q}^{-1}\left(\beta-\underline{\beta}\right)+\left(\beta-\hat{\beta}\right)^{T}Hx^{T}x\left(\beta-\hat{\beta}\right)\right]$$

After the simplification of the above equation we get the following the equation

$$H^{a_1-1}exp^{(-Hb_1)}\exp\left[-\frac{1}{2}(\underline{Q}^{-1}+Hx^Tx)\left(\beta-\frac{\underline{Q}^{-1}\underline{\beta}+Hx^Tx\hat{\beta}}{\underline{Q}^{-1}+Hx^Tx}\right)^2\right] \quad (16)$$

$a_1 *= a+\frac{N}{2}$ and $b_1 *= b+\frac{SSE}{2}$

As we know that

$$\overline{Q}^{-1}=\left(\underline{Q}^{-1}+Hx^Tx\right)$$
$$\overline{Q}=(\underline{Q}^{-1}+Hx^Tx)^{-1}$$
$$\overline{\beta}=\frac{\underline{Q}^{-1}\underline{\beta}+Hx^Tx\hat{\beta}}{\underline{Q}^{-1}+Hx^Tx}=\frac{\underline{Q}^{-1}\underline{\beta}+Hx^Tx\hat{\beta}}{\underline{Q}^{-1}}$$
$$\overline{\beta}=\overline{Q}\left(\underline{Q}^{-1}\underline{\beta}+Hx^Tx\hat{\beta}\right)$$

where, $\overline{\beta}$ is the mean and $\overline{Q}$ is var-cov matrix of posterior distribution. However, it is must be stressed that both distribution of normal Gamma prior and likelihood function does not related to the posterior and here the posterior simulator is known as Gibbs sampler which follows multivariate normal distribution and gamma distribution. After this we will go towards Monte Carlo integration.

3.3.2 Bayesian Derivation of Fixed Effect Model

$$Y_{it}=\ \beta oi+\beta_1x_{1it}+\beta_2x_{3it}+\beta_3x_{3it}+\beta_4x_{4it}+\beta_5x_{5it}+\beta_6x_{6it}+\varepsilon it \quad (1)$$

We can write the above model in the matrix form

$Y_{it}=[y_1,y_2,\dots y_n], X_{it}=[1,1\dots 1,\ x_{11},\ x_{12}\dots\ x_{1n},\ x_k\dots\ x_{kn}\ ], \beta=[\beta_0\dots\beta_k]$, $\epsilon=[\epsilon_1\dots\epsilon_n]$

The model 1 is written as follows

$$Y=X\beta+\in$$

where, $\in\sim NID(0,\delta^2)$

$$f(y)=\frac{1}{\sqrt{2\pi\delta^2}}exp^{\frac{-\in^t\in}{2\delta^2}}$$
$$\mathrm{l}L=\prod_{i=1}^{n}\left[\left(\frac{1}{\sqrt{2}\pi\delta^2}exp^{\frac{-\in^t\in}{2\delta^2}}\right)\right]$$

$$L = \left(\frac{1}{2\pi\delta^2}\right)^{\frac{n}{2}} exp^{\frac{-\in^t \in}{2\delta^2}} \tag{2}$$

$$Y = X\beta - \in$$

$$\in = Y - X\beta$$

Put the of ∈ in model 2 we get the following equation

$$L = \left(\frac{1}{2\pi\delta^2}\right)^{\frac{n}{2}} exp^{[-\frac{1}{2\delta^2}[(Y-X\beta)^t(Y-X\beta)]}$$

Now, Appling kernel density method, we get the following equation

$$L \propto \left(\frac{1}{\delta^2}\right)^{\frac{n}{2}} exp^{[-\frac{1}{2\delta^2}(Y-X\beta)^t(Y-X\beta)]} \tag{3}$$

where, $\hat{y} = x\hat{\beta}$

Now consider

$$(y - x\beta)^T(y - x\beta)$$

Adding and subtracting $\hat{y}$

$$(y - \hat{y} + \hat{y} - x\beta)^T(y - \hat{y} + \hat{y} - x\beta)$$

Now put the value of $\hat{y}$ in the above equation

$$[(y - x\hat{\beta} + x\hat{\beta} - x\beta)^T(y - x\hat{\beta} + x\hat{\beta} - x\beta)]$$
$$[(y - x\hat{\beta}) - x(\beta - \hat{\beta})]^T[(y - x\hat{\beta}) - x(\beta - \hat{\beta})]$$
$$[(y - x\hat{\beta})^T\left(y - x\hat{\beta}\right) + (\beta - \hat{\beta})^T x' x\left(\beta - \hat{\beta}\right)] \tag{4}$$

<u>For general information:</u>

$$\delta^2 = \frac{(y-\hat{y})^2}{n-1} \text{ as } \hat{y} = x\hat{\beta}$$

$$\delta^2 = \frac{(y-x\hat{\beta})^2}{n-1}$$

When we write in matrix form the square term will be written in matrix form

$$A^2 = A^T A$$

$$\delta^2 = \frac{(y - x\hat{\beta})^T(y - x\hat{\beta})}{n-k}$$

$$(n-k)\delta^2 = (y - x\hat{\beta})^T(y - x\hat{\beta})$$

Since,

$$(\beta - \hat{\beta})^T x^T(y - x\hat{\beta})$$

where,

$$\hat{\beta} = (x^T x)x^T y$$

Put the value of $\hat{\beta}$ in the above equation

$$(\beta-\hat{\beta})^T x'(y-x(x'x)^{-1}x'y)$$

$$(\beta-\hat{\beta})^T (x'y-x'y)$$

$$(\beta-\hat{\beta})^T (0)=0$$

Now

$$(\beta-\hat{\beta})^T x^T(y-x\hat{\beta})$$

Similarly,

$$(y-x\hat{\beta})^T x(\beta-\hat{\beta})^T=0$$

Put the above equation in equation (4) we will get the following equation

$$=(n-k)\delta^2-0-0+(\beta-\hat{\beta})^T x'x(\beta-\hat{\beta})$$
$$=(n-k)\delta^2+(\beta-\hat{\beta})^T x'x(\beta-\hat{\beta})$$

where, $v=n-k$ put in equation 4

$$L(y)\propto(\tfrac{1}{\delta^2})^{\frac{n}{2}}exp^{[-\frac{1}{2\delta^2}[v\delta^2+(\beta-\widehat{\beta})^T x'x(\beta-\widehat{\beta})]} \quad (5)$$

To derive posterior distribution of $\beta\ and\ \delta^2$ we have to specify prior distribution.

Assuming that β_i follows normal distribution with hyper parameter (β_{0i},δ_i^2)

$$\beta_i\sim NID(\beta_{0i},\delta_i^2)$$

i.e.

$$\beta_0\sim N(\beta_0,\delta_0^2)$$
$$\beta_1\sim N(\beta_1,\delta_1^2)$$
$$\beta_k\sim N(\beta_k,\delta_k^2)$$
$$\beta\sim MN(\beta_0,\varepsilon_0)$$

$$P(\beta)=(2\pi)^{-\frac{K}{2}}|\varepsilon|^{1/2}exp^{\left[-\frac{1}{2}(\beta-\beta_0)\varepsilon_0{}^{-1}(\beta-\beta_0)\right]}$$

Now, assuming that δ^2 follow Gamma distribution with hyper parameter (a, b)

$$\delta^2\sim Gamma\ (a,b)$$

[v_0 is the prior degree of freedom]

$$p(\delta^2)=\frac{b^a}{\sqrt{a}}(\delta^2)^{a-1}e^{-b/\delta^2}$$

Put the value of “a” and “b” in the above equation

$$p(\delta^2)=\frac{(\frac{v_0\delta_0}{2})^{v_0/2}}{\sqrt{v_0/2}}(\delta^2)^{v_0/2-1}e^{-v_0\delta_0^2/2/\delta^2}$$

Now

$$p(\beta, {}^{\delta^2}/_{y}\ , x) \propto L(y) \times p(\beta) \times p(\delta^2)$$

$$\propto (\frac{1}{\delta^2})^{\frac{n}{2}} exp^{[-\frac{1}{2\delta^2}[v\delta^2+(\beta-\widehat{\beta})^T x' x(\beta-\widehat{\beta})]} \times exp^{\left[-\frac{1}{2}(\beta-\beta_0)\varepsilon_0{}^{-1}(\beta-\beta_0)\right]}$$

$$\times (\delta^2)^{{}^{v_0}/_{2}-1} e^{-{}^{v_0\delta_0^2/2}/_{\delta^2}}$$

$$\propto [(\delta^2)^{\frac{v_0}{2}-1-\frac{n}{2}} exp^{\left[-\frac{v_0\delta_0^2-v\delta^2}{2\delta^2}\right]} \times exp^{\left[-\frac{1}{2}(\beta-\beta_0)\varepsilon_0{}^{-1}(\beta-\beta_0)+(\beta-\widehat{\beta})^T\frac{x'x}{\delta^2}(\beta-\widehat{\beta})\right]}$$

$$\propto (\delta^2)^{-\left(\frac{v_0}{2}+\frac{n}{2}\right)-1} exp^{-\left[\frac{v_0\delta_0^2-v\delta^2}{2\delta^2}\right]} \times exp^{\left[-\frac{1}{2}(\beta-\beta_0)\varepsilon_0{}^{-1}(\beta-\beta_0)+(\beta-\widehat{\beta})^T\frac{x'x}{\delta^2}(\beta-\widehat{\beta})\right]} \quad (6)$$

Let

$$a^* = \frac{v_0 + n}{2}, b^* = \frac{v_0\delta_0^2 + v\delta^2}{2}$$

$$M^* = \left(\varepsilon_0{}^{-1}\beta_0 + \frac{x'x\hat{\beta}}{\delta^2}\right)$$

$$V^* = \left(\varepsilon_0{}^{-1} + \frac{x'x}{\delta^2}\right)^{-1}\left(\varepsilon_0{}^{-1}\beta_0 + \frac{x'x\hat{\beta}}{\delta^2}\right)$$

Put in equation (6) we will get

$$p\left(\beta, {}^{\delta^2}/_{y}\ , x\right) \propto MN{\sim}G(M^*, V^*, a^*, b^*).$$

4. RESULTS AND DISCUSSION

4.1 Estimation of Panel Data Model under Classical Framework

The following table shows the results of two panel data models under classical framework.

Table 4.1

Models	**Common Effect Model**	**Fixed Effect Model**
Coefficients	Estimates [std. Error] (P-value)	Estimates [std. Error] (P-value)
Intercept	1676.78 70.34 0.000	578.2992 33.95059 0.000
Capital stock	636.3076 37.30354 0.000	31.50383 12.52307 0.012
Trade openness	1.350232 0.3795391 0.000	1.536507 0.4794899 0.001
Financial Institutions	119.3162 17.48536 0.000	9.951428 24.94693 0.090
Economic Institutions	86.83288 18.3888 0.000	27.75865 33.29542 0.004
Social Institutions	-1.559234 14.60306 0.915	11.80383 18.92561 0.533
Political Institutions	58.74759 16.26041 0.000	20.26469 22.53992 0.069
$\boldsymbol{R^2}$	0.3696	0.0243
Adjusted $\boldsymbol{R^2}$	0.3982	0.0300
F test	0.23	0.38

On the basis of probability value we can decide about significance and non-significance of the parameters. We can take 5 percent level of significance, if probability value is less than 0.05 then the parameter estimate is significant, but if probability value is more than 0.05 then the parameter estimate is insignificant.

In the above table results indicates that common effect model is the best model among the two models. F test is used between the compression of common effect and fixed effect. According to F test results Common effect is the best model in the classical approach. One unit change in gross capital stock causes 636.3076 unit increase in GDP value added with standard error 37.30354 and shows significant effect on dependent variable while one unit increase in trade openness causes 1.350232 unit increase in GDP value added with standard error 0.3795391 and shows significant effect on dependent variable moreover one unit increase in Financial Institutions causes 119.3162 unit increase in GDP value added with standard error 17.48536 and shows significant effect on dependent variable. One unit change in economic institutions causes 86.83288 unit increase in GDP value added with standard error 18.3888 shows significant effect of dependent variable, furthermore one unit change in social institutions causes -1.559234 unit decrease in GDP value added with standard error 14.60306 shows negative insignificant effect of dependent variable and one unit change in political institutions causes 58.74759 unit increase in GDP value added with standard error 16.26041 and shows significant effect of dependent variable.

4.2 Estimation of Panel Data Model under Bayesian Framework

The following table shows the results of two panel data models under Bayesian framework

Table 4.2

Models	Common Effect Model	Fixed Effect Model
Coefficients	Mean [Std. Error] C.D.I[2.5-97.5%]	Mean [Std. Error] C.D.I[2.5-97.5%]
Intercept	1671.0 0.002093 16071.0-16071.0	140.8 0.000862 126.1-156.7
Capital Stock	636.3 0.02684 636.3-6.36.4	0.00802 0.01102 0.02948-0.01342
Trade Openness	10.34 0.0946 10.16-10.53	0.06127 0.02053 0.02122-0.1016
Financial Institutions	119.3 0.05439 119.2-119.4	0.02744 0.003989 0.01961-0.0352
Economic Institutions	86.83 0.05439 86.72-86.94	0.02156 0.004303 0.01308-0.03004
Social Institutions	1.592 0.06845 1.727-1.457	0.00196 0.003613 0.0051-0.00918
Political Institutions	58.75 0.06156 58.63-58.87	0.00439 0.003967 0.00322-0.0122

Based on 95% credible interval coefficient estimate of variables is significant as the interval doesn`t contain 0. In the above table results indicates that fixed effect model is the best model among all the models. On the basis of standard error, we conclude that fixed effect model is the best model in the Bayesian methodology of analysis. Fixed effect model shows batter results as compare to other model that`s why we only interpret fixed effect model in this research study. One unit change in gross capital stock causes 0.000862 unit increase in GDP value added with standard error 0.01102 and shows significant effect on dependent variable while one unit increase in trade openness causes 0.06127 unit increase in GDP value added with standard error 0.2053 and shows significant effect on dependent variable moreover one unit increase in Financial Institutions causes 0.02744 unit increase in GDP value added with standard error 0.0034303 and shows significant effect on dependent variable. One unit change in economic institutions causes 0.02156 unit increase in GDP value added with standard error 0.004303 shows significant effect of dependent variable, furthermore one unit change in social institutions causes 0.00196 unit increase in GDP value added with standard error 0.003613 shows significant effect of dependent variable and one unit change in political institutions causes 0.00439 unit increase in GDP value added with standard error 0.003967 and shows significant effect of dependent variable. On the basis of 95% credible interval coefficient estimate of all variables show significant effect on the dependent variable.

4.4 Diagnostic plots

Trace Series Plot

From the Fig 1 we illustrated that the trace plot of samples viruses the simulation number running the multiple chain 2, each chain indicating a different color. In this situation, we can be reasonably confident that convergence has been achieved because all chain appears to be overlapping one another. The trace tells that the chain is converted to the stationary distribution after long burns in the period. The feature of stationary that most familiar from trace plot is a relatively constant mean and variance. This figure also shows a perfect trace plot. Because the center of the chain appears to be around constant mean values with very small fluctuations. This indicates that the chain would reach the target (right, stationary) distribution. We concluded that the mixing is sufficient good for each parameter.

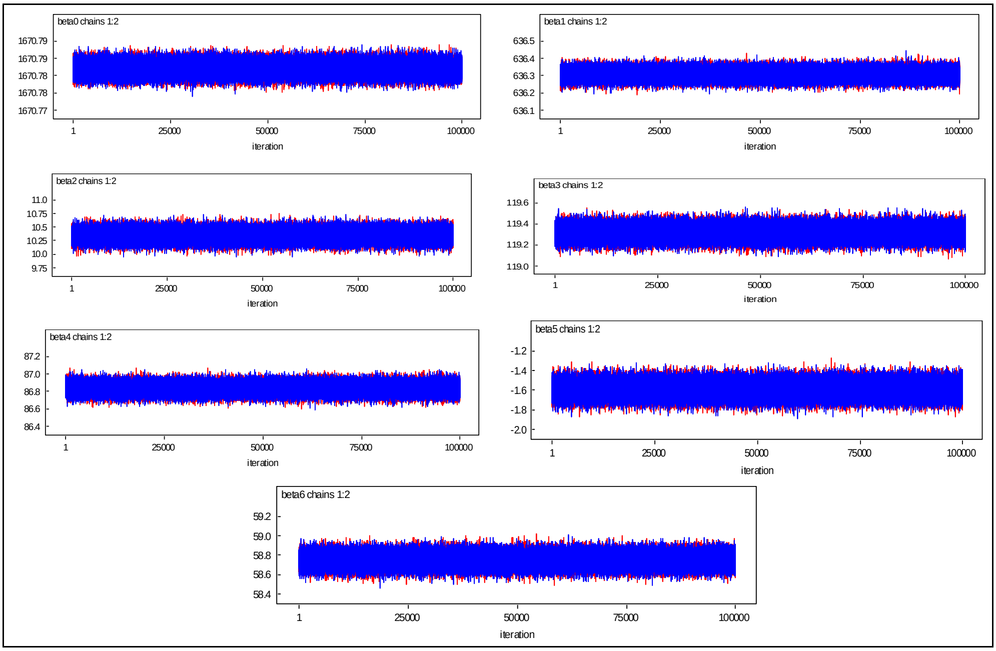


Figure 1: Trace Plot

Kernel Density Plots

From Figure 2, the kernel density plots alternative visualization of the simulated marginal posterior distribution of the parameters. The marginal posterior distribution of $\beta o, \beta_1 \dots \beta_6$ under non-informative prior is normal. Kernel density plots indicates that the Bayesian point estimate typical posterior mean or posterior median and the range between 2.5th and 95.5th percentile represents 95% Bayesian confidence interval also called credible interval. The numerical outcomes of the variables give graphical representation which shows similar results. Hence the posterior kernel density designed to stabilize and converges for all parameters.

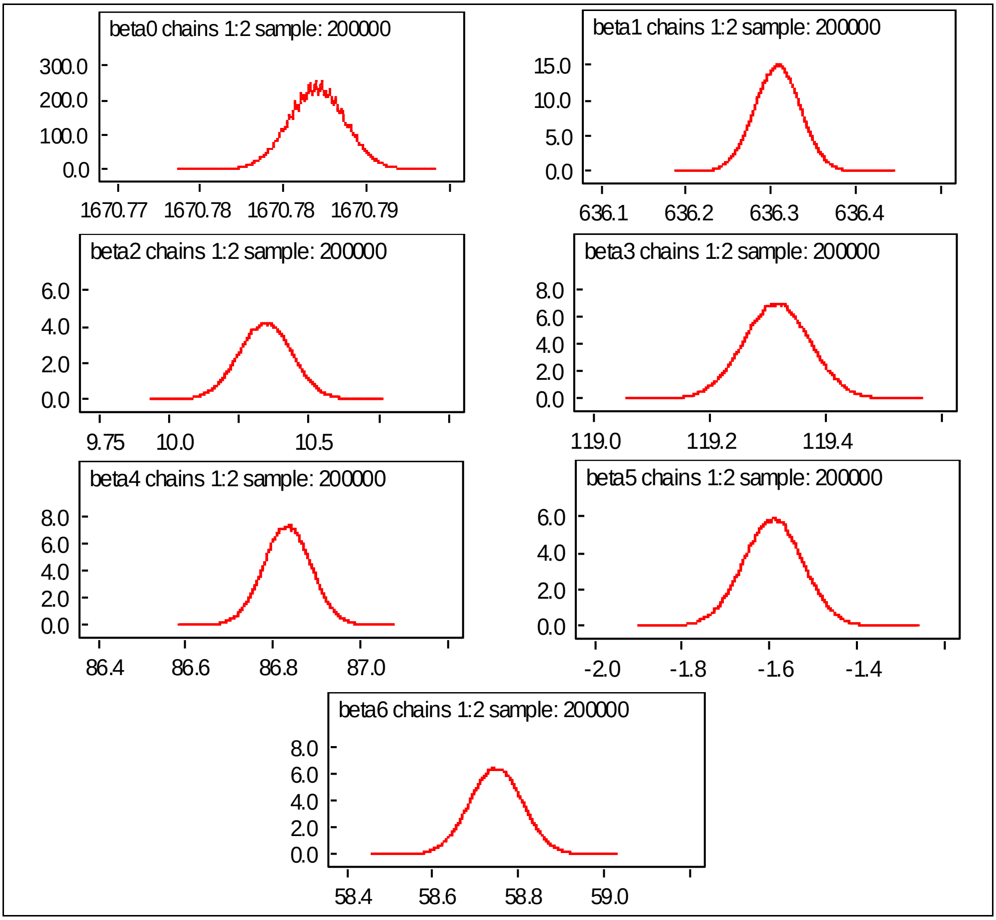


Figure 2: Kernel Density Plots

Auto Correlation Function Plots

The auto correlation function plots from Fig 3 indicates for chain of every parameter and the dimensions of the posterior distribution that are maxing. Rapidly mixing is often associated with low posterior correlation between parameters. The plots indicate that all parameters are mixing well with auto correlation vanishing before five lags in each case.

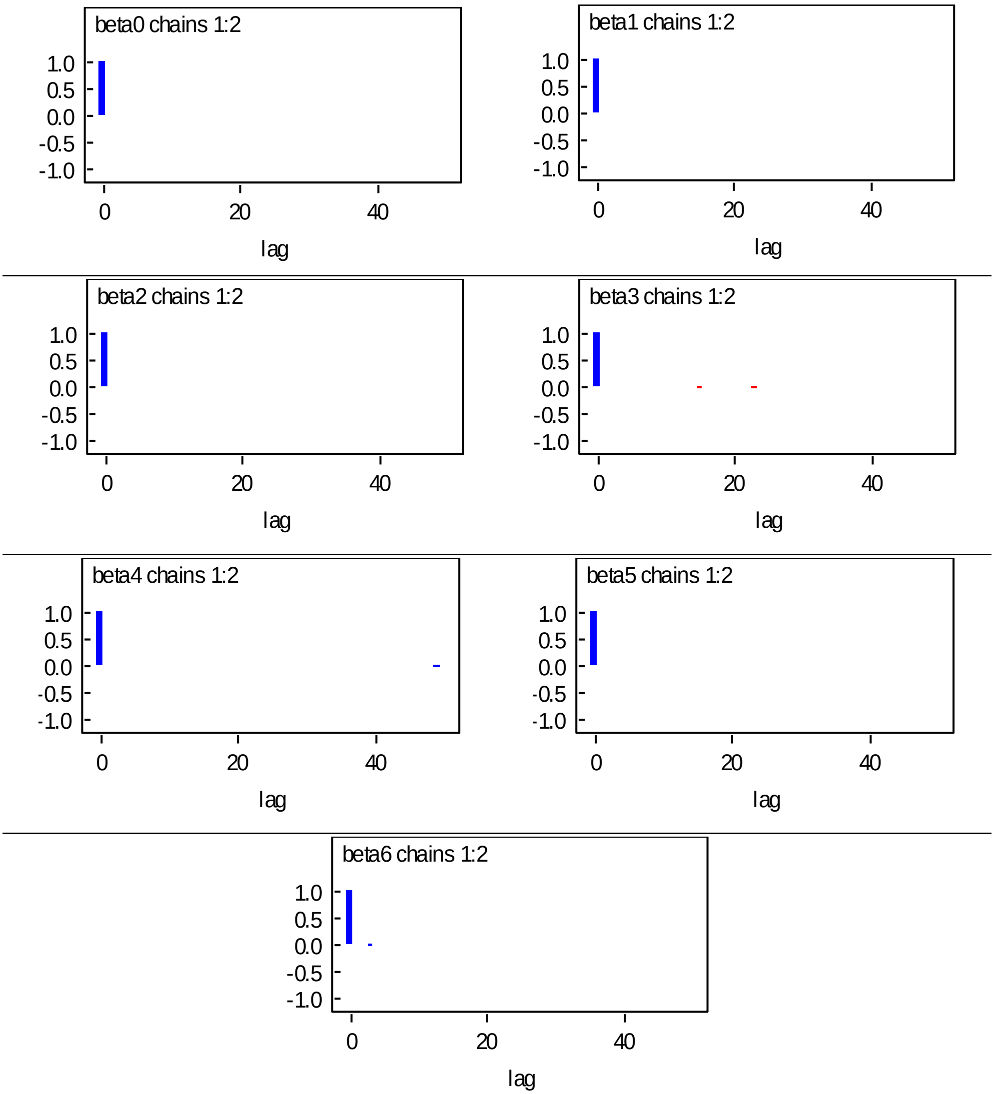


Figure 3: Auto Correlation Function Plots

BGR Plots Statistic

Fig 4 BGR-plots indicates that the diagnostic plot generated for the 2500000 values of β sampled from the coins and discarding the 1st 90,000. The blue line represents the average width of 80% credible interval computed from 3 separate chains. The green line represent computed from the pooled data and the red lines in the ratio of these two values. The fig indicates that the ratio is 1 and the 3 chain converges to its desired distribution.

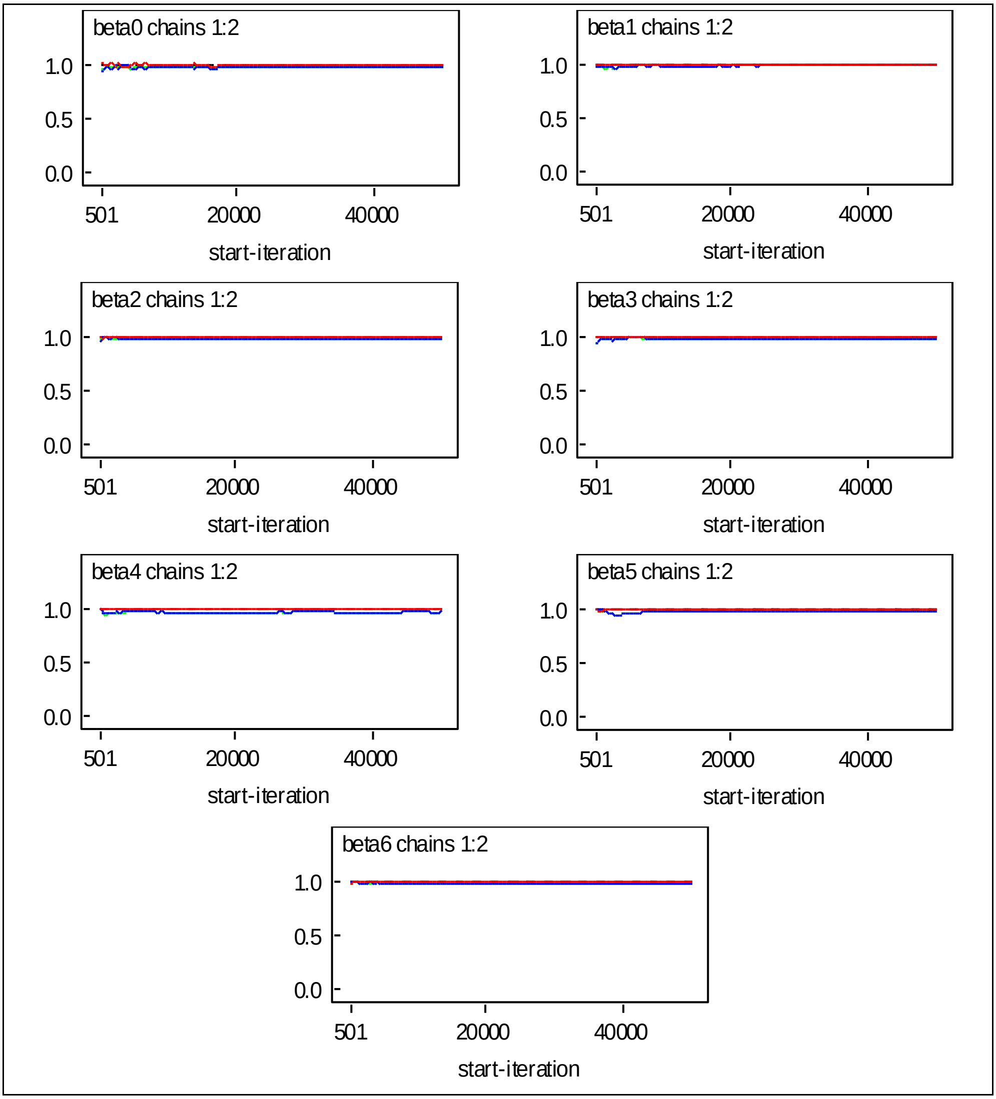


Figure 4: BGR Plots Statistic

5. CONCLUSIONS

This study examined the classical and Bayesian approach of panel data models which are common effect model, fixed effect model. This study investigated that, the classical models are not valid models due to many reasons. Common effect model is not selected as a valid model because the estimate of political institution is insignificant. In the Bayesian methodology the institutions have shown positive effect on the GDP. On the basis of standard error, we would like to say that fixed effect model is the best model in Bayesian estimation of panel data model. Because in fixed effect model have low standard error as compare to other models. Finally, Bayesian fixed effect models have shown a batter results as compare to the classical estimation of panel data models. At last we decided that Bayesian panel models are the best and appropriate models.

REFERENCES

1. Abrahamson, N.A. and Youngs, R.R. (1992). A stable algorithm for regression analyses using the random effects model. *Bulletin of the Seismological Society of America*, 82(1), 505-510.
2. Acemoglu, D. and Robinson, J. (2010). *The role of institutions in growth and development* (p. 135). World Bank Publications.
3. Alesina, A., Özler, S., Roubini, N. and Swagel, P. (1996). Political instability and economic growth. *Journal of Economic growth*, 1(2), 189-211.
4. Arellano, M. and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29-51.
5. Arbia, G., Battisti, M. and Di Vaio, G. (2010). Institutions and geography: Empirical test of spatial growth models for European regions. *Economic Modelling*, 27(1), 12-21.
6. Bhargava, A. and Sargan, J.D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica: Journal of the Econometric Society*, 1635-1659.
7. Blundell, R. and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
8. Berger, J.O. and Wolpert, R.L. (1988). The Likelihood Principle. Institute of Mathematical Statistics. *Lecture Notes*, 6.
9. Coles, S., Pericchi, L.R. and Sisson, S. (2003). A fully probabilistic approach to extreme rainfall modeling. *Journal of Hydrology*, 273(1-4), 35-50.
10. Caselli, F., Esquivel, G. and Lefort, F. (1996). Reopening the convergence debate: A new look at cross-country growth empirics. *Journal of Economic Growth*, 1(3), 363-389.
11. Chauffour, J.P., Saborowski, C. and Soylemezoglu, A.I. (2010). Trade Finance in Crisis: Should Developing Countries Establish Export Credit Agencies? The World Bank.
12. Dias, J. and Tebaldi, E. (2012). Institutions, human capital, and growth: The institutional mechanism. *Structural Change and Economic Dynamics*, *23*(3), 300-312.
13. Drury, A.C., Krieckhaus, J. and Lusztig, M. (2006). Corruption, democracy, and economic growth. *International Political Science Review*, 27(2), 121-136.

14. Dutta, N. and Williamson, C.R. (2016). Aiding economic freedom: Exploring the role of political institutions. *European Journal of Political Economy*, 45, 24-38.
15. Greenland, S. and Robins, J.M. (1985). Estimation of a common effect parameter from sparse follow-up data. *Biometrics*, 41(1), 55-68.
16. Hall, R.E. and Jones, C.I. (1999). Why do some countries produce so much more output per worker than others? *The Quarterly Journal of Economics*, 114(1), 83-116.
17. Kaushik, L.S., Singh, D.P. and Paroda, R.S. (1984). Line x tester analysis for fixed effect model in cotton (Gossypium hirsutum L.). *TAG Theoretical and Applied Genetics*, 68(6), 487-491.
18. Lee, K. and Kim, B.Y. (2009). Both institutions and policies matter but differently for different income groups of countries: determinants of long-run economic growth revisited. *World Development*, 37(3), 533-549.
19. Massa, I. (2011). *Impact of multilateral development finance institutions on economic growth*. London: ODI.
20. Valeriani, E. and Peluso, S. (2011). The impact of institutional quality on economic growth and development: An empirical study. *Journal of Knowledge Management, Economics and Information Technology*, 1(6), 1-25.
21. Moral-Benito, E. (2012). Determinants of economic growth: a Bayesian panel data approach. *Review of Economics and Statistics*, 94(2), 566-579.
22. Vieira, F.V. and MacDonald, R. (2012). A panel data investigation of real exchange rate misalignment and growth. *Estudos Econômicos (São Paulo)*, 42(3), 433-456.
23. Mauro, P. (1995). Corruption and growth. *The Quarterly Journal of Economics*, 110(3), 681-712.
24. Nawaz, S. (2015). Growth effects of institutions: A disaggregated analysis. *Economic Modelling*, 45, 118-126.
25. North, D.C. (1987). Institutions, transaction costs and economic growth. *Economic Inquiry*, 25(3), 419-428.
26. North, D.C. and Thomas, R.P. (1973). *The rise of the western world: A new economic history*. Cambridge University Press.
27. Siddiqui, D.A. and Ahmed, Q.M. (2013). The effect of institutions on economic growth: A global analysis based on GMM dynamic panel estimation. *Structural Change and Economic Dynamics*, 24, 18-33.
28. Nawaz, S., Iqbal, N. and Khan, M.A. (2014). The impact of institutional quality on economic growth: Panel evidence. *The Pakistan Development Review*, 15-31.
29. Barro, R.J. (1998). *Human capital and growth in cross-country regressions*. Harvard University.
30. Yıldırım, A. and Gökalp, M.F. (2016). Institutions and economic performance: A review on the developing countries. *Procedia Economics and Finance*, 38, 347-359.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 49-61

ON ESTIMATION OF REFLECTED WEIBULL DISTRIBUTION USING BAYESIAN ANALYSIS UNDER NON-INFORMATIVE PRIOR

Umme Kalsoom, Wajiha Nasir and **Alina Syed**
Department of Statistics, GC Women University, Sialkot, Pakistan
Email: ummekalsoom894@gmail.com
wajiha.nasir@gcwus.edu.pk
alinasyed97@gmail.com

ABSTRACT

In this paper, we have studied Reflected Weibull Distribution by using Bayesian analysis. Reflected Weibull Distribution has obtained by little transformation in Weibull Distribution. We have estimated the scale parameter of Reflected Weibull Distribution under Bayesian analysis by using Non-Informative prior and derived its Posterior distribution. We have used five loss function to study the performance of Posterior distribution under Uniform and Jeffreys prior. Simulation study has been utilized to check the performance of Bayes estimator and their corresponding risks by using different sample sizes. We have also drawn the graph of posterior distribution for different values of scale parameter.

1. INTRODUCTION

The Reflection Weibull Distribution was introduced by Cohen (1973). It has been shown the Reflected Weibull Distribution is widely employed as a model in reliability and in life testing applications. The modifications of this distribution consist of transformation of a continuous Weibull variate. It originates in the following linear transformation of classical Weibull variate X:

$$Y = -X$$

The probability density function of Reflected Weibull Distribution is as follows:

$$f = \alpha\theta(-y)^{\theta-1}\, exp\{-\alpha(-y)^{\theta}\}, -\infty < y < 0, \theta > 0$$

where $\alpha\ is$ scale parameter and $\theta\ is$ shape parameter.

It has also a version with three parameters. The probability density function of Reflected Weibull distribution with three parameters is as follows:

$$f(y; a, \theta) = \frac{\theta}{b}\left(\frac{a-y}{b}\right)^{\theta-1} exp\left\{-\left(\frac{a-y}{b}\right)^{\theta}\right\}, y < a, b\ \theta > 0$$

When α is equal to zero, then Reflected Weibull distribution reduces to two parameters distribution. This is also known as a type-III extreme value distribution.

In this study the Bayes Estimator and Posterior Risks of scale parameter are derived under Square Error Loss Function, Precautionary Loss Function, DeGroot Loss Function, Weighted Loss Function and Mean Square Loss Function using Uniform and Jeffreys Prior.

Cohen (1973), has introduced Reflected Weibull Distribution. He has proposed that when the Weibull distribution is reflected about the vertical axis y = -x, the Reflected Weibull Distribution is obtained. He has also included an illustrative example to demonstrate the use of the Reflected Weibull in fitting sample data and to provide comparisons with results obtained from fitting the Type VI, Type III, Type V and lognormal distribution. Feroze and Aslam (2012), have studied Bayesian analysis of Inverse Reyleigh distribution under singly and doubly type II censored data. They have derived Bayes estimator and corresponding risks under Non-Informative prior by using symmetric and asymmetric loss function. They have used inverse transformation method for simulation. They have used simulation study for analyze the properties of estimators by using different samples sizes and various parametric values. Feroze and Aslam (2012), have estimated the scale parameter of Log Gamma Distribution. They have used Bayesian and maximum likelihood framework. They have used uniform and Jeffreys priors to derived posterior distribution and also derived Bayes estimators and associated risks under five different loss function. They have used simulation study to derived numerical application of the results and have to compared the performance of different estimators. The purpose of this study is to compare the performance of the estimators based on Bayesian and maximum likelihood frameworks. Feroze and Aslam (2012), have estimated scale parameter of error function distribution using four different loss function. They have used six informative and non-informative priors to derived corresponding posterior distributions. They have derived Bayes estimators and associated risks under different loss functions. They have constructed Bayesian credible intervals under each prior. Feroze and Aslam (2012), have used Gumbel type II distribution to found out the appropriate combination of prior distribution and loss function. They have used the assumption of two non-informative priors and four loss functions (symmetric and asymmetric) to estimated parameters. They have also evaluated posterior predictive intervals. Feroze and Aslam (2012), have studied posterior analysis of exponentiated gamma distribution for type II censored samples. They have derived Bayes estimators and associated risks using different prior. They have used Entropy and Quadratic loss functions for estimation. Feroze and Aslam (2012), have estimated parameter of Burr type XI distribution. For this purpose, they have studied posterior risk under the assumption of eight priors (informative, non-informative, single and mixture of priors). They have used entropy and precautionary loss functions for estimation. Feroze and Aslam (2012), have studied Bayesian analysis of Burr type VII distribution. They have used censored data (three censuring schemes, namely, left censuring, singly type II censuring and doubly type II censuring) to estimate posterior risks and also used two non-informative priors (uniform and Jeffreys) to derived posterior distributions. Feroze and Aslam (2012), have studied Bayesian analysis of scale parameter of inverse Gaussian distribution. They used various loss function and class of prior to derived posterior risks. Nasir et al. (2015), have estimated scale parameter of Log logistic distribution using Bayesian approach. They have used non-informative prior to derived Posterior distribution. They have used different loss function to derive the Bayes estimators and their corresponding risk. Nasir and Aslam (2015), have studied Frechet distribution under Bayesian paradigm. They have used Gumbel Type-II and Levy prior to derived posterior distribution. They have used four loss function to obtained Bayes estimators and their risks. Feroze et al. (2015), have estimated parameters of single and mixture of Rayleigh distribution under type-II double censoring. They have used informative and non-informative prior under squared error loss function and k-loss function assumed for posterior estimation. Nasir (2017), have studied Frechet distribution by using Bayesian analysis. She has derived Posterior analysis by using Gamma and Exponential. She has also derived Bayes estimators and their Posterior risks by using five different loss function. Feroze and Aslam (2018), have studied Bayes estimation of the parameter of Generalized Gamma (GG) distribution using different priors and loss function. They have used WinBugs to obtain numerical solutions for point and interval estimators of parameters.

2. POSTERIOR DISTRIBUTION

We have used two Non-Informative prior (Uniform, Jeffreys) to estimate the scale parameter of posterior distribution. The posterior distribution of Non-Informative prior are given below.

2.1 Posterior Distribution using Uniform Prior

The uniform prior for parameter α is define as

$$p(\alpha) \propto 1 \ ; \ 0 < \alpha < \infty$$

The posterior distribution for parameter α is

$$p(\alpha \mid y) = \frac{\left(\sum_{i=1}^{n} (-y_i)^{\theta}\right)^{n+1} \alpha^{(n+1)-1} e^{-\alpha \sum_{i=1}^{n} (-y_i)^{\theta}}}{\Gamma(n+1)}$$

and So $p(\alpha \mid y) \sim G(\alpha_1, \beta_1)$, where $\alpha_1 = n+1$ and $\beta_1 = \sum_{i=1}^{n} (-y_i)^{\theta}$

The Bayes estimator using SELF is

$$\alpha_{SELF} = \frac{n+1}{\sum_{i=1}^{n} (-y_i)^{\theta}}$$

The Posterior Risk using SELF is

$$P\left(\alpha_{SELF}\right) = \frac{n+1}{\left(\sum_{i=1}^{n} (-y_i)^{\theta}\right)^2}$$

The Bayes estimator using PLF is

$$\alpha_{PLF} = \frac{\sqrt{(n+1)(n+2)}}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using PLF is

$$P(\alpha_{PLF}) = \frac{2\left(\sqrt{(n+1)(n+2)}-(n+1)\right)}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Bayes estimator using DELF is

$$\alpha_{DeLF} = \frac{(n+2)}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using DELF is

$$P(\alpha_{DeLF}) = \frac{1}{n+2}$$

The Bayes estimator using WLF is

$$\alpha_{WLF} = \frac{n}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using WLF is

$$P(\alpha_{WLF}) = \frac{1}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Bayes estimator using MELF is

$$\alpha_{MELF} == \frac{n-1}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using MELF is

$$P(\alpha_{MELF}) = \frac{1}{n}$$

2.2 Posterior Distribution using Jeffreys Prior

The Jeffreys Prior for parameter α is defined as

$$p(\alpha) \propto \frac{1}{\alpha}$$

The Posterior distribution for parameter α is

$$p(\alpha \mid y) = \frac{\left(\sum_{i=1}^{n}(-y_i)^{\theta}\right)^{n} \alpha^{n-1} e^{-\alpha \sum_{i=1}^{n}(-y_i)^{\theta}}}{\Gamma(n)}$$

So $p(\alpha \mid y) \sim G(\alpha_2, \beta_2)$, where $\alpha_2 = n$ and $\beta_2 = \sum_{i=1}^{n}(-y_i)^{\theta}$

The Bayes estimator using SELF is

$$\alpha_{SELF} = \frac{n}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using SELF is

$$P(\alpha_{SELF}) = \frac{n}{\left(\sum_{i=1}^{n}(-y_i)^{\theta}\right)^{2}}$$

The Bayes estimator using PLF is

$$\alpha_{PLF} = \frac{\sqrt{n(n+1)}}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using PLF is

$$P(\alpha_{PLF}) = \frac{2\left(\sqrt{n(n+1)} - n\right)}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Bayes estimator using DELF is

$$\alpha_{DELF} = \frac{n+1}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using DELF is

$$P(\alpha_{DELF}) = \frac{1}{n+1}$$

The Bayes estimator using WLF is

$$\alpha_{WLF} = \frac{n-1}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using WLF is

$$P\left(\alpha_{WLF}\right) = \frac{1}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Bayes estimator using MELF is

$$\alpha_{MELF} = \frac{n-2}{\sum_{i=1}^{n}(-y_i)^{\theta}}$$

The Posterior Risk using MELF is

$$P\left(\alpha_{MELF}\right) = \frac{1}{n-1}$$

3. SIMULATION

The simulation study has been conducted to evaluate the behavior and performance of different estimators. The purpose of the simulation study is to compare the magnitude of Bayes estimators and Posterior risk under different priors and loss functions. If the estimator has minimum posterior risk, then consider it better among all estimators. In this paper, we have used simulation study to check the performance of priors under different five loss function. For this purpose, we can repeat simulation process 10,000 times considering generating of random sample of size 30, 70, 100 and 200 from Reflected Weibull distribution by using α = 3, 5, 7 and β =2, 5, 8.

Table 1
Simulation of Bayes Estimator and Posterior Risk using Uniform Prior

	N	SELF	PLF	DELF	WLF	MELF
$\alpha = 3, \beta = 2$	30	3.2092 (0.3319)	3.2517 (0.1024)	3.3116 (0.0313)	3.0929 (0.1031)	2.9943 (0.0333)
	70	3.0886 (0.1341)	3.1014 (0.0432)	3.1348 (0.0139)	3.0456 (0.0435)	2.9941 (0.0143)
	100	3.0637 (0.0928)	3.0739 (0.0302)	3.0934 (0.0098)	3.0271 (0.0303)	2.9967 (0.01)
	200	3.0315 (0.0457)	3.0381 (0.0151)	3.0408 (0.0050)	3.0175 (0.0151)	3.0026 (0.005)
$\alpha = 3, \beta = 5$	30	3.1995 (0.3329)	3.2719 (0.1027)	3.3076 (0.0313)	3.0968 (0.1032)	3.0043 (0.0333)
	70	3.0854 (0.1343)	3.1089 (0.0433)	3.1259 (0.0139)	3.0492 (0.0435)	2.9993 (0.0143)
	100	3.0579 (0.0928)	3.0791 (0.0302)	3.0914 (0.0098)	3.0326 (0.0303)	3.0045 (0.01)
	200	3.0301 (0.0457)	3.0362 (0.0150)	3.0436 (0.0050)	3.0176 (0.0151)	2.9994 (0.005)
$\alpha = 3, \beta = 8$	30	3.2129 (0.3338)	3.2627 (0.1027)	3.3016 (0.0313)	3.1047 (0.1035)	2.9989 (0.0333)
	70	3.0892 (0.1346)	3.1093 (0.0433)	3.1375 (0.0139)	3.0420 (0.0435)	3.0039 (0.0143)
	100	3.0604 (0.0928)	3.0796 (0.0302)	3.0885 (0.0098)	3.0303 (0.0303)	2.9995 (0.01)
	200	3.0296 (0.0457)	3.0386 (0.0151)	3.0471 (0.0050)	3.0107 (0.0151)	3.0018 (0.005)
$\alpha = 5, \beta = 2$	30	5.3507 (0.9224)	5.4558 (0.1708)	5.5211 (0.0313)	5.1707 (0.1724)	4.9985 (0.0333)
	70	5.1455 (0.3705)	5.1815 (0.0721)	5.2116 (0.0139)	5.0784 (0.0725)	5.0109 (0.0143)
	100	5.1030 (0.2575)	5.1258 (0.0503)	5.1474 (0.0098)	5.0568 (0.0506)	4.9940 (0.01)
	200	5.0541 (0.1267)	5.0596 (0.0251)	5.0744 (0.0050)	5.0259 (0.0251)	5.0029 (0.005)
$\alpha = 5, \beta = 5$	30	5.3384 (0.9250)	5.4134 (0.1705)	5.5097 (0.0313)	5.1568 (0.1719)	5.0121 (0.0333)
	70	5.1386 (0.3723)	5.1818 (0.0721)	5.2138 (0.0139)	5.0892 (0.0726)	5.0019 (0.0143)
	100	5.0915 (0.2575)	5.1202 (0.0503)	5.1554 (0.0098)	5.0414 (0.0504)	4.9934 (0.01)
	200	5.0562 (0.1266)	5.0643 (0.0251)	5.0843 (0.0050)	5.0255 (0.0251)	4.9983 (0.005)

	N	SELF	PLF	DELF	WLF	MELF
$\alpha = 5, \beta = 8$	30	5.3566 (0.9234)	5.4217 (0.1708)	5.5295 (0.0313)	5.1587 (0.1717)	5.0112 (0.0333)
	70	5.1341 (0.3744)	5.1793 (0.0722)	5.2292 (0.0139)	5.0791 (0.0726)	4.9916 (0.0143)
	100	5.1036 (0.2572)	5.1204 (0.0503)	5.1512 (0.0099)	5.0464 (0.0505)	4.9877 (0.01)
	200	5.0528 (0.1270)	5.0646 (0.0251)	5.0746 (0.0050)	5.0223 (0.0251)	4.9992 (0.005)
$\alpha = 7, \beta = 2$	30	7.4771 (1.8155)	7.5759 (0.2386)	7.7332 (0.0313)	7.2494 (0.2416)	7.0140 (0.0333)
	70	7.1923 (0.7352)	7.2537 (0.1010)	7.3096 (0.0139)	7.1080 (0.1014)	7.0052 (0.0143)
	100	7.1412 (0.5056)	7.1766 (0.0705)	7.2110 (0.0098)	7.0662 (0.0707)	6.9944 (0.01)
	200	7.0755 (0.2491)	7.0937 (0.0352)	7.1069 (0.0050)	7.0375 (0.0352)	6.9979 (0.005)
$\alpha = 7, \beta = 5$	30	7.4764 (1.8673)	7.6134 (0.2398)	7.7236 (0.0313)	7.2101 (0.2403)	7.0022 (0.0333)
	70	7.1922 (0.7316)	7.2410 (0.1009)	7.3031 (0.0139)	7.0882 (0.1013)	7.0104 (0.0143)
	100	7.1342 (0.5073)	7.1801 (0.0706)	7.2205 (0.0098)	7.0775 (0.0708)	7.0071 (0.01)
	200	7.0701 (0.2489)	7.0805 (0.0351)	7.1046 (0.0050)	7.0394 (0.0352)	6.9965 (0.005)
$\alpha = 7, \beta = 8$	30	7.4538 (1.8133)	7.5900 (0.2391)	7.7488 (0.0313)	7.2249 (0.2408)	6.9935 (0.0333)
	70	7.2055 (0.7305)	7.2516 (0.1010)	7.2929 (0.0139)	7.0871 (0.1012)	6.9919 (0.0143)
	100	7.1378 (0.5042)	7.1689 (0.0705)	7.2098 (0.0098)	7.0835 (0.0707)	7.0096 (0.01)
	200	7.0638 (0.2482)	7.0865 (0.0351)	7.1023 (0.0050)	7.0364 (0.0352)	7.0020 (0.005)

Table 2
Simulation of Bayes Estimator and Posterior Risk using Jeffreys Prior

	N	SELF	PLF	DELF	WLF	MELF
$\alpha=3, \beta=2$	30	3.1004 (0.3443)	3.1545 (0.1026)	3.1974 (0.0323)	2.9957 (0.1033)	2.8988 (0.0345)
	70	3.0412 (0.1363)	3.0638 (0.0433)	3.0800 (0.0141)	3.0062 (0.0436)	2.9573 (0.0145)
	100	3.0306 (0.0939)	3.0420 (0.0302)	3.0612 (0.0099)	3.0030 (0.0303)	2.9687 (0.0101)
	200	3.0162 (0.0460)	3.026 (0.0151)	3.0304 (0.0050)	3.0020 (0.0151)	2.9811 (0.0050)
$\alpha=3, \beta=5$	30	3.1058 (0.3420)	3.1583 (0.1031)	3.2028 (0.0323)	3.0094 (0.1038)	2.9056 (0.0345)
	70	3.0437 (0.1360)	3.0621 (0.0433)	3.0887 (0.0141)	3.0011 (0.0436)	2.9569 (0.0145)
	100	3.0312 (0.0935)	3.0435 (0.0303)	3.0593 (0.0099)	3.0014 (0.0303)	2.9692 (0.0101)
	200	3.0165 (0.0459)	3.0235 (0.0151)	3.0284 (0.0050)	3.0012 (0.0151)	2.9804 (0.0050)
$\alpha=3, \beta=8$	30	3.1091 (0.3448)	3.1580 (0.1028)	3.2100 (0.0322)	3.0018 (0.1035)	2.8977 (0.0345)
	70	3.0471 (0.1363)	3.0636 (0.0433)	3.0800 (0.0141)	3.0029 (0.0435)	2.9610 (0.0145)
	100	3.0302 (0.0937)	3.0480 (0.0303)	3.0574 (0.0099)	3.0025 (0.0303)	2.9683 (0.0101)
	200	3.0147 (0.0459)	3.0220 (0.0151)	3.0316 (0.0050)	3.0008 (0.0151)	2.9860 (0.0050)
$\alpha=5, \beta=2$	30	5.1706 (0.9566)	5.2524 (0.1718)	5.3430 (0.0323)	4.9906 (0.1721)	4.8212 (0.0345)
	70	5.0562 (0.3783)	5.0950 (0.0722)	5.1470 (0.0141)	4.9905 (0.0723)	4.9179 (0.0145)
	100	5.0498 (0.2605)	5.0713 (0.0504)	5.1063 (0.0099)	4.9926 (0.0504)	4.9451 (0.0101)
	200	5.0194 (0.1277)	5.0338 (0.0251)	5.0487 (0.0050)	4.9987 (0.0251)	4.9711 (0.0050)
$\alpha=5, \beta=5$	30	5.1734 (0.9524)	5.2592 (0.1710)	5.3405 (0.0323)	4.9975 (0.1723)	4.8226 (0.0345)
	70	5.0688 (0.3772)	5.1011 (0.0722)	5.1477 (0.0141)	5.0096 (0.0727)	4.9320 (0.0145)
	100	5.0486 (0.2593)	5.0750 (0.0504)	5.1002 (0.0099)	5.0052 (0.0506)	4.9563 (0.0101)
	200	5.0192 (0.1278)	5.0354 (0.0251)	5.0467 (0.0050)	5.0072 (0.0252)	4.9721 (0.0050)

	N	SELF	PLF	DELF	WLF	MELF
$\alpha = 5, \beta = 8$	30	5.1731 (0.9590)	5.2637 (0.1712)	5.3510 (0.0323)	4.9782 (0.1720)	4.8339 (0.0345)
	70	5.0824 (0.3766)	5.1086 (0.0722)	5.1486 (0.0141)	5.0186 (0.0727)	4.9204 (0.0145)
	100	5.0460 (0.2605)	5.0789 (0.0504)	5.1158 (0.0099)	4.9970 (0.0505)	4.9416 (0.0101)
	200	5.0276 (0.1276)	5.0367 (0.0251)	5.0537 (0.0050)	4.9962 (0.0251)	4.9754 (0.0050)
$\alpha = 7, \beta = 2$	30	7.2475 (1.8693)	7.3415 (0.2388)	7.4925 (0.0323)	7.0054 (0.2416)	6.7591 (0.0345)
	70	7.1208 (0.7392)	7.1425 (0.1011)	7.2098 (0.0141)	6.9955 (0.1015)	6.8998 (0.0145)
	100	7.0745 (0.5101)	7.1065 (0.0705)	7.1384 (0.0099)	7.0064 (0.0708)	6.9141 (0.0101)
	200	7.0402 (0.2503)	7.0580 (0.0352)	7.0675 (0.0050)	6.9953 (0.0352)	6.9735 (0.0050)
$\alpha = 7, \beta = 5$	30	7.2358 (1.8672)	7.3769 (0.2399)	7.4723 (0.0323)	6.9831 (0.2408)	6.7564 (0.0345)
	70	7.1033 (0.7397)	7.1452 (0.1010)	7.2203 (0.0141)	6.9865 (0.1013)	6.9072 (0.0145)
	100	7.0859 (0.5091)	7.1250 (0.0707)	7.1440 (0.0099)	6.9951 (0.0707)	6.9313 (0.0101)
	200	7.0378 (0.2499)	7.0550 (0.0351)	7.0686 (0.0050)	7.0004 (0.0352)	6.9667 (0.0050)
$\alpha = 7, \beta = 8$	30	7.2478 (1.8564)	7.3737 (0.2398)	7.4862 (0.0323)	7.0185 (0.2520)	6.7559 (0.0345)
	70	7.0996 (0.7416)	7.1467 (0.1011)	7.2003 (0.0141)	6.9892 (0.1013)	6.9110 (0.0145)
	100	7.0645 (0.5095)	7.1167 (0.0706)	7.1408 (0.0099)	7.0000 (0.0708)	6.9345 (0.0101)
	200	7.0281 (0.2495)	7.0473 (0.0351)	7.0655 (0.0050)	7.0006 (0.0352)	6.9629 (0.0050)

4. GRAPHS OF POSTERIOR DISTRIBUTION

In this graph, we check the behavior of Posterior distribution by using different values of scale parameter.

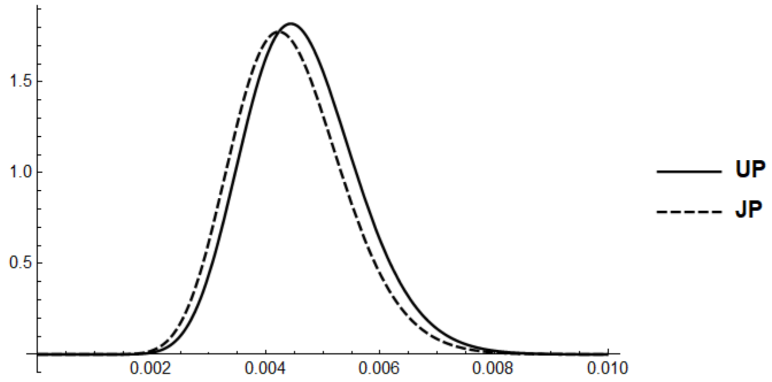


Figure 1: Posterior Distribution When α = 3

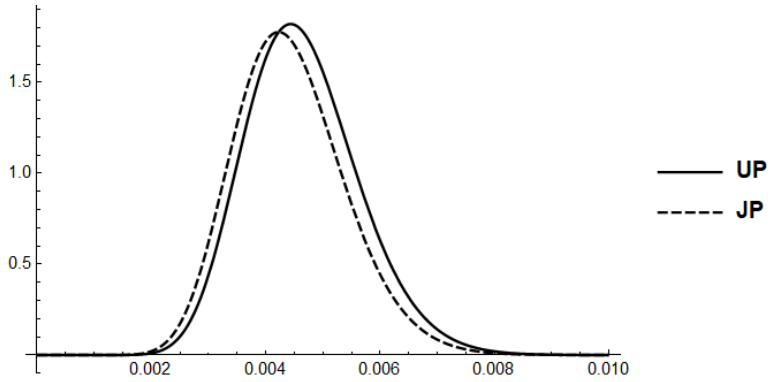


Figure 2: Posterior Distribution When α = 5

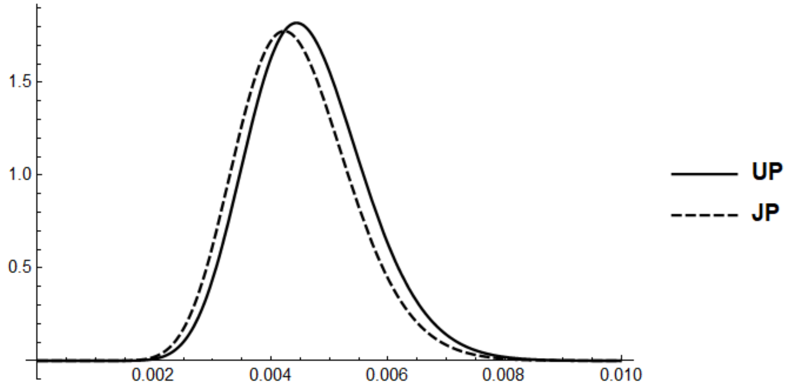


Figure 3: Posterior Distribution When α = 7

Figure 1, 2 and 3 shows the result of posterior distribution. The above graphs show that posterior distribution under non-Informative priors (Uniform and Jeffreys prior) are slightly positively skewed. The behavior of posterior distribution under Uniform and Jeffreys prior is almost same.

5. CONCLUSION

In this paper, we have used simulation study to check the behavior of scale parameter of Reflected Weibull distribution. After simulation and comparing different loss function, we conclude that by increasing sample size Bayes estimators approaches to their true value of parameter and Posterior risks is decreases. Uniform prior is performing better than Jeffrey prior as its posterior risk is minimum. DELF (DeGroot Loss Function) is performing better as its posterior risk is minimum among all other loss functions. So, Uniform prior with DELF (DeGroot loss function) provides minimum Posterior risks as compared to other loss function and priors. The graphs show that posterior distribution under Non-Informative prior are positively skewed.

REFERENCES

1. Aslam, M. (2003). An application of prior predictive distribution to elicit the prior density. *Journal of Statistical Theory and Applications*, 2(1), 70-83.
2. Berger, J.O. (1985). *Statistical decision theory and Bayesian analysis*. Springer.
3. Bernardo, J.M. (2011). Bayesian statistics. In *International Encyclopedia of Statistical Science*, pages 107-133. Springer.
4. Bolstad, W.M. (2007). *Introduction to Bayesian statistics*. John Wiley & Sons.
5. Box, G.E. and Tiao, G.C. (2011). *Bayesian inference in statistical analysis*, Vol. 40, John Wiley & Sons.
6. Cohen, A.C. (1973). The reflected Weibull distribution. *Technometrics*, 15(4), 867-873.
7. DeGroot, M. H. (2005). *Optimal statistical decisions*, volume 82. John Wiley & Sons.
8. Feroze, N. (2012). Estimation of scale parameter of inverse Gaussian distribution under a Bayesian framework using different loss functions. *Scientific Journal of Review*, 1(3), 39-52.
9. Feroze, N. and Aslam, M. (2012). A note on Bayesian analysis of error function distribution under different loss functions. *International Journal of Probability and Statistics*, 1(5), 153-159.
10. Feroze, N. and Aslam, M. (2012). A Note on Bayesian and Maximum Likelihood Estimation of Scale Parameter of Log Gamma Distribution. *International Journal of Statistics and Applications*, 2(5), 73-79.
11. Feroze, N. and Aslam, M. (2012). Bayesian Analysis of Burr Type XI Distribution under Single and Mixture of Priors. Advances in Agriculture, Sciences and Engineering Research, 2(11), 487-502.
12. Feroze, N. and Aslam, M. (2012). Bayesian analysis of exponentiated gamma distribution under type II censored samples. *Sci. J. Pure. Appl. Sci.*, 1(1), 30-39.
13. Feroze, N. and Aslam, M. (2012). On Bayesian analysis of Burr type VII distribution under different censoring schemes. *International Journal of Quality, Statistics, and Reliability*, 2012.

14. Feroze, N. and Aslam, M. (2012). On posterior analysis of inverse Rayleigh distribution under singly and doubly type II censored data. *International Journal of Probability and Statistics*, 1(5), 145-152.
15. Feroze, N. and Aslam, M. (2012). On Posterior Analysis of mixture of two components of Gumbel Type II Distribution. *International Journal of Probability and Statistics*, 1(4), 119-132.
16. Feroze, N. and Aslam, M. (2018). A Family of Bayes Estimators for the Parameters of the Generalized Gamma Distribution. *Pakistan Journal of Statistics and Operation Research*, 14(4), 975-994.
17. Feroze, N., Sindhu, T.N., Aslam, M. and Yaseen, W. (2015). Doubly censored mixture of two Rayleigh distributions: Properties and applications. *Maejo International Journal of Science and Technology*, 9(3), 312-327.
18. Gelman, A. (2004). *Bayesian data analysis*. Chapman Hall/CRC, Boca Raton, Fla.
19. Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007), 453-461.
20. Mood, A.M. and Graybill, F.A. (1950). *Introduction to the Theory of Statistics*, New York: McGraw-H ill Book Company.
21. Norstrom, J.G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Transactions on Reliability*, 45(3), 400-403.
22. Samaniego, F.J. (2010). *A comparison of the Bayesian and frequentist approaches to estimation*. Springer.
23. Tummala, V. and Sathe, P.T. (1978). Minimum expected loss estimators of reliability and parameters of certain lifetime distributions. *IEEE Transactions on Reliability*, 27(4), 283-285.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 63-92

FACTORS AFFECTING THE PERFORMANCE OF HEADS OF GOVERNMENT SECONDARY SCHOOLS IN PUNJAB

Muhammad Yousuf[1] and **Ali Shan**[2]

[1] Department Education, Allama Iqbal Open University, Islamabad, Pakistan. Email: epmaioupk@yahoo.co.uk

[2] Department of Statistics, Allama Iqbal Open University, Islamabad, Pakistan. Email: mirza.ali.shan685@gmail.com

ABSTRACT

This study critically examines the factors affecting the performance of secondary hoods of government school in Punjab. For this purpose, a critical review of existing literature was carried out. The study found that there are so many internal and external factors which directly or indirectly affect the performance of government secondary school heads in Punjab province. Secondary Education is affected by different factors. There is lack of physical facilities. There is political interference in the system which breeds corruption, favoritism and nepotism. The system of supervision and monitoring is weak. There is no an effective mechanism for teacher training. If they are tried to train, they do not implement the learnt educational methodologies and techniques. There is poor system of accountability. Teachers are professionally less motivated. They feel that their socio-economic system is poor. There is a problem of medium of instruction. Parents are less involved in the educational process which affects the results of the students. Teachers' performance depends on the performance of the results. The performance of school heads depends on the performance of their teachers including other factors. On the basis of this study, it is recommended that proper heed should be given to the factors which negatively or positively affect the performance of school heads.

KEYWORDS

Secondary Education, Factors, Management, Leadership, Performance, Environment, Efficiency, Supervision, Monitoring, Impact.

INTRODUCTION

Terry & Franklin, (2002) describe, "Management is a different process consisting of activities of planning, organizing, and controlling performed to establish and accomplish stated objectives with the use of human beings and other resources."

Jones, (2004) opines that both education and experience enable managers to recognize and develop the skills they need to put organizational resources to their best use.

Adeyemi, (2010) says that here is close relationship between the leadership styles of heads and teachers' performance in secondary schools. There is also a direct relationship between leadership styles used by heads and their performance. Malik, Danish, &Usman, (2010) give their view that in an organization success can only be achieved by the satisfied and self-motivated employees and good leadership and professional vision of a school head.

Griffin, (2005) says that the performance of an individual is determined by three factors i.e. professional motivation, work environment and ability to do work. The workplace environment impacts on an employee’s morale, productivity and job performance both positively and negatively. If the work place environment is not liked by the employees so they get de-motivated and their performance also affected.

STATEMENT OF THE PROBLEM

The problems of secondary school education and school heads in Punjab province of Pakistan can be traced back to the very inception of the country. The system could not be reformed according to the needs and social genre of the people (Sheikh and Rasool, 1998). In other words, the performance of school heads depends upon many internal and external factors. This study attempts to analyze the factors affecting the performance of secondary school heads of government schools in Punjab.

RESEARCH QUESTIONS

Following questions were designed to explore the problem and to solve the issue.

1. What are the different types of factors affecting the performance of heads of government secondary schools?
2. What is the ranking order of factors affecting the performance of heads of government secondary schools?

OBJECTIVES OF THE STUDY

The objectives of the study were:

1. To identify the factors affecting the performance of heads of government secondary schools.
2. To split into different categories the factors affecting the performance of heads of government secondary schools.
3. To compare the factors affecting the performance of the heads of rural and urban government secondary schools.
4. To prepare a list of factors in ranking order affecting the performance of heads of government secondary schools.

SIGNIFICANCE OF THE STUDY

The proposed study may be helpful to improve their managerial skills which may be helpful for the improvement and progress of their institutions. The study may be of significance because its findings may provide the guidelines to the policy makers, educational planners, and educational administrators at provincial level for the betterment of educational set up at secondary level.

REVIEW OF RELATED LITERATURE

The present study aimed at investigating into different factors affecting the performance of heads of government secondary schools in Punjab. It is believed that performance in a school depends upon the indicators of school performance. Assessment of performance through indicators is an actuality. It consists of administration, school council, physical facilities, co-curricular activities, school environment, and many other internal and external factors.

Khan (2000), highlights some of the serious problems faced by the heads of government secondary schools in Pakistan and to examine major policy interventions by the Government of Pakistan to tackle them.

The quality of education at secondary level depends upon many factors internal and external as well. Most important of them is the optimal utilization of available human and physical resources, which has direct influence on the performance of school heads. Since the inception of Pakistan, different national education policies have been implemented with a view to improve the quality of education in the country which consequently depends upon school performance (Chande, 2006, Shami & Shah, 2005).

Ayub (2001), indicated that parents' involvement in the educational activities had positive impact on the achievements of their children. It had also been found that the parents' and family environment were important factors responsible for improving the achievement level of students in schools.

Penkajam (2005), states that education is an important tool to bring about socio-economic transformation of any society. There is a vital role of academic institutions, qualified staff, and facilities which are helpful in the uplift of school students.

Ahmed (2005), claims that teachers are the builders of personality of the students from kindergarten to the highest level of education. In fact the destiny of a nation depends upon teachers who could make maximum contribution to the proper utilization of national human resources through quality education.

James & Connolly (2000), explained the effectiveness of change in school with respect to improvement of school projects. It is an important aim of the project to promote and facilitate the sharing of experiences among schools and the results showed.

Drop out is defined as a student who left an educational institution for any reason except death, before completion of a programme of studies and without transferring to another school. The term is used to designate an elementary or secondary school student who has been in membership during the regular academic year (Ahmed, 2006).

Stephen (2012), says that the most outstanding factor which affects the performance of a secondary school head is his/her administrative vision. The basic cause of inefficiency of any institution is lack in administration and incompetent vision of its head. Head's vision, professional competency and administrative abilities directly affect the school process and school results.

Iqbal (2013), concludes that the process of school supervision in Pakistani schools is ineffective and is considered to be more a dictation than an improvement. There are some deep rooted assumptions concerning the concept of school supervision in Pakistan.

Hameed (2013), says that the influence of political figures, partiality, and community domination are the factors that impede productivity level of educational institutions and a school head. The local political pressure on the school heads creates an obstacle on the way of the learning process. In Pakistan, political interference in educational system is being practiced.

Ashraf (2008), conducted a research study on "Impact of Management on Teacher's Performance at Secondary Level in Punjab," and conclude that heads of institutions should promote teacher’s efficiency and use the resources properly.

Din (2008), says that constructive and optimistic attitude enhance the performance of teachers and school heads. The study also shows that the majority of the heads is very strict in their dealing with the staff. He concludes that stiff dealing of heads, reproach the teachers on their mistakes and discriminative in assigning duties, is the technique that negatively affects the performance of school teachers and their head.

Tareen (2008), concludes that irrespective of difference in education, experience, locality, and management controls majority of primary school teachers have positive attitude towards their educational administrators. It is recommended that educational heads may adopt the administrative style according to the situation.

Sultana (2005), concludes that performance and effectiveness of teachers may be defined as “The attribute of a teacher who has the capabilities, potentials of having a positive impact on students’ learning/ behavior and attitude”.

Ndana (2011), says that administrative challenges that public schools face contends that discipline and teachers’ relations can be the cause of poor performance in those institutions. He observes that teachers play a role in students’ indiscipline.

Wawira (2011), says that school heads' job and teaching experiences influence teachers' perception towards heads' instructional supervision practices. This means that such teachers are readily willing to accept instructional guidance from experienced school heads.

Somech (2006), states that impact of directive leadership on teachers’ performance is contingent in nature. The positive effect of participative leadership on their performance is above and beyond the specific conditions.

Retallick (2005), quotes that the head’s behavior and activities directly affect school processes, schools outcomes, and the performance of a school head.

Onyango (2001), asserts that the management of material resources entails planning, acquisition, allocating, distribution and even controlling maintenance and use of those materials.

RESEARCH METHODOLOGY

The purpose of the present study was to identify "Factors affecting the performance of heads of government secondary schools in Punjab". For carrying out the study, the descriptive research method was the proper choice. This chapter deals with the method of study that covers population, selection sample, development of research tool, their administration and statistical techniques used for data analysis.

3.1 Population

The targeted population consisted of the 277 working male heads of Government Secondary Schools of three districts i.e. Gujrat, Jhelum, and Mandi Baha-ud-Din of the Punjab province. The description of the population is as under:

S#	Name of District	Male Heads Rural	Male Heads Urban	Male Heads Total
1	**Gujrat**	68	70	138
2	**Jhelum**	35	37	072
3	**Mandi Baha-ud-Din**	32	35	067
	G.Total	**135**	**142**	**277**

3.2 Sample

A sample size of 160 (approximately 57% of the population) Heads of Government Secondary School from three districts Gujrat, Jhelum and Mandi Baha-ud-Din was sampled.

The description of sample from each category of population is as under:

Table 3.2.1
Description of Sample
Sample of Secondary Schools heads

S#	Categories of Population	Population	Sample Size	Heads Rural	Heads Urban	% age
1	**Male Heads (Gujrat)**	138	82	41	41	60%
2	**Male Heads (Jhelum)**	072	40	20	20	56%
3	**Male Heads (M.B.D.)**	067	38	19	19	57%

3.3 Framework of Research Tool

In the light of the major identified factors a questionnaire as an instrument for data collection was developed. The questionnaire for the heads of government secondary schools of rural and urban area (comprises 38 open ended and 04 closed ended questions).Questionnaire comprises 10 factors with items and the number of open ended questions comprises 04 items.

3.3.1 Validation of Instrument

Before the actual launch, questionnaire was validated personally to the experts of education. The questionnaire was developed in the light of the related literature. The questionnaire was improved in the light of the observations and suggestions.

3.3.2 Pilot Testing

For pilot testing, 10 Heads were selected from 3 districts of the Punjab which were not included in sample. The mid-point of the tool was 3. To find the reliability of the questionnaire, Cronbach's coefficient Alpha formula was applied by using SPSS Version 17 to estimate the internal consistency of items. The obtained Cronbach's alpha was 0.84.

Table 3.3.3
Scoring of Responses

Description	Strongly Agree	Agree	Undecided	Disagree	Strongly Disagree
Score	5	4	3	2	1

3.4 Collection of Data

The investigator started data collection by visiting each sampled school and administering the questionnaire in person. The response rate was 100 percent.

3.5 Analysis and Interpretation of Data

The collected data were tabulated and interpreted. Data collected from school heads through questionnaire were arranged and analyzed by using Percentage, Mean score and Weighted Mean. Following scale was used to calculate the mean score:

Strongly Agree SA 5 point, Agree A 4 points, Uncertain UNC 3 points,

Disagree, DA 2 points, Strongly Disagree SDA 1 point

Percentage

Percentage was worked out by using the following formula:

$$\% = \frac{F}{N} \times 100$$

Mean

While the mean score of each item was calculated by using the following formula:

$$\textbf{Mean Score} = \frac{\text{FSAX5 + FAX4 + FUNCX3 + FDAX2 + FSDAX1}}{\textbf{N}}$$

Weighted Mean Formula

$$\text{Weighted Mean } = \frac{\sum wx}{\sum w}$$

W = weight of the response

4. DATA ANALYSIS AND ITS INTERPRETATION

This chapter deals with the tabulation, analysis and interpretation of the data. The main objective of the study was to investigate the “Factors Affecting the Performance of Heads of Government Secondary Schools in Punjab”. This chapter consists of the presentation of results based on the data obtained by the researcher.

The data were collected from the Heads of Government Secondary Schools for Boys of three districts i.e. Gujrat, Jhelum and Mandi Baha-ud-Din of the province Punjab. The data were studied as many angles as possible to explore the different factors which affect the performance of a school head. The data obtained from the samples with the help of a questionnaire which was tabulated, analyzed and interpreted in the form of tables. The focus of the data analysis was to reduce and describe large volume of data to produce information that is useful and meaningful for the discussion and the interpretation leading to findings, conclusions and recommendations.

The heads were given a questionnaire to indicate the factors which affect their professional performance. The respondents were treated as overall male heads of Government Secondary Schools of the above mentioned three districts.

The foregoing pages of this study include the major factors which affect the performance of the heads of government secondary schools. Different school heads exchanged their different opinions to be analyzed. The overall analysis of the heads’ opinion and the major factors affecting their performance is as under.

4.1 Analysis based on Rural and Urban Locality

Table 4.1
Professional Vision of a School Head

<table>
<tr><th>Category</th><th>Sector</th><th>Responses</th><th>SA</th><th>A</th><th>UNC</th><th>DA</th><th>SDA</th><th>Total</th><th>Mean</th></tr>
<tr><td rowspan="4">Head</td><td rowspan="2">Urban</td><td>Frequency</td><td>52</td><td>25</td><td>3</td><td>0</td><td>0</td><td>80</td><td rowspan="2">4.2</td></tr>
<tr><td>Percentage</td><td>65%</td><td>31.25%</td><td>3.75%</td><td>0%</td><td>0%</td><td>100%</td></tr>
<tr><td rowspan="2">Rural</td><td>Frequency</td><td>56</td><td>23</td><td>1</td><td>0</td><td>0</td><td>80</td><td rowspan="2">3.9</td></tr>
<tr><td>Percentage</td><td>70%</td><td>28.75%</td><td>1.25%</td><td>0%</td><td>0%</td><td>100%</td></tr>
<tr><td colspan="2">Total</td><td></td><td>108</td><td>48</td><td>4</td><td>0</td><td>0</td><td>160</td><td></td></tr>
</table>

Table 4.1 indicates the opinion of school heads about the professional vision of a school head. 96.25% urban respondents are agreed to the statement, 3.75% respondents are uncertain. The mean score 4.2 also supports the statement whereas 98.75% rural respondents are agreed, 01 % respondents are uncertain. The mean score 3.9 also supports the statement.

Table 4.2
Lack of Vision Affects Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	47	29	2	2	0	**80**	**4.3**
		Percentage	58.75%	36.25%	2.50%	2.50%	0%	**100%**	
	Rural	**Frequency**	45	31	2	2	0	**80**	**3.96**
		Percentage	56.25%	38.75%	2.50%	2.50%	0%	**100%**	
Total			**92**	**60**	**4**	**4**	**0**	**160**	

Table 4.2 shows the opinion of school heads about the Lack of Vision Affects Performance. 95% urban respondents are agreed to the statement, 2.50% respondents are disagreed. The mean score 4.3 also supports the statement whereas 95% rural respondents are agreed, 2.5 % respondents are uncertain and 2.5% respondents are disagreed. The mean score 3.96 also supports the statement.

Table 4.3
Professional Training Influences Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	41	36	0	0	3	**80**	**3.8**
		Percentage	51.25%	45%	0%	0%	3.75%	**100%**	
	Rural	**Frequency**	32	45	0	0	3	**80**	**4.23**
		Percentage	40%	56.25%	0%	0%	3.75%	**100%**	
Total			**73**	**81**	**0**	**0**	**6**	**160**	

Table 4.3 highlights the opinion of school heads about the Professional Training Influences Head's Performance. 96.25% urban respondents are agreed to the statement, 3.75% respondents are disagreed. The mean score 3.8 also supports the statement whereas 96.25% rural respondents are agreed, 3.75% respondents are uncertain are disagreed. The mean score 4.23 also supports the statement.

Table 4.4
Leadership Style Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	29	44	5	2	0	**80**	**4**
		Percentage	36.25%	55%	6.25%	2.50%	0%	**100%**	
	Rural	**Frequency**	26	47	5	2	0	**80**	**4.19**
		Percentage	32.50%	58.75%	6.25%	2.50%	0%	**100%**	
Total			**55**	**91**	**10**	**4**	**0**	**160**	

Table 4.4 indicates the opinion of school heads about the Leadership Style Affects Head's Performance. 91.25% urban respondents are agreed to the statement, 6.25% respondents are uncertain and 2.50% are disagreed. The mean score 4 also supports the statement whereas 91.25% rural respondents are agreed, 6.25 % respondents are uncertain and 2.50% are disagreed. The mean score 4.19 supports the statement.

Table 4.5
Democratic Style Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	30	39	5	5	1	80	4.2
		Percentage	37.50%	48.75%	6.25%	6.25%	1.25%	100%	
	Rural	Frequency	32	39	5	3	1	80	4.23
		Percentage	40%	48.75%	6.25%	3.75%	1.25%	100%	
Total			72	78	10	8	2	160	

Table 4.5 shows the opinion of school heads about the Democratic Style Affects Head's Performance. 86.25% urban respondents are agreed to the statement, 6.25% respondents are uncertain and 7.50% respondents are disagreed. The mean score 4.2 also supports the statement whereas 88.75% rural respondents are agreed, 6.25 % respondents are uncertain and 05% respondents are disagreed. The mean score 4.23 also supports the statement.

Table 4.6
Bureaucratic Style Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	27	43	4	1	5	80	4.2
		Percentage	33.75%	53.75%	5.0%	1.25%	6.25%	100%	
	Rural	Frequency	25	44	7	1	3	80	3.4
		Percentage	31.25%	55%	8.75%	1.25%	3.75%	100%	
Total			52	87	11	2	8	160	

Table 4.6 highlights the opinion of school heads about the Bureaucratic Style Affects Head's Performance. 87.5% urban respondents are agreed to the statement, 5.0% respondents are uncertain and 7.50% respondents are disagreed. The mean score 4.2 also supports the statement whereas 86.25% rural respondents are agreed, 07.0 % respondents are uncertain and 05% respondents are disagreed. The mean score 3.4 also supports the statement.

Table 4.7
Autocratic Style Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	11	43	8	12	6	80	3
		Percentage	13.75%	53.75%	10%	15%	7.5%	100%	
	Rural	Frequency	12	43	7	12	6	80	2.8
		Percentage	15%	53.75%	8.75%	15%	7.5%	100%	
Total			23	86	15	24	12	160	

Table 4.7 indicates the opinion of school heads about the Autocratic Style Affects Head's Performance. 67.25% urban respondents are agreed to the statement, 10.0% respondents are uncertain and 22.5% respondents are disagreed. The mean score 3 also supports the statement whereas 68.75% rural respondents are agreed, 8.75 % respondents are uncertain and 22.5% respondents are disagreed. The mean score 2.8 also supports the statement.

Table 4.8
Lack of Physical Facilities Has Adverse Effect on Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	26	45	0	6	3	**80**	**4.29**
		Percentage	32.5%	56.25%	0%	7.5%	3.75%	**100%**	
	Rural	Frequency	40	31	2	6	1	**80**	**4.29**
		Percentage	50%	38.75%	2.5%	7.5%	1.25%	**100%**	
Total			**66**	**76**	**2**	**12**	**4**	**160**	

Table 4.8 shows the opinion of school heads about the Lack of Physical Facilities Has Adverse Effect on Head's Performance. 85.75% urban respondents are agreed to the statement, and 11.25% respondents are disagreed. The mean score 4.29 also supports the statement whereas 88.75% rural respondents are agreed, 2.5 % respondents are uncertain and 8.75% respondents are disagreed. The mean score 4.29 also supports the statement.

Table 4.9
Lack of Physical Facilities Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	24	32	4	20	0	**80**	**4**
		Percentage	30%	40%	5%	25%	0%	**100%**	
	Rural	Frequency	20	34	2	24	0	**80**	**4**
		Percentage	25%	42.5%	2.5%	30%	0%	**100%**	
Total			**44**	**66**	**6**	**44**	**0**	**160**	

Table 4.9 highlights the opinion of school heads about the Lack of Physical Facilities Affects Head's Performance. 70% urban respondents are agreed to the statement, 5% respondents are uncertain and 25% respondents are disagreed. The mean score 4 also supports the statement whereas 67.50% rural respondents are agreed, 2.5% respondents are uncertain and 30% respondents are disagreed. The mean score 4 also supports the statement.

Table 4.10
Lack of Physical Facilities Affects Educational Process

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	42	34	1	3	0	**80**	**3**
		Percentage	52.5%	42.5%	1.25%	3.75%	0%	**100%**	
	Rural	Frequency	37	39	2	2	0	**80**	**4**
		Percentage	46.25%	48.75%	2.5%	2.5%	0%	**100%**	
Total			**79**	**73**	**3**	**5**	**0**	**160**	

Table 4.10 indicates the opinion of school heads about the Lack of Physical Facilities Affects Educational Process. 95% urban respondents are agreed to the statement, 1.25% respondents are uncertain and 3.75% respondents are disagreed. The mean score 3 also supports the statement whereas 95% rural respondents are agreed, 2.5 % respondents are uncertain and 2.5% respondents are disagreed. The mean score 4 also supports the statement.

Table 4.11
Lack of Learning Resources Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	35	41	3	1	0	**80**	**4.31**
		Percentage	43.75%	51.25%	3.75%	1.25%	0%	**100%**	
	Rural	Frequency	30	46	3	1	0	**80**	**4.32**
		Percentage	37.5%	57.5%	3.75%	1.25%	0%	**100%**	
Total			**65**	**87**	**6**	**2**	**0**	**160**	

Table 4.11 shows the opinion of school heads about the Lack of Learning Resources Affects Head's Performance. 95% urban respondents are agreed to the statement, 3.75% respondents are uncertain and 1.25% respondents are disagreed. The mean score 4.31 also supports the statement whereas 95% rural respondents are agreed, 3.75 % respondents are uncertain and 1.25% respondents are disagreed. The mean score 4.32 also supports the statement.

Table 4.12
Learning Resources Increase Professional Motivation among Teachers

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	36	41	2	1	0	**80**	**4.2**
		Percentage	45%	51.25%	2.5%	1.25%	0%	**100%**	
	Rural	Frequency	35	44	0	1	0	**80**	**4.38**
		Percentage	43.75%	55%	0%	1.25%	0%	**100%**	
Total			**71**	**85**	**2**	**2**	**0**	**160**	

Table 4.12 highlights the opinion of school heads about the Learning Resources Increase Professional Motivation among Teachers. 96.25% urban respondents are agreed to the statement, 2.5% respondents are uncertain and 1.25% respondents are disagreed. The mean score 4.2 also supports the statement whereas 98.75% rural respondents are agreed, and 1.25% respondents are disagreed. The mean score 4.38 also supports the statement.

Table 4.13
Dual Medium of Instruction Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	26	43	7	3	1	**80**	**4.12**
		Percentage	32.5%	53.75%	8.75%	3.75%	1.25%	**100%**	
	Rural	Frequency	37	34	4	3	2	**80**	**4.44**
		Percentage	46.25%	42.5%	5%	3.75%	2.5%	**100%**	
Total			**63**	**77**	**11**	**6**	**3**	**160**	

Table 4.13 indicates the opinion of school heads about the Dual Medium of Instruction Affects Head's Performance. 86.25% urban respondents are agreed to the statement, 8.75% respondents are uncertain and 5% respondents are disagreed. The mean score 4.12 also supports the statement whereas 88.75% rural respondents are agreed, 5%

respondents are uncertain and 6.25% respondents are disagreed. The mean score 4.44 also supports the statement.

Table 4.14
Shortage of Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	51	29	0	0	0	**80**	**4.66**
		Percentage	63.75%	36.25%	0%	0%	0%	**100%**	
	Rural	Frequency	60	19	1	0	0	**80**	**4.66**
		Percentage	75%	23.75%	1.70%	0%	0%	**100%**	
Total			**111**	**48**	**1**	**0**	**0**	**160**	

Table 4.14 shows the opinion of school heads about the Shortage of Teachers Affects Head's Performance. 100% urban respondents are agreed to the statement. The mean score 4.66 also supports the statement whereas 98.75% rural respondents are agreed, 1.75% respondents are uncertain. The mean score 4.66 also supports the statement.

Table 4.15
Shortage of Subject Specialists Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	37	34	3	6	0	**80**	**4.14**
		Percentage	46.25%	42.5%	3.75%	7.50%	0%	**100%**	
	Rural	Frequency	36	39	0	5	0	**80**	**4.66**
		Percentage	45%	48.75%	0%	6.25%	0%	**100%**	
Total			**73**	**73**	**3**	**11**	**0**	**160**	

Table 4.15 highlights the opinion of school heads about the Shortage of Subject Specialists Affects Head's Performance. 88.75% urban respondents are agreed to the statement, 3.75% respondents are uncertain and 7.50% respondents are disagreed. The mean score 4.14 also supports the statement whereas 93.75% rural respondents are agreed, 6.25% respondents are disagreed. The mean score 4.66 also supports the statement.

Table 4.16
Low Professional Motivation among Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	28	48	2	2	0	**80**	**4.12**
		Percentage	35%	60%	2.50%	2.50%	0%	**100%**	
	Rural	Frequency	30	45	2	3	0	**80**	**4.4**
		Percentage	37.50%	56.25%	2.50%	3.75%	0%	**100%**	
Total			**58**	**93**	**4**	**5**	**0**	**160**	

Table 4.16 indicates the opinion of school heads about the Low Professional Motivation among Teachers Affects Head's Performance. 95% urban respondents are agreed to the statement, 2.5% respondents are uncertain and 2.5% respondents are disagreed. The mean score 4.12 also supports the statement whereas 93.75% rural

respondents are agreed, 2.5 % respondents are uncertain and 3.75% respondents are disagreed. The mean score 4.4 also supports the statement.

Table 4.17
Low Motivation among Teachers Affects Head's performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	28	41	5	5	1	80	4.2
		Percentage	35%	51.25%	6.25%	6.25%	1.25%	100%	
	Rural	Frequency	33	38	3	6	0	80	4.3
		Percentage	41.25%	47.50%	3.75%	7.50%	0%	100%	
Total			61	79	8	11	1	160	

Table 4.17 shows the opinion of school heads about the Low Motivation among Teachers Affects Head's performance. 86.25% urban respondents are agreed to the statement, 6.25% respondents are uncertain and 7.50% respondents are disagreed. The mean score 4.2 also supports the statement whereas 88.75% rural respondents are agreed, 3.75 % respondents are uncertain and 7.5% respondents are disagreed. The mean score 4.3 also supports the statement.

Table 4.18
Absenteeism among Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	34	40	2	2	2	80	4.22
		Percentage	42.54%	50%	2.50%	2.50%	2.50%	100%	
	Rural	Frequency	38	36	2	1	3	80	4.25
		Percentage	47.50%	45%	2.50%	1.25%	3.75%	100%	
Total			72	76	4	3	5	160	

Table 4.5 highlights the opinion of school heads about the Absenteeism among Teachers Affects Head's Performance. 92.54% urban respondents are agreed to the statement, 2.5% respondents are uncertain and 5% respondents are disagreed. The mean score 4.22 also supports the statement whereas 92.5% rural respondents are agreed, 2.5 % respondents are uncertain and 05% respondents are disagreed. The mean score 4.25 also supports the statement.

Table 4.19
Weak Coordination of Head with Parents Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	43	31	2	4	0	80	4.01
		Percentage	53.75%	38.75%	2.50%	5%	0%	100%	
	Rural	Frequency	41	38	1	0	0	80	4.34
		Percentage	51.25%	47.50%	1.25%	0%	0%	100%	
Total			84	69	3	4	0	160	

Table 4.19 indicates the opinion of school heads about the Weak Coordination of Head with Parents Affects Head's Performance. 92.5% urban respondents are agreed to

the statement, 2.5% respondents are uncertain and 5% respondents are disagreed. The mean score 4.01 also supports the statement whereas 98.75% rural respondents are agreed, and 1.25 % respondents are uncertain. The mean score 4.34 also supports the statement.

Table 4.20
Less Participation of Parents Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	32	40	2	6	0	**80**	**4.14**
		Percentage	40%	50%	2.50%	7.50%	0%	**100%**	
	Rural	Frequency	35	40	1	4	0	**80**	**4.2**
		Percentage	43.75%	50%	1.25%	5%	0%	**100%**	
Total			**67**	**80**	**3**	**10**	**0**	**160**	

Table 4.20 shows the opinion of school heads about the Less Participation of Parents Affects Head's Performance administrative vision of a school head. 90% urban respondents are agreed to the statement, 2.5% respondents are uncertain and 7.5% respondents are disagreed. The mean score 4.14 also supports the statement whereas 93.75% rural respondents are agreed, 1.25 % respondents are uncertain and 05% respondents are disagreed. The mean score 4.2 also supports the statement.

Table 4.21
Parents' Create Problems for School Head

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	21	44	4	9	2	**80**	**4.12**
		Percentage	26.25%	55%	5%	11.25%	2.50%	**100%**	
	Rural	Frequency	21	41	5	11	2	**80**	**4.31**
		Percentage	26.25%	51.25%	6.25%	13.75%	2.50%	**100%**	
Total			**42**	**85**	**9**	**20**	**4**	**160**	

Table 4.21 highlights the opinion of school heads about the Parents' Create Problems for School Head. 81.25% urban respondents are agreed to the statement, 5% respondents are uncertain and 13.75% respondents are disagreed. The mean score 4.12 also supports the statement whereas 77.5% rural respondents are agreed, 6.25 % respondents are uncertain and 16.25% respondents are disagreed. The mean score 4.31 also supports the statement.

Table 4.22
Local Politics Negatively Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	39	27	4	10	0	**80**	**3.8**
		Percentage	48.75%	33.75%	5%	12.50%	0%	**100%**	
	Rural	Frequency	36	31	2	11	0	**80**	**3.96**
		Percentage	45%	38.75%	2.50%	13.75%	0%	**100%**	
Total			**75**	**58**	**6**	**21**	**0**	**160**	

Table 4.22 indicates the opinion of school heads about the Local Politics Negatively Affects Head’s Performance. 82.5% urban respondents are agreed to the statement, 5% respondents are uncertain and 12.50% respondents are disagreed. The mean score 3.8 also supports the statement whereas 83.75% rural respondents are agreed, 2.5% respondents are uncertain and 13.75% respondents are disagreed. The mean score 3.96 also supports the statement.

Table 4.23
Local Politics Positively Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	43	23	11	1	2	**80**	**3.9**
		Percentage	53.75%	28.75%	27.50%	1.25%	2.50%	**100%**	
	Rural	**Frequency**	28	38	10	3	1	**80**	**4.17**
		Percentage	35%	47.50%	12.50%	3.75%	1.25%	**100%**	
Total			**71**	**61**	**21**	**4**	**3**	**160**	

Table 4.23 shows the opinion of school heads about the Local Politics positively Affects Head's Performance. 82.5% urban respondents are agreed to the statement, 27.5% respondents are uncertain and 3.75% respondents are disagreed. The mean score 3.9 also supports the statement whereas 82.5% rural respondents are agreed, 12.5% respondents are uncertain and 05% respondents are disagreed. The mean score 4.17 also supports the statement.

Table 4.24
Attitude of Higher Authorities Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	28	40	3	9	0	80	**3.8**
		Percentage	35%	50%	3.75%	11.25%	0%	100%	
	Rural	**Frequency**	38	28	6	8	0	80	**4.25**
		Percentage	47.5%	36%	7.50%	11.25%	0%	100%	
Total			**66**	**68**	**9**	**17**	**0**	**160**	

Table 4.24 highlights the opinion of school heads about the Attitude of Higher Authorities Affects Head's Performance. 85% urban respondents are agreed to the statement, 3.75% respondents are uncertain and 11.25% respondents are disagreed. The mean score 3.8 also supports the statement whereas 83.5% rural respondents are agreed, 7.5% respondents are uncertain and 11.25% respondents are disagreed. The mean score 4.25 also supports the statement.

Table 4.25
Vigilant Supervision of Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	42	32	1	3	2	**80**	**4.11**
		Percentage	52.5%	40%	1.25%	12%	2.50%	**100%**	
	Rural	Frequency	51	26	1	2	0	**80**	**4.12**
		Percentage	63.75%	32.50%	1.25%	2.50%	0%	**100%**	
Total			**93**	**58**	**2**	**5**	**2**	**160**	

Table 4.25 indicates the opinion of school heads about the Vigilant Supervision of Teachers Affects Head's Performance. 92.5% urban respondents are agreed to the statement, 1.25% respondents are uncertain and 14.50% respondents are disagreed. The mean score 4.11 also supports the statement whereas 96.25% rural respondents are agreed, 1.25 % respondents are uncertain and 2.5% respondents are disagreed. The mean score 4.12 also supports the statement.

Table 4.26
Weak Process of Evaluation Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	36	40	1	3	1	**80**	**4.12**
		Percentage	45%	50%	1.25%	3.75%	1.25%	**100%**	
	Rural	Frequency	36	37	3	3	1	**80**	**4.48**
		Percentage	45%	46.25%	3.75%	3.75%	1.25%	**100%**	
Total			**71**	**77**	**4**	**6**	**2**	**160**	

Table 4.26 shows the opinion of school heads about the Weak Process of Evaluation Affects Head's Performance. 95% urban respondents are agreed to the statement, 1.25% respondents are uncertain and 5% respondents are disagreed. The mean score 4.12 also supports the statement whereas 91.25% rural respondents are agreed, 3.75% respondents are uncertain and 5% respondents are disagreed. The mean score 4.48 also supports the statement.

Table 4.27
Coordination among Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	28	29	3	17	3	**80**	**3.15**
		Percentage	35%	36.25%	3.75%	21.25%	3.75%	**100%**	
	Rural	Frequency	24	29	6	20	1	**80**	**3.16**
		Percentage	30%	36.25%	7.50%	25%	1.25%	**100%**	
Total			**52**	**58**	**9**	**37**	**4**	**160**	

Table 4.27 highlights the opinion of school heads about the Coordination among Teachers Affects Head's Performance. 71.25% urban respondents are agreed to the statement, 3.75% respondents are uncertain and 25% respondents are disagreed. The mean score 3.15 also supports the statement whereas 66.25% rural respondents are

agreed, 7.50% respondents are uncertain and 26.25% respondents are disagreed. The mean score 3.16 also supports the statement.

Table 4.28
Teachers' Union Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	29	23	14	11	3	**80**	**4.12**
		Percentage	36.25%	28.75%	17.5%	13.75%	3.75%	**100%**	
	Rural	**Frequency**	31	24	11	12	2	**80**	**3.94**
		Percentage	38.75%	30%	13.75%	15%	2.50%	**100%**	
Total			**60**	**47**	**25**	**23**	**5**	**160**	

Table 4.28 indicates the opinion of school heads about the Teachers' Union Affects Head's Performance. 65% urban respondents are agreed to the statement, 17.5% respondents are uncertain and 17.50% respondents are disagreed. The mean score 4.12 also supports the statement whereas 68.75% rural respondents are agreed, 13.75 % respondents are uncertain and 17.5% respondents are disagreed. The mean score 3.94 also supports the statement.

Table 4.29
Environment of a School Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	49	25	4	2	0	**80**	**4.12**
		Percentage	61.25%	31.25%	5%	2.5%	0%	**100%**	
	Rural	**Frequency**	53	23	2	1	1	**80**	**3.56**
		Percentage	66.25%	28.75%	2.5%	1.25%	1.25%	**100%**	
Total			**102**	**48**	**6**	**3**	**1**	**160**	

Table 4.29 shows the opinion of school heads about the Environment of a School Affects Head's Performance. 92.5% urban respondents are agreed to the statement, 5% respondents are uncertain and 2.5% respondents are disagreed. The mean score 4.12 also supports the statement whereas 95% rural respondents are agreed, 2.5% respondents are uncertain and 2.5% respondents are disagreed. The mean score 3.56 also supports the statement.

Table 4.30
Poor Socio-Economic Status of Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	53	24	3	0	0	**80**	**4.23**
		Percentage	66.25%	30%	3.75%	0%	0%	**100%**	
	Rural	**Frequency**	45	33	2	0	0	**80**	**4.44**
		Percentage	57.5%	41.25%	2.50%	0%	0%	**100%**	
Total			**98**	**57**	**5**	**0**	**0**	**160**	

Table 4.30 highlights the opinion of school heads about the Poor Socio-Economic Status of Teachers Affects Head's Performance. 96.25% urban respondents are agreed to

the statement, and 3.75% respondents are uncertain. The mean score 4.23 also supports the statement whereas 98.75% rural respondents are agreed, and 2.5% respondents are uncertain. The mean score 4.44 also supports the statement.

Table 4.31
Inadequate Funding Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	27	37	16	0	0	**80**	**4.43**
		Percentage	33.75%	46.25%	20%	0%	0%	**100%**	
	Rural	Frequency	21	46	11	0	2	**80**	**4.13**
		Percentage	26.25%	57.5%	13.75%	0%	2.50%	**100%**	
Total			**48**	**83**	**27**	**0**	**2**	**160**	

Table 4.31 indicates the opinion of school heads about the Inadequate Funding Affects Head's Performance. 80% urban respondents are agreed to the statement, and 20% respondents are uncertain. The mean score 4.43 also supports the statement whereas 83.75% rural respondents are agreed, 2.5 % respondents are uncertain and 2.5% respondents are disagreed. The mean score 4.13 also supports the statement.

Table 4.32
Slow Process of Sanctioning of Bills Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	28	35	14	0	3	**80**	**3.9**
		Percentage	35%	43.75%	17.5%	0%	3.75%	**100%**	
	Rural	Frequency	24	42	10	0	4	**80**	**4.2**
		Percentage	30%	52.5%	12.5%	0%	5%	**100%**	
Total			**52**	**77**	**24**	**0**	**7**	**160**	

Table 4.32 shows the opinion of school heads about the Slow Process of Sanctioning of Bills Affects Head's Performance. 78.75% urban respondents are agreed to the statement, 17.5% respondents are uncertain and 3.75% respondents are disagreed. The mean score 3.9 also supports the statement whereas 82.5% rural respondents are agreed, 12.5 % respondents are uncertain and 5% respondents are disagreed. The mean score 4.2 also supports the statement.

Table 4.33
Limited Authority to Use School Budget Affects Head’s Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	34	32	3	3	8	**80**	**3.9**
		Percentage	42.5%	40%	3.75%	3.75%	10%	**100%**	
	Rural	Frequency	34	34	3	4	5	**80**	**4.26**
		Percentage	42.5%	42.5%	3.75%	5%	6.25%	**100%**	
Total			**68**	**66**	**6**	**7**	**13**	**160**	

Table 4.33 highlights the opinion of school heads about the Limited Authority to Use School Budget Affects Head’s Performance. 82.5% urban respondents are agreed to the

statement, 3.75% respondents are uncertain and 13.75% respondents are disagreed. The mean score 3.9 also supports the statement whereas 85% rural respondents are agreed, 3.75% respondents are uncertain and 11.25% respondents are disagreed. The mean score 4.26 also supports the statement.

Table 4.34
Poor Socio-Economic Status of Heads Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	27	37	6	0	10	**80**	**3.89**
		Percentage	33.75%	46.25%	7.50%	0%	12.5%	**100%**	
	Rural	Frequency	24	39	7	0	10	**80**	**4.13**
		Percentage	30%	48.75%	8.75%	0%	12.5%	**100%**	
Total			**51**	**76**	**13**	**0**	**20**	**160**	

Table 4.34 indicates the opinion of school heads about the Poor Socio-Economic Status of Heads Affects Head's Performance. 80% urban respondents are agreed to the statement, 7.5% respondents are uncertain and 12.50% respondents are disagreed. The mean score 3.98 also supports the statement whereas 78.75% rural respondents are agreed, 8.75 % respondents are uncertain and 12.5% respondents are disagreed. The mean score 4.13 also supports the statement.

Table 4.35
Poor Socio-Economic Status of Teachers Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	32	34	2	8	4	**80**	**4.25**
		Percentage	40%	42.5%	2.50%	10%	5%	**100%**	
	Rural	Frequency	25	32	7	14	2	**80**	**4.09**
		Percentage	31.25%	40%	8.75%	17.5%	2.25%	**100%**	
Total			**57**	**66**	**9**	**22**	**6**	**160**	

Table 4.35 shows the opinion of school heads about the Poor Socio-Economic Status of Teachers Affects Head's Performance. 82.5% urban respondents are agreed to the statement, 2.5% respondents are uncertain and 15% respondents are disagreed. The mean score 4.25 also supports the statement whereas 71.25% rural respondents are agreed, 8.75 % respondents are uncertain and 19.75% respondents are disagreed. The mean score 4.09 also supports the statement.

Table 4.36
Extra Workload Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	Urban	Frequency	41	25	9	0	5	**80**	**4.1**
		Percentage	51.25%	31.25%	11.25%	0%	6.25%	**100%**	
	Rural	Frequency	38	32	4	0	6	**80**	**3.67**
		Percentage	47.5%	40%	5%	0%	7.5%	**100%**	
Total			**79**	**57**	**13**	**0**	**11**	**160**	

Table 4.36 highlights the opinion of school heads about the Extra Workload Affects Head's Performance. 82.5% urban respondents are agreed to the statement, 11.25% respondents are uncertain and 6.25% respondents are disagreed. The mean score 4.1 also supports the statement whereas 87.50% rural respondents are agreed, 5% respondents are uncertain and 7.5% respondents are disagreed. The mean score 3.67 also supports the statement.

Table 4.37
Extra Workload Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	32	41	6	1	0	**80**	**4.24**
		Percentage	40%	51.25%	7.50%	1.25%	0%	**100%**	
	Rural	**Frequency**	36	39	3	1	1	**80**	**3.9**
		Percentage	45%	48.75%	3.75%	1.25%	1.25%	**100%**	
Total			**68**	**80**	**9**	**2**	**1**	**160**	

Table 4.37 indicates the opinion of school heads about the Extra Work Load Affects the Performance of Teachers. 91.25% urban respondents are agreed to the statement, 7.5% respondents are uncertain and 1.25% respondents are disagreed. The mean score 4.24 also supports the statement whereas 93.75% rural respondents are agreed, 3.75 % respondents are uncertain and 2.5% respondents are disagreed. The mean score 3.9 also supports the statement.

Table 4.38
Involvement of Teachers in Non-Teaching Tasks Affects Head's Performance

Category	Sector	Responses	SA	A	UNC	DA	SDA	Total	Mean
Head	**Urban**	**Frequency**	30	44	0	4	2	**80**	**4.3**
		Percentage	37.5%	55%	0%	5%	2.5%	**100%**	
	Rural	**Frequency**	31	43	0	5	1	**80**	**4.32**
		Percentage	38.75%	53.75%	0%	6.25%	1.25%	**100%**	
Total			**61**	**87**	**0**	**9**	**3**	**160**	

Table 4.38 shows the opinion of school heads about the Involvement of Teachers in Non Teaching Tasks Affects Head's Performance. 92.5% urban respondents are agreed to the statement, and 7.5% respondents are disagreed. The mean score 4.3 also supports the statement whereas 92.5% rural respondents are agreed, 7.5% respondents are disagreed. The mean score 4.32 also supports the statement.

4.2 Overall Analysis of Factors Affecting the Performance of Heads

Data collected through the questionnaire for the heads of government secondary schools were analyzed in terms of percentage and mean score. Factor wise analyzed data is given below.

Table 4.39
Administrative Vision Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	66	76	08	06	04	160	**3.89**
Percentage	41%	47.50%	05%	04%	2.5%	100%	

Table 4.39 indicates the opinion of school heads about the administrative vision factors. 88.50% of the respondents agreed, 05% respondents are uncertain and 6.50% respondents disagreed to the statement "administrative vision factors affect the performance of a school head". The mean score 3.89 also supported the statement.

Table 4.40
Physical Facilities Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	63	72	04	20	01	160	**3.93**
Percentage	39%	45%	2.5%	12.5%	01%	100%	

Table 4.40 shows the opinion of school heads about the physical facilities factor. 84% of the respondents agreed, 2.5% respondents are uncertain and 13.5% respondents disagreed to the statement "physical factors affect the performance of a school head". The mean score 3.93 also supported the statement.

Table 4.41
Learning Resources Factors

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	66	84	06	03	01	160	**4.21**
Percentage	41%	52%	04%	02%	01%	100%	

Table 4.41 highlights the opinion of school heads about the learning resources factors. 93% of the respondents agreed, 04% respondents are uncertain and 03% respondents disagreed to the statement "learning resources factors affect the performance of a school head". The mean score 4.21 also supported the statement.

Table 4.42
Teacher Related Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	75	74	04	06	01	160	**4.26**
Percentage	47%	46%	2%	04%	01%	100%	

Table 4.42 indicates the opinion of school heads about the teacher related factor. 93% of the respondents agreed, 02% respondents are uncertain and 05% respondents disagreed

to the statement “teacher related factors affect the performance of a school head”. The mean score 4.26 also supported the statement.

Table 4.43
Participation of Community Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	65	78	05	11	01	160	**4.08**
Percentage	41%	49%	03%	07%	01%	100%	

Table 4.43 shows the opinion of school heads about the participation of community factor. 90% of the respondents agreed, 03% respondents are uncertain and 08% respondents disagreed to the statement “participation of community factors affect the performance of a school head”. The mean score 4.08 also supported the statement.

Table 4.44
Political interference Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	71	62	12	14	01	160	**4.05**
Percentage	44%	39%	7%	09%	01%	100%	

Table 4.44 highlights the opinion of school heads about the political interference factor. 83% of the respondents agreed, 07% respondents are uncertain and 10% respondents disagreed to the statement “political interference factors affect the performance of a school head”. The mean score 4.05 also supported the statement.

Table 4.45
Supervision of Teachers Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	81	68	03	06	02	160	**3.77**
Percentage	51%	42%	02%	04%	1%	100%	

Table 4.27 indicates the opinion of school heads about the teachers’ supervision factor. 93% of the respondents agreed, 02% respondents are uncertain and 05% respondents disagreed to the statement “supervision of teachers factors affect the performance of a school head”. The mean score 3.77 also supported the statement.

Table 4.46
Relations among Teachers Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	78	53	10	16	03	160	**4.12**
Percentage	49%	33%	6%	10%	02%	100%	

Table 4.46 shows the opinion of school heads about the relations among teachers factor. 82% of the respondents agreed, 06% respondents are uncertain and 12% respondents disagreed to the statement “relations among teachers factors affect the performance of a school head”. The mean score 4.12 also supported the statement.

Table 4.47
Financial Resources Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	54	74	16	06	10	160	**4.03**
Percentage	34%	46%	10%	04%	06%	100%	

Table 4.47 highlights the opinion of school heads about the financial resources factor. 80% of the respondents agreed, 10% respondents are uncertain and 10% respondents disagreed to the statement “financial resources factors affect the performance of a school head”. The mean score 4.03 also supported the statement.

Table 4.48
Extra Workload Factor

Frequency/ Percentage	SA	A	UNC	DA	SDA	Total	Mean
Frequency	69	75	07	04	05	160	**4.23**
Percentage	43%	47%	04%	03%	03%	100%	

Table 4.48 indicates the opinion of school heads about the extra workload factor. 90% of the respondents agreed, 04% respondents are uncertain and 06% respondents disagreed to the statement “extra workload factors affect the performance of a school head”. The mean score 4.23 also supported the statement.

4.4 Overall Comparison of Factors Affecting Heads’ Performance Based on Rural and Urban Locality

S#	Table No.	FACTORS	Urban Mean	Rural Mean	Overall Weighted Mean
1.	4.1	Administrative Vision Factor	**3.95**	**3.83**	**3.89**
2.	4.2	Physical Facilities Factor	**3.77**	**4.10**	**3.93**
3.	4.3	Learning Resources Factor	**4.11**	**4.32**	**4.21**
4.	4.4	Teacher Related Factor	**4.13**	**4.39**	**4.26**
5.	4.5	Participation of Community Factor	**4.02**	**4.15**	**4.08**
6.	4.6	Political Interference Factor	**3.93**	**4.18**	**4.05**
7.	4.7	Supervision of Teachers Factor	**3.63**	**3.82**	**3.72**
8.	4.8	Relations Among Teachers Factor	**4.22**	**4.01**	**4.12**
9.	4.9	Financial Resources Factor	**4.00**	**4.07**	**4.03**
10.	4.10	Extra Workload Factor	**4.33**	**4.14**	**4.23**

1. **Sr. No. 4.1** indicates that the mean score 3.95 in urban area and 3.83 in rural area support the statement that administrative vision factor affects the performance of secondary school heads.
2. **Sr. No. 4.2** shows that the mean score 3.77 in urban area and 4.10 in rural area support the statement that physical factors affect the performance of secondary school heads.
3. **Sr. No. 4.3** narrates that the mean score 4.11 in urban area and 4.32 in rural area support the statement that learning resources factors affect the performance of secondary school heads.
4. **Sr. No. 4.4** highlights that the mean score 4.13 in urban area and 4.39 in rural area support the statement that teacher related factors affect the performance of secondary school heads.
5. **Sr. No. 4.5** indicates that the mean score 4.02 in urban area and 4.15 in rural area support the statement that participation of community factors affect the performance of secondary school heads.
6. **Sr. No. 4.6** shows that the mean score 3.93 in urban area and 4.18 in rural area support the statement that political interference factors affect the performance of secondary school heads.
7. **Sr. No. 4.7** narrates that the mean score 3.63 in urban area and 3.82 in rural area support the statement that supervision of teachers factors affect the performance of secondary school heads.
8. **Sr. No. 4.8** highlights that the mean score 4.22 in urban area and 4.01 in rural area support the statement that relations among teachers factors affect the performance of secondary school heads.
9. **Sr. No. 4.9** indicates that the mean score 4.0 in urban area and 4.07 in rural support area the statement that financial resources factors affect the performance of secondary school heads.
10. **Sr. No. 4.10** highlights that the mean score 4.33 in urban area and 4.14 in rural area support the statement that extra workload factors affect the performance of secondary school heads.

4.5 Factors Affecting the Performance of Heads in Ranking Order

S#	Table No.	Factors	Overall Weighted Mean
1	4.4	Teacher Related Factor	**4.26**
2	4.10	Extra Workload Factor	**4.23**
3	4.3	Learning Resources Factor	**4.21**
4	4.8	Relations Among Teachers Factor	**4.12**
5	4.5	Participation of Community Factor	**4.08**
6	4.6	Political Interference Factor	**4.05**
7	4.9	Financial Resources Factor	**4.03**
8	4.2	Physical Facilities Factor	**3.93**
9	4.7	Supervision of Teachers Factor	**3.72**
10	4.1	Administrative Vision Factor	**3.89**

1. **Sr. No. 4.4** shows that the factors at 1st rank affecting the performance of school heads are teacher related factors. The weighted mean score 4.26 supports the statement.
2. **Sr. No. 4.10** indicates that the factors at 2nd rank affecting the performance of school heads are the extra workload factors. The weighted mean score 4.23 supports the statement.
3. **Sr. No. 4.3** highlights that the factors at 3rd rank affecting the performance of school heads are learning resources factors. The weighted mean score 4.21 supports the statement.
4. **Sr. No. 4.8** narrates that the factors at 4th rank affecting the performance of school heads are relations among teachers factors. The weighted mean score 4.12 supports the statement.
5. **Sr. No. 4.5** shows that the factors at 5th rank affecting the performance of school heads are participation of community factors. The weighted mean score 4.08 supports the statement.
6. **Sr. No. 4.6** narrates that the factors at 6th rank affecting the performance of school heads are political interference factors. The weighted mean score 4.05 supports the statement.
7. **Sr. No. 4.9** highlights that the factors at 7th rank affecting the performance of school heads are financial resources factors. The weighted mean score 4.03 supports the statement.
8. **Sr. No. 4.2** indicates that the factors at 8th rank affecting the performance of school heads are physical facilities factors. The weighted mean score 3.93 supports the statement.
9. **Sr. No. 4.7** shows that the factors at 9th rank affecting the performance of school heads are supervision of teachers factors. The weighted mean score 3.72 supports the statement.
10. **Sr. No. 4.1** narrates that the factors at 10th rank affecting the performance of school heads are administrative vision factors. The weighted mean score 3.89 supports the statement.

5. SUMMARY, FINDINGS, CONCLUSIONS AND RECOMMENDATIONS

5.1 Summary

The intension of the study was to find “Factors Affecting the Performance of Heads of Government Secondary School in Punjab”. The main objective of the study was to identify the factors affecting the performance of heads of government secondary schools in Punjab. The study was conducted at Punjab level in secondary schools. The population of the proposed research was the male heads of government secondary schools of three districts i.e. Gujrat, Jhelum and Mandi Baha-ud-Din. The total population size of the said secondary schools for boys of three districts was 277. The sample of the study was consisted on 160 male heads of government secondary schools of three districts.

The design of this study was descriptive in nature so the Descriptive Method of research was adopted for the study and Random Sampling Technique was used. After the review of related literature one questionnaire was constructed to collect data. One questionnaire for the male heads of secondary schools on five point Likert scale was developed.

The collected data were organized, tabulated and analyzed through SPSS technique by applying percentage, mean and weighted mean. Findings were put forward, redrawn and conclusions were made and in the last, the relevant recommendations were also put forward.

5.2 Findings of Factors in Ranking Order

1. The factor affecting the performance of school heads at **1st rank** was ***teacher related factor***. The weighted mean score 4.26 also supports the statement that teacher related factors affect the performance of government secondary school heads.
2. The factor affecting the performance of school heads at **2nd rank** was the ***extra workload factor***. The weighted mean score 4.23 also supports the statement that extra workload factors affect the performance of government secondary school heads.
3. The factors affecting the performance of school heads at **3rd rank** was ***learning resources factor***. The weighted mean score 4.21 also supports the statement that learning resources factors affect the performance of government secondary school heads.
4. The factors affecting the performance of school heads at **4th rank** was ***relations among teachers factor***. The weighted mean score 4.12 also supports the statement that relations among teachers' factors affect the performance of government secondary school heads.
5. The factors affecting the performance of school heads at **5th rank** was ***participation of community factor***. The weighted mean score 4.08 also supports the statement that participation of community factors affect the performance of government secondary school heads.
6. The factors affecting the performance of school heads at **6th rank** was ***political interference factor***. The weighted mean score 4.05 also supports the statement that political interference factors affect the performance of government secondary school heads.
7. The factors affecting the performance of school heads at **7th rank** was ***financial resources factor***. The weighted mean score 4.03 also supports the statement that financial resources factors affect the performance of government secondary school heads.
8. The factors affecting the performance of school heads at **8th rank** was ***physical facilities factor***. The weighted mean score 3.93 also supports the statement that physical facilities factors affect the performance of government secondary school heads.
9. The factors affecting the performance of school heads at **9th rank** was ***teachers' supervision factor***. The weighted mean score 3.72 also supports the statement that teachers' supervision factors affect the performance of government secondary school heads.

10. The factors affecting the performance of school heads at **10th rank** was ***administrative factor***. The weighted mean score 3.89 also supports the statement that administrative factors affect the performance of government secondary school heads.

5.3 Objective Wise Conclusions

In the light of statistical analysis and findings of the study, the objective wise conclusions were drawn based on rural and urban locality as under:

The Objective No. 1 was to identify the factors affecting the performance of heads of government secondary schools.

It was concluded that administrative factor, physical facilities factor, learning resources factor, teacher related factor, community related factor, political interference factor, supervision of teachers factor, relation among teachers factor, financial factor and workload on heads factor were **identified** as the factors affecting the performance of heads.

The Objective No.2 was to categorize the factors affecting the performance of heads of government secondary schools.

It was concluded that there were ten main factors affecting the performance of government secondary school heads so the identified factors were then **categorized** into six categories. The six categories of factors were located as under:

1. **Category No.1** *Administration Related Factors*
2. **Category No.2** *Teacher Related Factors*
3. **Category No.3** *Student Related Factors*
4. **Category No.4** *Community Related Factors*
5. **Category No.5** *Political Related Factors*
6. **Category No.6** *Financial Related Factors*

The Objective No.3 was to compare the factors affecting the performance of heads of government secondary schools of urban and rural areas.

It was concluded that different factors affect the performance of head of government secondary schools. They vary from area to area. The comparison based on urban and rural locality was as under:

1. The study revealed that ***Factor of Administration*** including professional vision, professional training, and leadership style of a head in urban area as compared to the rural area highly affects the performance of government secondary school heads.
2. The study demonstrated that ***Factor of Physical facilities*** highly affects the performance of government secondary school heads in rural area as compared to the urban area. It also adversely affects the educational process as well in both of the locality.
3. The study exposed that ***Factor of learning resources*** including dual medium of instruction in rural area as compared to the urban area highly affects the performance of government secondary school heads.

4. The study revealed that ***Factor related to teachers*** including shortage of teachers, shortage of subject specialists, low professional motivation among teachers and absenteeism among teachers in rural area as compared to the urban area highly affects the performance of government secondary school heads.
5. The study demonstrated that ***Factor of participation*** of community including weak or less participation of parents in PTMs, and weak coordination of a school head with parents in rural area as compared to the urban area highly affects the performance of government secondary school heads.
6. The study revealed that ***Factor of political interference*** including local politics, negative interference of local politics, and negative attitude of higher authorities in rural area as compared to the urban area highly affects the performance of government secondary school heads.
7. The study stated that ***Factor of teachers' supervision*** including weak process of supervision and monitoring and evaluation in rural area as compared to the urban area highly affects the performance of government secondary school heads.
8. The study revealed that ***Factor of relations among teachers*** including weak relations among school staff, teachers' unions, poor working environment at a school, and weak coordination of a school head with his staff in urban area as compared to the rural area highly affects the performance of government secondary school heads.
9. The study demonstrated that ***Factor of financial resources*** including inadequate funding, slow process of sanctioning of bills, limited authority to use the school budget and poor socio-economic status of teachers and school heads as well in rural area as compared to the urban area highly affects the performance of government secondary school heads.
10. The study stated that ***Factor of extra workload*** on heads and teachers and involvement of school heads and teachers in urban area as compared to the rural area highly affects the performance of government secondary school heads.

The Objective No.4 was to prepare the list of factors affecting the performance of government in ranking order.

1. It was concluded that different factors affect the performance of heads of government secondary schools in different intensity. The factors in ranking order are as under:
2. The study revealed that the factor at 1st rank affecting the performance of school heads was found as teacher related factor.
3. The study discovered that the factor at 2nd rank affecting the performance of school heads was found as the extra workload factor.
4. The study exposed that the factor at 3rd rank affecting the performance of school heads was found as the learning resources factor.
5. The study highlighted that the factor at 4th rank affecting the performance of school heads was found as the relations among teachers factor.
6. The study stated that the factor at 5th rank affecting the performance of school heads was found as the participation of community factor.
7. The study revealed that the factor at 6th rank affecting the performance of school heads was found as the political interference factor.

8. The study highlighted that the factor at 7th rank affecting the performance of school heads was found as the financial resources factor.
9. The study exposed that the factor at 8th rank affecting the performance of school heads was found as the physical facilities factor.
10. The study revealed that the factor at 9th rank affecting the performance of school heads was found as the teachers' supervision factor.
11. The study highlighted that the factor at 10th rank affecting the performance of school heads was found as administrative factor.

5.4 Recommendations

Following recommendations were made.

1. The government should appointment the subject specialists in the schools where posts are vacant.
2. Extra workload on school heads and teachers should be eradicated.
3. The rural and urban schools should be equipped with relevant teaching learning resources.
4. School heads and teachers should create a working environment at their respective schools.
5. Participation of parents/community in the schools' affairs like PTMs should be encouraged.
6. All the schools should be given financial resources so that to fulfill their educational requirements.
7. Rural schools should be equipped with basic physical facilities.
8. Government should appoint the school heads possessing administrative and managerial vision.

REFERENCES

1. Adeyemi, T. (2010). Principals' Leadership styles and Teachers' job performance in senior secondary schools in Oudo State. *Nigeria: Journal of Education Administration and Policy Studies*, 2(6), 83-91
2. Ashraf, M. (2008). *Impact of management on teachers' performance at secondary level in Punjab.* Unpublished M.Phil. Thesis, Allama Iqbal Open University, Islamabad.
3. Ayub, M.R. (2001). *Impact of principals' attitude on student academic achievement at secondary level in District Islamabad*. Islamabad: Unpublished M.Phil. Thesis, Allama Iqbal Open University, Islamabad.
4. Din, N. (2008). *A study of motivation techniques used by heads of institutions of higher education and their impact on the performance of teachers.* Rawalpindi: Unpublished doctoral thesis, Arid Agriculture University, Faisalabad.
5. Griffin, G. (2004). Issues that KCSE results should raise. *The Educational insight. Issues*, 07, p24.
6. Hameed, A. (2013). *Implementing inclusive education in Pakistan*: A concept paper. Lahore: Department of Special Education.

7. Iqbal, M. (2013). *A comparative study of organizational structure, leadership style and physical facilities of public and private secondary schools in Punjab and their effect on school effectiveness.* Unpublished Doctor Dissertation, Lahore: Institute of Education and Research, University of the Punjab, Lahore.
8. James, C. and Connolly, U. (2000). *Effective Change in Schools.* Routledge Falmer, London.
9. Jones, A. (2004). *A Review of the Research Literature on Barriers to the Uptake of ICT by Teachers.* British Educational Communications and Technology Agency.
10. Khan, H. (2000). *Impact Study on the Contribution of Private Sector towards Human Resource Development in Islamabad*, Academy of Educational Planning and Management, Islamabad.
11. Malik, A.H. (2010). *Sociology of Gender Relations*. Teachers Students Welfare Forum, Multan.
12. Ndana, J.M. (2011). *The administrative Problems that Public Secondary School Heads Face: A Case of Kitui District*. Unpublished MBA Thesis, Kenyatta University.
13. Onyango, M.A. (2008). Factors influencing participation rates in Public primary school in Langwe Division, Homa-Bay District: Unpublished thesis, Nairobi University, Nairobi, Kenya.
14. Penkamjam, G. (2005). *Education and Development*. Delhi: Gyan Publishing House.
15. Retallick, J. and Farah, I. (2005). *Transforming schools in Pakistan: towards learning community*. New York: Oxford University Press.
16. Shah, D. (2003). *Decentralization in the education system of Pakistan: Policies and strategies*. Paper for Academy of Educational Planning Management Ministry of Education, Government of Pakistan, Islamabad.
17. Shami, P.A. (2007). Office management (module). Academy of Educational Planning and Management, Islamabad.
18. Smith, J.B. (2003). The effects of the social organization of schools on teacher satisfaction. *Sociology of Education*, 64, 190-208.
19. Smith, B.C. (1993). *Choices in the design of decentralization*. Common wealth Secretariat, London.
20. Sultana, N. and Bashir, M. (2005). *Educational management and supervision* (4th ed). Allama Iqbal Open University, Islamabad.
21. Wawira, M.G. (2012). *Heads' characteristics influencing instructional supervision in public primary schools in Kasarani District*. Kenya: Unpublished M.E.d project. University of Nairobi, Nairobi, Kenya.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 93-105

ON ESTIMATION OF REFLECTED WEIBULL DISTRIBUTION USING BAYESIAN UNDER INFORMATIVE PRIOR

Alina Syed, Wajiha Nasir and **Umme Kalsoom**
Government College Women University, Sialkot, Pakistan
Email: alinasyed97@gmail.com
wajiha.nasir@gcwus.edu.pk
ummekalsoom894@gmail.com

ABSTRACT

This paper gives the summary of Reflected Weibull Distribution by using Bayesian Analysis. The Reflected Weibull Distribution has been obtained by transformation in Weibull Distribution. We have estimated the scale parameter of Reflected Weibull Distribution under Bayesian Analysis by using Informative prior and derived its Posterior distribution. We studied the performance of posterior distribution by using different loss functions under exponential and gamma prior. Simulation study has been utilized to check the performance of Bayes estimator and their corresponding risks by using different sample size. We have also drawn the graphs of the posterior distribution for different values of scale parameter.

1. INTRODUCTION

The Reflected Weibull distribution was introduced by AC Cohen (1973). It has been shown to be widely employed as a model in reliability and in life testing applications. The modifications of this distribution consist of transformation of a continuous Weibull variate. It originates in the following linear transformation of a classical Weibull variate X:

$$Y = -X$$

The probability density function of Reflected Weibull distribution is as follows:

$$f(y; a, \theta) = \alpha\theta(-y)^{\theta-1}\, exp\{-\alpha(-y)^{\theta}\}, -\infty < y < 0, \theta > 0.$$

where α is scale parameter and θ is shape parameter.

It also has a version with three parameters. The probability density function of Reflected Weibull distribution with three parameters is as follows:

$$f(y; a, \theta) = \frac{\theta}{b}\left(\frac{a-y}{b}\right)^{\theta-1} exp\left\{-\left(\frac{a-y}{b}\right)^{\theta}\right\}, y < a, b\ \theta > 0.$$

When a is equal to zero, then Reflected Weibull distribution reduces to two parameters distribution. This is also known as a type-III extreme-value distribution.

In this study the Bayes Estimator and Posterior Risk of scale parameter are considered under Square Error Loss Function, Precautionary Loss Function, DeGroot Loss Function, Weighted Loss Function and Mean Square Loss Function using Exponential Prior and Gamma Prior.

Cohen (2012), has introduced Reflected Weibull Distribution. He proposed that when the Weibull distribution is reflected about the vertical axis $y = -x$, the Reflected Weibull Distribution is obtained. He has also included an illustrative example to demonstrate the use of the Reflected Weibull in fitting sample data and to provide comparisons with results obtained from fitting the Type VI, Type III, Type V and lognormal distributions. Connolly and Hirschey (1988), have studied the effect of patent statistics on firm market values. They found that these effects are robust to a variety of alternate specifications of their underlying model. And concluded that patent data provide economically meaningful information on the scope and relative effectiveness of a firm's R&D program. Basu and Ebrahimi (1991), have derived Bayes estimator of mean lifetime and reliability function in the exponential life testing model under number of prior distributions by using asymmetric loss function. They have also compared Bayesian estimates with corresponding estimates with squared error loss function. Feroze and Aslam (2012), have estimated the parameters of error function distribution. They have used four different loss functions to calculate Bayes estimators and their associated risks under six informative and non- informative priors using Squared Error Loss Function (SELF), Precautionary Loss Function (PLF), Quadratic Loss Function (QLF) and Entropy Loss Function (ELF). Guure and Ibrahim (2012), have determined the best method, by comparing the classical maximum likelihood against the Bayesian estimators. They have used an informative prior and proposed a data-dependent prior known as generalized non-informative prior. Then they considered Bayesian estimation under three loss functions and employed Lindley's approximation procedure to reduce the ratio of the integrals. They have obtained the MSE (mean squared error) and the absolute bias for the purpose of comparison. Feroze and Aslam (2012), have proposed posterior analysis of exponentiated gamma distribution for type II censored samples. They have derived the expressions for Bayes estimators and their associated risks under different priors and assumed the entropy and quadratic loss functions for estimation. They have also obtained posterior predictive distributions and constructed corresponding intervals. Feroze and Aslam (2012), have considered the Bayesian analysis of Inverse Rayleigh distribution under singly and doubly type II censored data. Then they have used symmetric and asymmetric loss functions to derive Bayes estimators and their corresponding risks under the assumption of non-informative priors. Also, they have constructed credible intervals under the assumption of non-informative priors. In order to compare the performance of the derived point and interval estimators they have used inverse transformation method of simulation for data generation and utilized it. Singh et al., (2013) have proposed Bayes estimators of the parameter and reliability function of inverted exponential distribution under the general entropy loss function for complete type-I and type-II censored samples. They have compared proposed estimators with the corresponding maximum-likelihood estimators for their simulated risks under different priors. Elbatal et al., (2014) have introduced a new model of Generalized Inverse Weibull distribution referred to as the Exponentiated generalized inverse distribution. They have derived the moment generating functions and rth moment. They have also obtained the expressions of density function, moment generating function and rth moment of the order statistics. Further, they have discussed the estimation of parameters by maximum likelihood and provided the information matrix. Almalki and Nadarajah (2014), have studied that the hazard function of Weibull distribution cannot exhibit non-monotonic shapes like the bathtub shape or the unimodal shape. They have given an extensive review of some discrete and continuous versions of the modifications of the Weibull distribution.

2. POSTERIOR DISTRIBUTION

The Bayesian Analysis of Reflected Weibull Distribution is discussed below. We have used two informative priors to estimate the scale parameter of posterior distribution.

2.1 Posterior Distribution using Exponential Prior

The Exponential prior of parameter α is defined as

$$P(\alpha|y) = \frac{1}{\lambda} e^{-\alpha/\lambda}, 0 < \alpha < \infty.$$

The posterior distribution for parameter α is

$$P(\alpha|y) = \frac{\left[\frac{1}{\lambda} + \sum_{i=1}^{n}(-y)^{\theta}\right]^{n+1} \alpha^{(n+1)-1} exp\left\{-\alpha\left(\frac{1}{\lambda} + \sum_{i=1}^{n}(-y)^{\theta}\right)\right\}}{\Gamma(n+1)}$$

and so $P(\alpha|y) \sim G(\alpha_1, \beta_1)$, where $\alpha_1 = n+1$ and $\beta_1 = \frac{1}{\lambda} + \sum_{i=1}^{n}(-y)^{\theta}$.

The Bayes estimator using SELF is

$$\alpha_{SELF} = \frac{n+1}{\frac{1}{\lambda} + \sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using SELF is

$$P(\alpha_{SELF}) = \frac{n+1}{\left\{\frac{1}{\lambda} + \sum_{i=1}^{n}(-y_i)^{\theta}\right\}^2}.$$

The Bayes estimator using PLF is

$$\alpha_{PLF} = \frac{\sqrt{(n+1)(n+2)}}{\frac{1}{\lambda} + \sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using PLF is

$$P(\alpha_{PLF}) = 2 * \left\{\frac{\sqrt{(n+1)(n+2)} - (n+1)}{\frac{1}{\lambda} + \sum_{i=1}^{n}(-y_i)^{\theta}}\right\}.$$

The Bayes estimator using DeLF is

$$\alpha_{DeLF} = \frac{(n+2)}{\frac{1}{\lambda}+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using DeLF is

$$P(\alpha_{DeLF}) = \frac{1}{n+2}.$$

The Bayes estimator using WLF is

$$\alpha_{WLF} = \frac{n}{\frac{1}{\lambda}+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using WLF is

$$P(\alpha_{WLF}) = \frac{1}{\frac{1}{\lambda}+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Bayes estimator using MELF is

$$\alpha_{MELF} = \frac{n-1}{\frac{1}{\lambda}+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using MELF is

$$P(\alpha_{MELF}) = \frac{1}{n}.$$

2.2 Posterior Distribution using Gamma Prior

The Gamma prior of parameter α is defined as

$$P(\alpha|y) = \frac{b^{a}}{\Gamma(a)}\alpha^{(a-1)}e^{-b\alpha}, 0 < \alpha < \infty.$$

The posterior distribution for parameter α is

$$P(\alpha|y) = \frac{\left[b + \sum_{i=1}^{n}(-y)^{\theta}\right]^{n+a}\alpha^{(n+a)-1}exp\left\{-\alpha\left(b + \sum_{i=1}^{n}(-y)^{\theta}\right)\right\}}{\Gamma(n+a)}.$$

and so $P(\alpha|y) \sim G(\alpha_2, \beta_2)$, where $\alpha_2 = n + a$ and $\beta_2 = b + \sum_{i=1}^{n}(-y)^{\theta}$

The Bayes estimator using SELF is

$$\alpha_{SELF} = \frac{n+a}{b+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using SELF is

$$P\left(\alpha_{SELF}\right) = \frac{n+a}{\left\{b+\sum_{i=1}^{n}(-y_i)^{\theta}\right\}^2}.$$

The Bayes estimator using PLF is

$$\alpha_{PLF} = \frac{\sqrt{(n+a)(n+a+1)}}{b+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using PLF is

$$P\left(\alpha_{PLF}\right) = 2*\left\{\frac{\sqrt{(n+a)(n+a+1)}-(n+a)}{b+\sum_{i=1}^{n}(-y_i)^{\theta}}\right\}.$$

The Bayes estimator using DeLF is

$$\alpha_{DeLF} = \frac{(n+a+1)}{b+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using DeLF is

$$P\left(\alpha_{DeLF}\right) = \frac{1}{n+a+1}.$$

The Bayes estimator using WLF is

$$\alpha_{WLF} = \frac{n+a-1}{b+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Posterior Risk using WLF is

$$P\left(\alpha_{WLF}\right) = \frac{1}{b+\sum_{i=1}^{n}(-y_i)^{\theta}}.$$

The Bayes estimator using MELF is

$$\alpha_{MELF} = \frac{n+a-2}{b+\sum_{i=1}^{n}\left(-y_i\right)^{\theta}}.$$

The Posterior Risk using MELF is

$$P\left(\alpha_{MELF}\right) = \frac{1}{n+a-1}.$$

3. ELICITATION OF HYPER-PARAMETERS

In this section, elicitation of hyper-parameters has been carried out by using prior predictive distributions, proposed by Aslam (2003). Prior predictive distribution is derived by

$$p(y) = \int_{-\infty}^{\infty} p(\alpha)\, f(x;\alpha)\, d\alpha.$$

3.1 Prior predictive using exponential prior

The prior predictive under exponential prior is

$$p(y) = \int_{0}^{\infty} \frac{1}{\theta} e^{-\alpha\theta} \alpha\theta\left(-y\right)^{\theta-1} e^{-\alpha(-y)^{\theta}}\, d\alpha .$$

3.2 Prior predictive using gamma prior

The prior predictive under gamma prior is

$$p(y) = \int_{0}^{\infty} \frac{b^{a}}{\Gamma(a)} \alpha^{a-1} e^{-b\alpha} \alpha\theta\left(-y\right)^{\theta-1} e^{-\alpha(-y)^{\theta}}\, d\alpha .$$

4. SIMULATION

The simulation can be helpful to approach the problems in the Bayesian analysis. We have conducted the simulation study to investigate the performance of Bayes estimator and Posterior risks using different priors under five different loss functions. Based on risk estimator the behavior of the loss functions is evaluated. Bayes estimator and the risks of these estimators have been computed and presented with respect to these loss functions. We have simulated the random sample of sizes n=50, 70, 100 and 200 of Reflected Weibull Distribution with parameter α=3, 5, 7 and parameter β=2, 5, 8 for 10,000 times.

Table 1
Simulation of Bayes Estimator and Posterior Risk using Exponential Prior

	N	SELF	PLF	DeLF	WLF	MELF
$\alpha = 3, \beta = 2$	30	3.1571 (0.3327)	3.2078 (0.1010)	3.2569 (0.0313)	3.0618 (0.1021)	2.9544 (0.0333)
	70	3.0687 (0.1346)	3.0957 (0.0431)	3.1123 (0.0139)	3.0225 (0.0432)	2.9794 (0.0143)
	100	3.0438 (0.0926)	3.0666 (0.0301)	3.0793 (0.0098)	3.0171 (0.0302)	2.9891 (0.0100)
	200	3.0249 (0.0458)	3.0322 (0.0150)	3.0383 (0.0050)	3.0092 (0.0150)	2.9932 (0.0050)
$\alpha = 3, \beta = 5$	30	3.1576 (0.3329)	3.2100 (0.1011)	3.2656 (0.0313)	3.0487 (0.1016)	2.9551 (0.0333)
	70	3.0643 (0.1342)	3.0926 (0.0431)	3.1169 (0.0139)	3.0231 (0.0432)	2.9797 (0.0142)
	100	3.0496 (0.0930)	3.0620 (0.0301)	3.0814 (0.0098)	3.0182 (0.0302)	2.9885 (0.0100)
	200	3.0213 (0.0456)	3.0325 (0.0150)	3.0411 (0.0050)	3.0124 (0.0151)	2.9929 (0.0050)
$\alpha = 3, \beta = 8$	30	3.1625 (0.3337)	3.2074 (0.1010)	3.2563 (0.0313)	3.0610 (0.1020)	2.9635 (0.0333)
	70	3.0659 (0.1343)	3.0941 (0.0431)	3.1079 (0.0139)	3.0256 (0.0432)	2.9816 (0.0142)
	100	3.0518 (0.0931)	3.0612 (0.0301)	3.0774 (0.0098)	3.0186 (0.0302)	2.9891 (0.0100)
	200	3.0223 (0.0457)	3.0304 (0.0150)	3.0361 (0.0050)	3.0088 (0.0150)	2.9949 (0.0050)
$\alpha = 5, \beta = 2$	30	5.2232 (0.9093)	5.2885 (0.1666)	5.3752 (0.0313)	5.0499 (0.1683)	4.8806 (0.0333)
	70	5.0844 (0.3691)	5.1255 (0.0714)	5.1638 (0.0139)	5.0129 (0.0716)	4.9536 (0.0142)
	100	5.0579 (0.2559)	5.0926 (0.0501)	5.1153 (0.0098)	5.0081 (0.0501)	4.9523 (0.0100)
	200	5.0358 (0.1268)	5.0485 (0.0250)	5.0578 (0.0050)	5.0101 (0.0251)	4.9854 (0.0050)
$\alpha = 5, \beta = 5$	30	5.2068 (0.9043)	5.3029 (0.1670)	5.4078 (0.0313)	5.0479 (0.1683)	4.8730 (0.0333)
	70	5.0868 (0.3697)	5.1350 (0.0716)	5.1668 (0.0139)	5.0218 (0.0717)	4.9445 (0.0142)
	100	5.0599 (0.2560)	5.0984 (0.0501)	5.1158 (0.0098)	5.0077 (0.0501)	4.9752 (0.0100)
	200	5.0281 (0.1264)	5.0450 (0.0250)	5.0548 (0.0050)	5.0047 (0.0250)	4.9784 (0.0050)

	N	SELF	PLF	DeLF	WLF	MELF
$\alpha = 5, \beta = 8$	30	5.2186 (0.9089)	5.2879 (0.1666)	5.3727 (0.0313)	5.0306 (0.1677)	4.8731 (0.0333)
	70	5.0886 (0.3700)	5.1256 (0.0714)	5.1582 (0.0139)	5.0258 (0.0718)	4.9403 (0.0142)
	100	5.0632 (0.2564)	5.0856 (0.0500)	5.1207 (0.0098)	5.0093 (0.0501)	4.9565 (0.0100)
	200	5.0300 (0.1265)	5.0494 (0.0250)	5.0539 (0.0050)	5.0057 (0.0250)	4.9831 (0.0050)
$\alpha = 7, \beta = 2$	30	7.2166 (1.7353)	7.3382 (0.2311)	7.4618 (0.0313)	6.9947 (0.2332)	6.7795 (0.0333)
	70	7.1064 (0.7215)	7.1410 (0.0995)	7.2142 (0.0139)	6.9925 (0.0999)	6.9061 (0.0143)
	100	7.0687 (0.4997)	7.1010 (0.0698)	7.1335 (0.0098)	6.9969 (0.0700)	6.9310 (0.0100)
	200	7.0309 (0.2471)	7.0418 (0.0349)	7.0650 (0.0050)	7.0003 (0.0350)	6.9723 (0.0050)
$\alpha = 7, \beta = 5$	30	7.2277 (1.7403)	7.3491 (0.1666)	7.4762 (0.0313)	6.9778 (0.2326)	6.7393 (0.0333)
	70	7.1055 (0.7215)	7.1527 (0.0714)	7.2024 (0.0139)	7.0013 (0.1000)	6.9020 (0.0143)
	100	7.0638 (0.4990)	7.1150 (0.0500)	7.1304 (0.0098)	6.9981 (0.0700)	6.9369 (0.0100)
	200	7.0331 (0.2473)	7.0452 (0.0250)	7.0740 (0.0050)	6.9907 (0.0350)	6.9702 (0.0050)
$\alpha = 7, \beta = 8$	30	7.2420 (1.7498)	7.3329 (0.2310)	7.4325 (0.0313)	6.9833 (0.2328)	6.7468 (0.0333)
	70	7.1058 (0.7212)	7.1591 (0.0998)	7.2136 (0.0139)	7.0006 (0.1000)	6.9067 (0.0143)
	100	7.0628 (0.4988)	7.1015 (0.0698)	7.1566 (0.0098)	6.9987 (0.0700)	6.9306 (0.0100)
	200	7.0339 (0.2474)	7.0453 (0.0349)	7.0719 (0.0050)	7.0006 (0.0350)	6.9693 (0.0050)

Table 2
Simulation of Bayes Estimator and Posterior Risk using Gamma Prior

	N	SELF	PLF	DeLF	WLF	MELF
$\alpha = 3, \beta = 2$	30	3.1976 (0.2831)	3.2420 (0.0859)	3.2919 (0.0263)	3.1095 (0.0864)	3.0226 (0.0278)
	70	3.0917 (0.1257)	3.1144 (0.0401)	3.1441 (0.0128)	3.0576 (0.0402)	3.0182 (0.0132)
	100	3.0756 (0.0892)	3.0807 (0.0286)	3.1002 (0.0093)	3.0389 (0.0287)	3.0118 (0.0094)
	200	3.0365 (0.0448)	3.0426 (0.0146)	3.0486 (0.0048)	3.0238 (0.0147)	3.0064 (0.0049)
$\alpha = 3, \beta = 5$	30	3.2114 (0.2857)	3.2423 (0.0859)	3.2765 (0.0263)	3.1110 (0.0864)	3.0248 (0.0278)
	70	3.0947 (0.1259)	3.1208 (0.0401)	3.1339 (0.0128)	3.0493 (0.0401)	3.0213 (0.0132)
	100	3.0724 (0.0890)	3.0843 (0.0286)	3.0995 (0.0093)	3.0422 (0.0287)	3.0084 (0.0094)
	200	3.0386 (0.0448)	3.0418 (0.0146)	3.0505 (0.0048)	3.0201 (0.0147)	3.0035 (0.0049)
$\alpha = 3, \beta = 8$	30	3.2024 (0.2841)	3.2389 (0.0858)	3.2938 (0.0263)	3.1076 (0.0863)	3.0236 (0.0278)
	70	3.0909 (0.1256)	3.1083 (0.0400)	3.1377 (0.0128)	3.0593 (0.0403)	3.0163 (0.0132)
	100	3.0681 (0.0888)	3.0844 (0.0286)	3.1020 (0.0093)	3.0403 (0.0287)	3.0136 (0.0094)
	200	3.0368 (0.0448)	3.0448 (0.0147)	3.0476 (0.0048)	3.0203 (0.0147)	3.0073 (0.0049)
$\alpha = 5, \beta = 2$	30	4.8123 (0.6385)	4.8706 (0.1290)	4.9371 (0.0263)	4.6793 (0.1300)	4.5449 (0.0278)
	70	4.9225 (0.2183)	4.9391 (0.0635)	4.9792 (0.0128)	4.8500 (0.0638)	4.7947 (0.0132)
	100	4.9481 (0.2308)	4.9615 (0.0460)	4.9894 (0.0093)	4.8970 (0.0462)	4.8455 (0.0094)
	200	4.9691 (0.1198)	4.9831 (0.0240)	4.9921 (0.0048)	4.9438 (0.0240)	4.9242 (0.0049)
$\alpha = 5, \beta = 5$	30	4.8105 (0.6379)	4.8802 (0.1293)	4.9296 (0.0263)	4.6861 (0.1302)	4.5580 (0.0278)
	70	4.9176 (0.3177)	4.9431 (0.0636)	4.9872 (0.0128)	4.8582 (0.0639)	4.7856 (0.0132)
	100	4.9349 (0.2296)	4.9626 (0.0461)	4.9791 (0.0093)	4.8909 (0.0461)	4.8451 (0.0094)
	200	4.9724 (0.1200)	4.9747 (0.0240)	4.9986 (0.0048)	4.9461 (0.0240)	4.9251 (0.0049)

	N	SELF	PLF	DeLF	WLF	MELF
$\alpha = 5, \beta = 8$	30	4.8232 (0.6412)	4.8676 (0.1290)	4.9263 (0.0263)	4.6814 (0.1300)	4.5511 (0.0278)
	70	4.9094 (0.3165)	4.9470 (0.0636)	4.9774 (0.0128)	4.8538 (0.0639)	4.7908 (0.0132)
	100	4.9348 (0.2295)	4.9528 (0.0460)	4.9833 (0.0093)	4.8957 (0.0461)	4.8477 (0.0094)
	200	4.9686 (0.1198)	4.9817 (0.0240)	4.9939 (0.0048)	4.9437 (0.0240)	4.2938 (0.0049)
$\alpha = 7, \beta = 2$	30	6.1367 (1.0343)	6.2154 (0.1647)	6.2957 (0.0263)	5.9818 (0.1662)	5.8030 (0.0278)
	70	6.5674 (0.5659)	6.6077 (0.0850)	6.6555 (0.0128)	6.4841 (0.0853)	6.3930 (0.0132)
	100	6.6881 (0.4214)	6.7235 (0.0624)	6.7443 (0.0093)	6.6264 (0.0625)	6.5653 (0.0094)
	200	6.8290 (0.2263)	6.8517 (0.0330)	6.8706 (0.0048)	6.8098 (0.0331)	6.7687 (0.0049)
$\alpha = 7, \beta = 5$	30	6.1199 (1.0288)	6.2153 (0.1647)	6.3025 (0.0263)	5.9747 (0.1660)	5.8069 (0.0278)
	70	6.5601 (0.5647)	6.6092 (0.0850)	6.6514 (0.0128)	6.4754 (0.0852)	6.3985 (0.0132)
	100	6.6828 (0.4207)	6.7350 (0.0625)	6.7469 (0.0093)	6.6190 (0.0624)	6.5701 (0.0094)
	200	6.8423 (0.2272)	6.8558 (0.0330)	6.8661 (0.0048)	6.8045 (0.0330)	6.7647 (0.0049)
$\alpha = 7, \beta = 8$	30	6.1338 (1.0337)	6.2127 (0.1646)	6.2990 (0.0263)	5.9814 (0.1662)	5.8006 (0.0278)
	70	6.5757 (0.5674)	6.6166 (0.0851)	6.6585 (0.0128)	6.4843 (0.0853)	6.3938 (0.0132)
	100	6.6881 (0.4214)	6.7131 (0.0623)	6.7496 (0.0093)	6.6194 (0.0624)	6.5673 (0.0094)
	200	6.8438 (0.2273)	6.8527 (0.0330)	6.8691 (0.0048)	6.8067 (0.0330)	6.7632 (0.0049)

5. GRAPHS

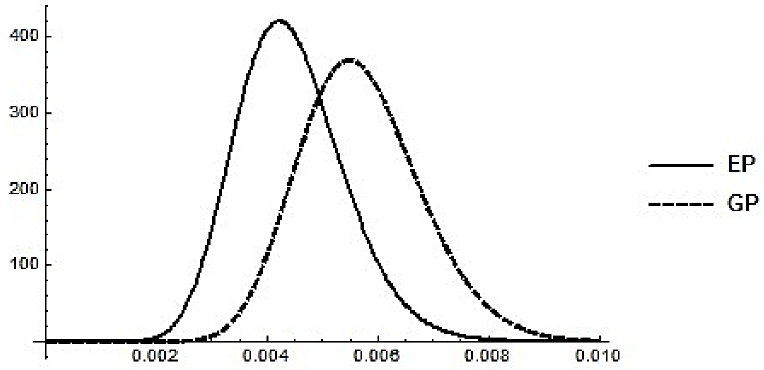


Figure 1: Posterior Distribution When α = 3

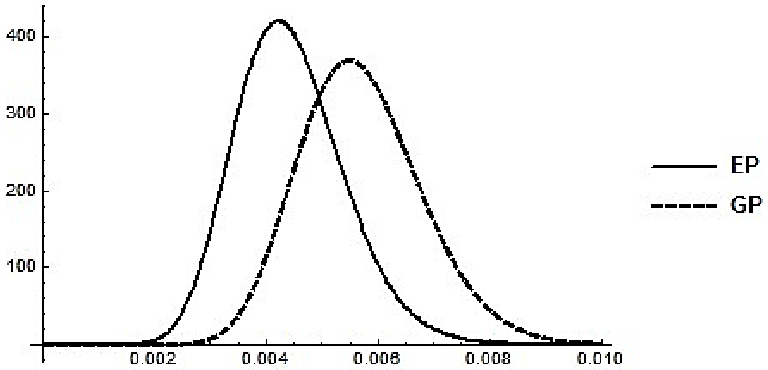


Figure 2: Posterior Distribution When α = 5

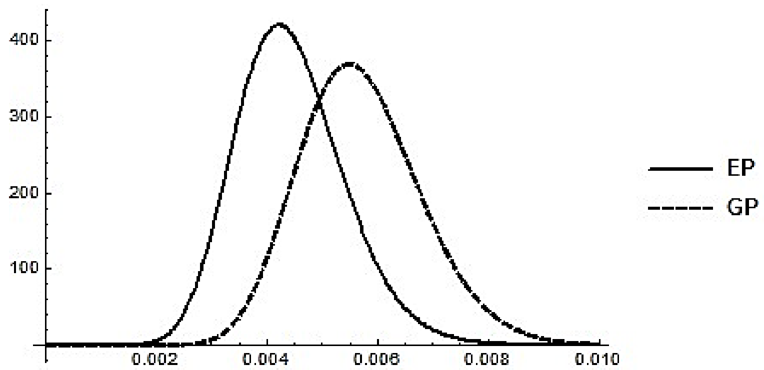


Figure 3: Posterior Distribution When α = 7

Figure 1, 2 and 3 shows the result of posterior distribution. The above graphs show that posterior distribution under informative priors (exponential and gamma prior) are symmetrical. The behavior of posterior distribution under exponential and gamma prior is almost same.

6. CONCLUSION

We have used this simulation technique to check the behavior of scale parameter of Reflected Weibull Distribution. The above tables show the result of posterior distribution under exponential and gamma prior by using different loss functions which are Square error loss function, Precautionary loss function, DeGroot loss function, Mean square loss function and weighted loss function. After the simulation and comparing these loss functions the results shows that as we increase sample size Bayes estimator approaches to its true value of parameter and posterior risk tends to decrease. From Table 1 we can see that DeLF is performing better because its posterior risk is minimum as compared to other loss functions. Similarly, in Table 2 DeLF is performing better. Whereas, the overall results show that the posterior distribution under gamma prior with DeLF is performing better because it is providing minimum posterior risk as compared to other loss functions and priors. The graphs show that posterior distribution under informative priors (exponential and gamma prior) is symmetrical.

REFERENCES

1. Almalki, S.J. and Nadarajah, S. (2014). Modifications of the weibull distribution: A review. *Reliability Engineering & System Safety*, 124, 32-55.
2. Aslam, M. (2003). An application of prior predictive distribution to elicit the prior density. *Journal of Statistical Theory and Applications*, 2(1), 70-83.
3. Basu, A. and Ebrahimi, N. (1991). Bayesian approach to life testing and reliability estimation using asymmetric loss function. *Journal of Statistical Planning and Inference*, 29(1-2), 21-31.
4. Berger, J.O. (1985). *Statistical decision theory and Bayesian analysis*. Springer.
5. Bernardo, J.M. (2011). Bayesian statistics. In *International Encyclopedia of Statistical Science*, pages 107-133. Springer.
6. Bolstad, W.M. (2007). *Introduction to Bayesian statistics*. John Wiley & Sons.
7. Box, G.E. and Tiao, G.C. (2011). *Bayesian inference in statistical analysis*, Vol. 40, John Wiley & Sons.
8. Cohen, A.C. (1973). The reflected Weibull distribution. *Technometrics*, 15(4), 867-873.
9. Connolly, R.A. and Hirschey, M. (1988). Market value and patents: A bayesian approach. *Economics Letters*, 27(1), 83-87.
10. DeGroot, M. H. (2005). *Optimal statistical decisions*, volume 82. John Wiley & Sons.
11. Elbatal, I. and Muhammed, H.Z. (2014). Exponentiated generalized inverse weibull distribution. *Applied Mathematical Sciences*, 8(81), 3997-4012.
12. Feroze, N. and Aslam, M. (2012). A note on Bayesian analysis of error function distribution under different loss functions. *International Journal of Probability and Statistics*, 1(5), 153-159.

13. Feroze, N. and Aslam, M. (2012). Bayesian analysis of exponentiated gamma distribution under type II censored samples. *Sci. J. Pure. Appl. Sci.*, 1(1), 30-39.
14. Feroze, N. and Aslam, M. (2012). On posterior analysis of inverse Rayleigh distribution under singly and doubly type II censored data. *International Journal of Probability and Statistics*, 1(5), 145-152.
15. Gelman, A. (2004). *Bayesian data analysis*. Chapman Hall/CRC, Boca Raton, Fla.
16. Guure, C.B. and Ibrahim, N.A. (2012). Bayesian analysis of the survival function and failure rate of weibull distribution with censored data. *Mathematical Problems in Engineering*, 2012.
17. Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007), 453-461.
18. Mood, A.M. and Graybill, F.A. (1950). *Introduction to the Theory of Statistics*, New York: McGraw-H ill Book Company.
19. Norstrom, J.G. (1996). The use of precautionary loss functions in risk analysis. *IEEE Transactions on Reliability*, 45(3), 400-403.
20. Samaniego, F.J. (2010). *A comparison of the Bayesian and frequentist approaches to estimation*. Springer.
21. Singh, S.K., Singh, U. and Kumar, D. (2013). Bayes estimators of the reliability function and parameter of inverted exponential distribution using informative and non-informative priors. *Journal of Statistical Computation and Simulation*, 83(12), 2258-2269.
22. Tummala, V. and Sathe, P.T. (1978). Minimum expected loss estimators of reliability and parameters of certain lifetime distributions. *IEEE Transactions on Reliability*, 27(4), 283-285.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 107-117

POVERTY LEADS TO CRIME: A CASE STUDY IN GC WOMEN UNIVERSITY, SIALKOT

Minahil Malik, Mahwish Rabia, Minahil Imran, Ramisha Irshad and **Kiran Bint-e-Khalid**
Government College Women University, Sialkot, Pakistan
Email: mahwish.rabia@gcwus.edu.pk

ABSTRACT

This study examines how poverty leads to crime. It demonstrates that crimes significantly contribute in the acknowledgment of crimes, its reasons, causes and consequences. Poverty talks about a person's mind and reveals the fact that what sort of crimes is expected from a certain person and what measures should be taken for its eradication. The data is collected by using questionnaire from university students. The results revealed that 91.6% people believed that poverty is the main cause of crime while rest of the people disagreed. 5.8% people claim that the main cause of poverty is lack of education and 11.2% claim that this happen because of less job opportunities. Moreover, people mostly commit robbery in poverty. The relationship between different variable is calculated by using chi-square test. The value of $\chi^2 = 17.664$ illustrates that there is significant association between main cause of crime and most likely group to commit crime. Therefore, it is concluded that poverty correlates with crime and is more often seen in poor people who are more attracted towards evil ways in order to fulfill their desired and basic needs.

KEYWORDS

Poverty, Crime, Hunger, Suicide, Government, Pakistan, Human, Opportunities, Success, Association.

1. INTRODUCTION

Poverty is a characteristic of the economic situation of the individuals or social group in which they cannot satisfy a certain range of the minimum requirements needed for life saving ability. Poverty is a relative concept and depends on the overall standard of living in this society. Poverty shares an intimate connection with crimes although it is true that all criminals are not poor and all the poor are not criminals. It is the likelihood of the people living in poverty to commit crime that brings up the correlation. The causes and effect of poverty are determining factors that make poverty often reason of crime. The facts that more white collar often go uninvestigated is probably one reason why more poor are associated with crime.

Poor people are intrinsically criminals which also suggest that they are inherently poor as well. If people aren't enjoying the best of life and were desperately needful ultimately, they would be driven to committing crime. Sometime despite various efforts made poverty is passed on from one generation to another and this poverty trap only increases the more ways towards crimes. Poverty does not limit economic possibilities but also cripples many opportunities and may have adverse consequences. The poor suffer not only because of lack of economic opportunities but also lack of studies and lack of health care. They are prey to more evidence. The incidence of crime is known to rise especially in times of unemployment. According to experts, not getting decent wages for their work, young people gradually lose the desire to marry, have children – they are full of fear for tomorrow, not sure that tomorrow they will not be fired because of another financial crisis. In an addition, low wages and lack of jobs, experts say, are pushing young people to commit crimes in order to get rich quickly.

It is therefore amongst the adolescents, aged 18 to 25, that there is an increase in years. As a fact, crimes begin from the childhood. As a rule, if a child was born in a poor family, there are more chances that he will start committing crimes. All over the world there is propaganda of a luxurious way of life, of successful people, of expensive mansions, cars, of beautiful journeys etc. When children and adolescents see that and understand that parents will not provide them with all these things, they start looking for the opportunity to get money fast to buy at least something. And consequently, they decide to commit a crime. The crimes committed by adolescents are considered the cruelest. The majority of cases of all crimes show that a driving force of the crimes is the money. The values of the society are built the way that those people with the problems in their minds decide to bring closer all these luxurious things. Thus, who might have never considered crime might take to it as a mean to survive only the eradications will bring a radical change. Eradication cannot be done overnight efforts can be made this issue needs to be taken seriously basic needs like food, shelter and sanitation are basic rights and should be provided by the government.

One of the main aspects that can prevent poverty is education. If children are educated properly, then they will perceive the reality as it is and they will try to succeed in legal ways to study, to get a job. Sometimes, crimes are the symptom of the culture and education can cure the cultural poverty and financial poverty will be cured itself, as a consequence. The state should make jobs available for the poor. If a dignity of men is restored half the battle is one counseling centers should be opened and proper ways should be adopted.

2. LITERATURE REVIEW

Kelly (2000) studied the relationship between poverty and crime using data from urban countries. He found that behavior of property and violent crime are quite different. He used regression analysis to investigate the relationship between poverty and crime and found a significant relationship between poverty and crime.

Melham (2005) has studied the crime induced poverty traps. Huang and Wang (2004) studied the theoretical relationship between crime and poverty. They studied the accumulation of human capital, crime and unemployment for searching of equilibrium model. Mauro and Carmike (2007) have studied poverty trap of crime and unemployment. Present an overlapping generation growth model with an imperfect labor market where the links among crime, growth and unemployment are jointly considered, both in an endogenous and exogenous set-up. They tested the major implications of our theory and verify the two model specifications through the Italian regional data, using the Pooled Mean Group estimator proposed by Persian, Shin and Smith (1999). The empirical results are in favor of the exogenous version of the model and suggest that crime and unemployment have long-run income level effects.

Carlen’s (1988) study of imprisoned women revealed that although poverty was common among imprisoned women. Hannon (2002) studied criminal opportunity theory and the relationship between poverty and property crime. The present study examined the relationship between economic deprivation and rates of burglary and motor vehicle theft for census tracts in two large American cities (Austin and Seattle). Regression analysis support the opportunity saturation hypothesis derived from criminal opportunity theory. This hypothesis suggests that the relationship between levels of deprivation and property crime is curvilinear where the positive effect of deprivation on property crime is stronger at low levels of neighborhood poverty than it is at high levels.

Ludwig (2000) studied the relationship by using randomized housing-mobility experiment. He studied the effects of relocating families from high to low poverty neighborhoods on juvenile crime by using randomized housing-mobility experiment. Their findings seem to suggest that providing families with the opportunity to move to lower-poverty neighborhoods reduces violent criminal behavior by teens.

Pare (2014) has examined the relationship between income inequality, poverty and different types of crime. His results showed that inequality is unrelated to homicide rates when poverty is controlled. In our multilevel analysis of the International Crime Victimization Survey we find that inequality is unrelated to assault, robbery, burglary and theft when poverty is controlled.

Wilkinson and Pickett (2009) link economic inequality to a wide range of social evils, including lower social trust, impaired mental and physical health, excessive consumption, drug addiction, obesity and failing education systems. For them, it is the context of inequality and not material poverty that generates these problems. Pride more (2008) then analyzed the relationship between inequality and homicide rates controlling for the infant mortality rate, an indirect measure of poverty. Based on a sample of 46 countries, he found that homicide rates are related to infant mortality but that the relationship between homicide rates and the Gini Index of income inequality is not statistically significant.

Pride more (2011) re-examined the inequality vs. poverty question using three different data sets from prior published studies (Fajnzylber et al. 1998, Savolainen 2000). He found a positive effect of infant mortality and no effect of the Gini Index based on Fajnzylber et al.’s data and on the Savolainen’s female (victims) homicide data. He also found, however, that the infant mortality and the Gini Index both have positive affects in Savolainen’s male (victims) homicide data.

Patterson (1991) examined the relationship between crime rates and aggregate economic conditions for 57 small social areas. The findings indicate that absolute poverty is more strongly associated with neighborhood crime rates, although the relationship is conditional on the type of crime considered.

3. RESEARCH METHODOLOGY

In this study, the population size of 2604 is considered taking BS students, of Government College Women University Sialkot, in which the number of female students of natural sciences are 1361 and social sciences are 1243. A sample of 347 has been drawn using Yamane (1967), where “N” is the total population size and “e” is the margin of error. The procedure is given below:

$$n = \frac{N}{1+Ne^2} \tag{1}$$

$$n = \frac{2604}{1+2604(0.05)^2} = 347.$$

3.1 Procedure for Data Collection

The Poverty Leads to Crime Questionnaire was distributed to the samples of 5th and 7th semester of some departments of natural sciences and social sciences in their classes. The samples of both semesters were required to answer the questionnaire in the class and collect them immediately upon completion. This is done in order to ensure a 100% return rate of the questionnaire from the samples of both semesters.

3.2 Analysis of Data

The following statistical techniques are used to analyze the data:

1. Frequency Distributions
2. Chi-Square Test

Significance of the Study

The significance of the study is:

1. To examine different reason of crimes in Pakistan.
2. To examine the thoughts of individuals about poverty.
3. To examine the association between the children who commit crime in poverty and causes of crime.
4. To study the association between the causes of crimes and group of committing crimes.
5. To study the different aspects of poverty.
6. To suggest different ways to overcome the crime.

4. RESULTS AND DISCUSSION

4.1 Frequency Distributions

Table 1
Univariate Data Analysis

Variable	Categories	Frequency	Percentage
Gender	Female	347	100.0
Area	Rural	229	66.0
	Urban	118	34.0
Age	16-25 years	322	92.8
	26-49 years	25	7.2
Father's Monthly Income	15,000-25,000	69	19.9
	25,000-45,000	132	38.0
	45,000-50,000	99	28.5
	50,000 to onwards	47	13.5
Current CGPA	2.00 to 2.50	21	6.1
	2.50 to 3.00	137	39.5
	3.00 to 3.50	159	45.8
	3.50 to 4.00	30	8.6

Table 1 shows the frequency distribution of different variables. The overall female in the whole university is 100.0% and 66.0% of them are from rural area and the percentage of females living in urban area is 34.0%. The percentage of age ranging from 16-25 years is 92.8% and from 26-49 years is 7.2%. It is clear that 20% respondents have 15000-25000 income, 38% from 25000-45000, 28.5% from 45,000-50,000 and 14% above 50,000. In addition, 6.1% respondents obtained 2.00-2.50 CGPA, 39.5% obtained 2.50-3.00 CGPA, 45.8% obtained 3.00-3.50 CGPA and 8.6% obtained 3.50-4.00 CGPA.

Table 2
Poverty is the Main Cause of Crime

Variable	Categories	Frequency	Percentage
Would you say poverty is the main cause of crime	Yes	318	91.6
	No	29	8.4
	Total	347	100

Table 2 displays that out of 347 respondents 91.6% believed that poverty is the main cause of crime while the rest of disagreed. The results proved that poverty is the main cause of crime since it instills a feeling of desperation in people, which forces them to commit crimes and to gain some money.

Table 3
Main Reasons of Poverty

Variable	Categories	Frequency	Percentage
What is the main reason of poverty in Pakistan	Lack of Education	20	5.8
	Overpopulation	35	10.1
	Epidemic Disease (AIDS, Malaria etc.)	42	12.1
	Lack of Job Opportunities	39	11.2
	Effects of Poverty on Society	46	13.3
	Child Labor	47	13.5
	Hunger	51	14.7
	Illness	36	10.4
	Lack of Housing	31	8.9
	Total	347	100

Table 3 shows that 5.8% respondents believed that the main cause of poverty is lack of education, 10.1% think that it is followed up by overpopulation, 12.1% thinks it is because of epidemic diseases. Moreover, other factors included lack of job opportunities, effects of poverty on society, child labor, hunger, illness and finally lack of housing are also source of poverty.

Table 4
Reason Behind the Poverty

Variable	Categories	Frequency	Percentage
Why would you say poverty is the main cause of crime	Depression and Poor Social Status	72	20.7
	Boredom (Feeling Bore)	104	30.0
	Hunger	136	39.2
	Unfair System to get Jobs	35	10.1
	Total	347	100

Table 4 demonstrates that 20.7% respondent's opinion is that depression and poor social status, while 30.0% believed that when someone feels boring and do not have any work to do, it leads to commit crime. On the other hand, 39.2% and 10.1% believed hunger and unfair system to get jobs respectively are the main cause of poverty which leads to crime. They believed so, because poor don't get education and not groomed enough to understand the drawbacks of crimes and ultimately they are the major victims of crime. Moreover, poor people feel they are owed money by the world and feel that crime is the way to get what they deserve.

Table 5
Most likely Group to Commit Crime

Variable	Categories	Frequency	Percentage
What do you think is the most likely group to commit crime	Upper class	25	7.2
	Middle class	78	22.5
	Lower class	98	28.2
	Working class	48	13.8
	Unemployed Men	30	8.6
	Unemployed women	44	12.7
	All	24	6.9
	Total	347	100

Table 5 represents the opinion of respondents about the most likely group to commit crime and 7.2% believed that upper class commits more crime, 22.5% believed that middle class is the most likely group to commit crime while 28.2% believed that lower class are often more victims of crimes and 13.8% believed that the working class people commit crimes. Furthermore, 8.6% and 12.7% believed that unemployed men and unemployed women commit more crimes respectively and finally 6.9% believed that all these groups commit crime. The results revealed that it is believed that lower class is mainly the most likely group to commit crime because people want to attain the luxurious life and they use illegal ways to do so. That’s why majority of the results showed that lower class is the mostly likely group to commit crimes.

Table 6
Kinds of Crime in Poverty

Variable	Categories	Frequency	Percentage
What kind of crime people can commit in poverty	Drugs and Alcohol	102	29.4
	Theft or Robbery	153	44.1
	Crime against property	92	26.5
	Total	347	100

Table 6 displays what kinds of crime people can commit in poverty. It is examined that 29.4% believed that people become addicted to drugs and alcohol while 44.1% believed that people commit crimes like theft and robbery in poverty and finally 26.5% believed that to fight for property is the major reason people can commit in poverty.

Table 7
Effects of Poverty in our Living Life Style

Variable	Categories	Frequency	Percentage
What do you think how the poverty effects our living life style	Bad reputation	153	44.1
	Fear of public	133	38.3
	Stereotype-Judged poorly from the neighborhood.	61	17.6
	Total	347	100

Table 7 appears 44.1% believed that bad reputation is caused due to poverty which affects the living life style. 38.3% believed that due to poverty fear of public is caused which affects the living life styles and 17.6% believed that the stereotypical judgment from the neighborhood affects the life style because of poverty. Though, all these factors effect a lot but the high percentage believed that poverty causes bad reputation which automatically disconnects the poor people from the upper class society and ultimately they are critically judged by the society.

Table 8
Children Commit Crime in Poverty

Variable	Categories	Frequency	Percentage
Do you feel that children who live in poverty are more likely to commit crime	Yes	265	76.4
	No	82	23.6
	Total	347	100

Table 8 displays that 76.4% respondents believed that children who live in poverty are more likely to commit crime and 23.6% disagreed to this opinion. The results examined that majority of respondents believed that children who live in poverty are more likely to become victims and commit crime as compare to others.

Table 9
Steps to Break the Cycle of Poverty

Variable	Categories	Frequency	Percentage
How did you think we can break the cycle of poverty	Educate children	128	36.9
	Job opportunities	147	42.4
	Proper Guidance about drawbacks of Crimes	72	20.7
	Total	347	100

Table 9 demonstrates that 36.9% respondents believed that if children get adequate education then the cycle of poverty can be break while 42.4% believed that more job opportunities are required in order to break the cycle of poverty and 20.7% believed that proper guidance about consequences of crimes are much more important to break the cycle of poverty. According to the results job opportunities play a major role to overcome poverty.

Table 10
Stealing in Hunger

Variable	Categories	Frequency	Percentage
Do you think that when people feel hungry and they find no money to buy anything then they will steal	Strongly Agree	83	23.9
	Agree	156	45.0
	Neutral	106	30.5
	Disagree	1	3
	Strongly Disagree	1	3
	Total	347	100

Table 10 demonstrates that 23.9% respondents strongly agreed to the fact that when people hungry and they find no money to buy anything then they steal while the percentage of agreed respondents is 45.0%. In addition, 30.5% ticked the neutral option and finally 6% respondents are disagreed that people feel hungry and they find no money to do anything then they steal.

Table 11
Punishment for Stealing

Variable	Categories	Frequency	Percentage
Do you think a man who steals for his family should go to jail	Strongly Agree	85	24.5
	Agree	160	46.1
	Neutral	100	28.8
	Disagree	1	3
	Strongly Disagree	1	3
	Total	347	100

Table 11 demonstrates that people should be punished if they commit crime. The results showed that 70% respondents are strongly agreed that a man who steals for his family should go to jail, while 28.8% gave a neutral decision that a man who steal for his family should go to jail and only 6% disagreed from this penalty. The results clearly depict that majority of the people believed and agreed that a man who steals for his family should go to jail. Stealing is highly prohibited; rules and regulations should always be same for everybody. There are some social laws designed by the government and everybody should follow it. So, stealing for any purpose is supposed to be a crime and should be prohibited.

Table 12
Job/Stealing

Variable	Categories	Frequency	Percentage
Do you agree with the statement that "A job offers little payment, whereas stealing offers a large amount of money"	Yes	87	25.1
	No	193	55.6
	Sometimes	66	19.0
	Total	347	100

Table 12 shows that 25.1% agreed to the statement that a job offers with little payment whereas, stealing offers a large amount of money while 55.6% disagreed to this fact. The results revealed that, since people get few amount of income while doing job while robbery offers large money that's why people prefer it.

Table 13
Chi-Square Test for Testing Relationship between Variables

Attributes	χ^2 -value	p-value	Conclusion
Poverty is the main cause of crime and most likely group to crime (upper, middle, lower etc.)	17.664	0.014	Significant
Our living life style and break the cycle of poverty	14.763	0.005	Significant
Children who live in poverty and Main cause of crime	8.297	0.040	Significant

Table 13 shows that there is association between poverty is the main cause of crime and the most likely group to commit crime such as upper, middle and lower class. The results indicates that value of χ^2=17.664 is significant as p-value is 0.014 which is less than the alpha value (0.05). The majority of the culprits belong to the lower class as they steal and adore wrong ways to end up their poverty. Moreover, the second value of χ^2=14.763 also provide significant result because p-value is less than significance level. It demonstrates that there is positive association between effect of poverty on living life style and steps to break the cycle of poverty like educate children, increase job opportunities etc. Most of the people think that poverty effect their lifestyle in a way of bad reputation. So, to maintain our living style different steps should be taken to break the poverty. Because of poverty reputation become low in society. So, in order to break the cycle of poverty we have to give the job opportunities or training to those people who are jobless or poor. Furthermore, the third value of χ^2=8.297 with p-value = 0.040 illustrates association between children who live in poverty and main cause of crime. It means that the poverty is the main cause of crime and because of poverty most of the children live in poverty. Since, they have no source of income to reduce their hunger. So, in order to fulfill their needs children, commit crimes.

5. CONCLUSION

A situation in which a person or household lacks the resources necessary to be able to consume a certain minimum baskets of goods is known as “Poverty”. This world is well populated with a large numbers of beggars, minimum wage workers and the financially unfortunate. Poverty is turned a blind eye to, in modern day and this is its impact on society is gone unnoticed, though in reality its effects are inconceivable. Poverty is one of the main driving gears on the road of crime and its high time this is brought to light.

The purpose of the research is to demonstrate the causes of crimes and poverty in Pakistan. The research is conducted through survey method and data is collected using questionnaire. The results examined that 91.6% people believed that poverty is the main cause of crime while rest of the people disagreed. Moreover, relationship between different variables is calculated by using chi-square test. The value of $\chi^2 = 17.664$ illustrates that there is significant association between main causes of crime like hunger, lack of education and most likely group to commit crime (upper, middle, lower). Therefore, it is concluded that poverty correlates with crime and is more often seen in poor people who are more attracted towards evil ways in order to fulfill their basic needs. Such evil ways lead to astray and crimes are often accompanied.

SUGGESTIONS

The research suggests that there are several ways to overcome poverty and this can only happen if government pays honest attention to this. The government should take severe actions against any provoking crimes happening in society. So, it’s high time that the government should enforce such laws that can help ending this nuisance.

1. Create jobs.
2. Raise the minimum wage.
3. Increase the Earned Income Tax Credit for childless workers.

4. Support pay equity.
5. Establish work schedules that work.
6. Invest in affordable, high-quality child care and early education.
7. Expand Medicaid.
8. Reform the criminal justice system and enact policies that support successful re-entry.
9. Do not harm.

REFERENCES

1. Kelly, M. (2000). Inequality and Crime. *Review of Economics and Statistics,* 82(4), 530-539.
2. Melham, H, Moene, K. and Torvik, R. (2005). Crime induced poverty traps. *Journal of Development Economics,* 77(2), 325-340.
3. Huang, C.C., Laing, D. and Wang, P. (2004). Crime and poverty: A search-theoretic approach. *International Economic Review,* 45(3), 909-938.
4. Mauro, L. and Carmeci, G. (2007). A poverty trap of crime and unemployment. *Review of Development Economics,* 11(3), 450-462.
5. Hannon, L. (2002). Criminal opportunity theory and the relationship between poverty and property crime. *Sociological Spectrum,* 22(3), 363-381.
6. Ludwig, J., Duncan, G.J. and Hirschfield, P. (2001). Urban poverty and juvenile crime: Evidence from a randomized housing-mobility experiment. *The Quarterly Journal of Economics,* 116(2), 655-679.
7. Pare, P.P. and Felson, R. (2014). Income inequality, poverty and crime across nations. *The British Journal of Sociology,* 65(3), 434-45.
8. Pridemore, W.A. (2008). A Methodological Addition to the Cross-national Literature on Social Structure and Homicide. A First Test of the Poverty-Homicide Thesis. *Criminology*, 46, 133-54.
9. Pridemore, W.A. (2011). Poverty Matters: A Reassessment of the Inequality-Homicide Relationship in Cross-national Studies. *British Journal of Criminology*, 51, 739-72.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 119-144

MODIFIED RIDGE REGRESSION ESTIMATORS: A SIMULATION STUDY

Muhammad Suhail[1§], Maryam Ilyas[2] and **Ayan Ullah[3]**
[1] Department of Statistics, The University of Agriculture
Amir Muhammad Khan Campus Mardan, Pakistan
[2] College of Statistical and Actuarial Sciences,
University of Punjab, Lahore, Pakistan
[3] Directorate of Information and Public Relations,
Peshawar, Pakistan
[§] Corresponding author: msuhail88@aup.edu.pk

ABSTRACT

Ridge regression is used to circumvent the effects of multicollinearity. This is done by introducing biasing constant 'k', called ridge parameter in the least square objective function. Ridge parameter shrinks the estimates and their variances. Selection and choice of the unknown ridge parameter 'k' is of prime importance in ridge regression analysis.

Khalaf, Mansson and Shukur (2013) proposed some modifications of existing ridge estimators k_1– k_{16} by multiplying them with the factor that make use of maximum eigenvalue associated with $\boldsymbol{(X^tX)}$ matrix and name resulting estimators as K1M–K16M. A new sets of ridge estimators are proposed in this study making use of geometric mean of eigenvalues associated with $\boldsymbol{(X^tX)}$ matrix denoted as K1G–K16G. The comparative performance of proposed sets of estimators and Khalaf et al. (2013) was evaluated by Mean Square Error (MSE) using simulated data sets. Data sets considering different levels of collinearity (r), sample size (n), number of predictor (p), error term variances and error term distributions were generated. It was observed that proposed estimators K1G–K16G outperform K1M–K16M when error terms were normally distributed for σ^2= 0.1, r = 0.80, 0.90, 0.95 and 0.99, p = 2, 4, 6 and 12 and when error terms were non-normally distributed, i.e., F (4, 20) for r = 0.80, 0.90 and p = 2, 4.

1. INTRODUCTION

In regression analysis, usually we consider that the predictors are not related to each other. In practice, there may be some type of relationships among the predictor variables. In this case the assumption of independence of predictors is no longer valid; violation of this assumption causes the problem of multicollinearity. Regression analysis is most powerful statistical tool that helps in investigating the relationships between response variable and explanatories. Prediction and description mainly depend on the estimated regression coefficients. Least squares method is the mostly used method for estimating the unknown regression coefficients. It gave good estimates only if the assumption of independence of explanatories is valid. The assumptions are that the explanatory variables are independent from each other and this is very difficult to hold in reality. When the purpose is to get more information about the outcome variable, there is need to add more predictors to regression model. By doing so, relationships between these variables occur and the magnitude of these relationships often increases. This type of linear relationships between the predictors is called the problem of multicollinearity. Chatterjee and Hadi (2006) and Gujarati (2003) highlighted that with the existence of multicollinearity in a data set, two or more explanatories give same or approximately same information. The existence of multicollinearity among explanatories causes many problems. It affects the model's ability to estimate unknown regression coefficients, t-test, computational accuracy, variance of LS estimated regression coefficients, and LS estimated regression coefficients, fitted values and predictions.

In order to overcome the problem of multicollinearity in multiple linear regression model among explanatories, Hoerl and Kennard (1970) suggested ridge regression method instead of OLS method in regression analysis. When explanatories are correlated, then for estimation of regression coefficients ridge regression is a technique that was initially recommended by Hoerl and Kennard.

Multiple linear regression model can be written in matrix form as,

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\epsilon}, \tag{1.1}$$

$\mathbf{y}$ is a vector of dependent variables with order n×1, $\mathbf{X}$ is matrix of explanatory of order n×p, $\boldsymbol{\beta}$ is a vector of unknown regression coefficients of order p×1 and $\boldsymbol{\epsilon}$ is vector of random errors of order n×1 that are distributed normally whose mean vector is zero while it's covariance matrix is $\sigma^2\mathbf{I}_n$ ($\mathbf{I}_n$ is identity matrix of n×n order). The OLS of the regression coefficients $\boldsymbol{\beta}$ is $\widehat{\boldsymbol{\beta}}_{OLS} = (\boldsymbol{X}^t\boldsymbol{X})^{-1}\boldsymbol{X}^t\boldsymbol{y}$, and variance-covariance matrix of $\boldsymbol{\beta}$ is Var ($\widehat{\boldsymbol{\beta}}_{OLS}$)= $\sigma^2(\boldsymbol{X}^t\boldsymbol{X})^{-1}$, both $\widehat{\boldsymbol{\beta}}$ and Var ($\widehat{\boldsymbol{\beta}}$) depend on characteristics of $\boldsymbol{X}^t\boldsymbol{X}$ matrix. If matrix $\boldsymbol{X}^t\boldsymbol{X}$ is near to singular then the variances of ordinary least square (OLS) estimates becomes large. To overcome this difficulty Hoerl and Kennard (1970) discovered RR (Ridge Regression) method and proposed that a small k positive number (≥0) to be added to diagonal of $\boldsymbol{X}^t\boldsymbol{X}$ matrix to counter the effects of multicollinearity.

The most popular method is Ridge regression which was developed by Hoerl and Kennard (1970). Draper and Smith (1981) stated that as a result of multicollinearity, the $\boldsymbol{X}^t\boldsymbol{X}$ matrix is near ill conditioned (singular) that leads to large standard errors for ordinary least squares (OLS) estimates. To solve this problem Hoerl and Kennard (1970) added a positive number (k) to the main diagonal of the matrix $\boldsymbol{X}^t\boldsymbol{X}$ such that the new estimates are,

$$\widehat{\boldsymbol{\beta}}_{\boldsymbol{RR}} = (\boldsymbol{X}^t\boldsymbol{X} + k\mathbf{I}_p)^{-1}\boldsymbol{X}^t\boldsymbol{y},\ k \geq 0 \tag{1.2}$$

This is called an RR estimator. For any positive value of 'k', this gave minimum mean square error (MMSE) as compared to LSE. The 'k' is known as "Ridge" or "biasing" parameter (constant) and will be finding out from data. When k=0, $\widehat{\boldsymbol{\beta}}_{\boldsymbol{RR}}$ becomes the ordinary least square estimates (OLSE) and as 'k' increases more bias is introduced but variance of the regression estimator stabilizes.

Now the MSE of ridge regression that is introduced by Hoerl and Kennard (1970) is defined as:

$$MSE\left(\hat{\beta}(K)\right) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i+k_i)^2} + \sum_{i=1}^{p} \frac{k^2{}_i\alpha^2{}_i}{(\lambda_i+k_i)^2} \tag{1.3}$$

The 1st term on right hand side of (1.3) is a variance and the second term is an amount of bias, where

$$k_i = \frac{\sigma^2}{\alpha_j^2} \tag{1.4}$$

σ^2 is the variance of the model (1.3) and α_i is ith element of α. The unbiased estimator of $k_i$is,

$$\hat{k}_i = \frac{\hat{\sigma}^2}{\hat{\alpha}^2{}_i} \tag{1.5}$$

where $\hat{\sigma}^2 = \frac{(\boldsymbol{y}-\hat{\boldsymbol{y}})^t\,(\boldsymbol{y}-\hat{\boldsymbol{y}})}{(n-p-1)}$ is the residual sums of square obtained from the OLS and is an unbiased estimator of σ^2 and $\hat{\alpha}^2{}_i$ is the ith elements of $\hat{\alpha}^2$ where $\hat{\boldsymbol{\alpha}} = \boldsymbol{V}^t\hat{\boldsymbol{\beta}}$, **V** is orthogonal matrix of order (p×p); the columns of **V** are the normalized eigenvectors of correlation matrix. The MSE of OLS is defined as:

$$MSE(\hat{\boldsymbol{\beta}}) = variance\,(\hat{\boldsymbol{\beta}}) + (bias(\hat{\boldsymbol{\beta}}))^{\mathbf{2}} \tag{1.6}$$

OLS estimators are generally unbiased so the 2nd term of right hand side of above equation (1.6) is zero and becomes,

$$MSE(\hat{\boldsymbol{\beta}}) = variance\,(\hat{\boldsymbol{\beta}}) \tag{1.7}$$

Many methods for estimation of ridge parameter 'k' have been described by many researchers. In which some well- known existing estimators are following. These estimators make use of the canonical form of regression model.

The canonical form of model (1.3) is (1.8). Consider orthogonal matrix **D** where $\boldsymbol{D}^T\boldsymbol{C}\boldsymbol{D} = \boldsymbol{\Lambda}$, where $\boldsymbol{C} = \boldsymbol{X}^T\boldsymbol{X}$ and $\boldsymbol{\Lambda} = diag(\lambda_1, \lambda_2, \dots, \lambda_p)$ containing eigenvalues 'λ_i' of matrix **C**. Model (1.3) in canonical form is,

$$\boldsymbol{y} = \boldsymbol{X}^*\boldsymbol{\alpha} + \boldsymbol{\epsilon} \tag{1.8}$$

$\boldsymbol{X}^* = \boldsymbol{X}\boldsymbol{D}$ and $\boldsymbol{\alpha} = \boldsymbol{D}^T\boldsymbol{\beta}$.

The least square estimators of the canonical form is,

$$\hat{\boldsymbol{\alpha}} = \boldsymbol{\Delta}^{-1}\boldsymbol{X}^{*T}\boldsymbol{y} \tag{1.9}$$

Ridge estimators in canonical form is

$$\hat{\boldsymbol{\alpha}}(k) = (\boldsymbol{X}^{*T}\boldsymbol{X}^* + K\boldsymbol{I})^{-1}\boldsymbol{X}^{*T}\boldsymbol{y} \tag{1.10}$$

K $= diag(\lambda_1, \lambda_2, \dots, \lambda_p)$ that is introduce by Hoerl and Kennard (1970). MSE of the above estimators which Hoerl and Kennard defined as:

$$MSE(\hat{\alpha}(k)) = \sigma^2 \sum_{i=1}^{p} \frac{\lambda_i}{(\lambda_i + k_i)^2} + \sum_{i=1}^{p} \frac{{k_i}^2{\alpha_i}^2}{(\lambda_i + k_i)^2} \tag{1.11}$$

First term on RHS (1.11) is variance and second term is amount of bias.

2. METHODOLOGY

It is evident from the study that multicollinearity is an important phenomenon in regression analysis that cannot be ignored while analysing data. Many researchers made different studies on this topic and it can be further explored. Different researchers have recommended different methods and estimators to overcome the problem of multicollinearity. In this study, some successful extensions of the existing work have been proposed to deal with multicollinearity problem.

Hoerl and Kennard Estimator

Hoerl and Kennard (1970) explored value of 'k' which minimizing the mean square error (MSE) is the following**:**

$K_1 = \hat{k}_{HK} = \frac{\hat{\sigma}^2}{\hat{\alpha}^2{}_{max}}$. Where $\hat{\alpha}^2{}_{max}$ is the square of the maximum value of $\hat{\boldsymbol{\alpha}}$.

Kibria Estimator

Kibria (2003) proposed the following estimators,

$$K_2 = \hat{k}_{GM} = \frac{\hat{\sigma}^2}{(\prod_{j=1}^{p} \hat{\alpha}^2{}_j)^{\frac{1}{p}}}\text{. And } K_3 = \hat{k}_{MED} = Median\{m^2{}_j\} \text{ where } m_j = \sqrt{\frac{\hat{\sigma}^2}{\hat{\alpha}^2{}_j}}.$$

Khalaf and Shukur Estimator

Khalaf and Shukur (2005) proposed a new estimator as a modification of k_{HK}

$$K_4 = \hat{k}_{KS} = \frac{t_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_{max}\hat{\alpha}^2{}_{max}}.$$

Where t_{max} is maximum eigenvalue of matrix X^tX.

Alkhamisi, Khalaf and Shukur Estimator

Alkhamisi et al. (2006) suggested that,

$$K_5 = k_{S3} = \hat{k}^{KS}{}_{max} = \max(s_j);\ K_6 = \hat{k}^{KS}{}_{md} = \mathrm{med}(s_j),$$

where $s_j = \frac{t_j\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_j\hat{\alpha}^2{}_j}$.

Alkhamisi and Shukur Estimator

Alkhamisi and Shukur (2008) suggested the estimators for 'k'as,

$$K_7 = k_{KM1} = \hat{k}^{KS}{}_{gm} = (\textstyle\prod_{j=1}^{p} s_j)^{\frac{1}{p}},\ K_8 = k_{KM2} = \max\left(\frac{1}{m_j}\right),$$

$$K_9 = k_{KM4} = (\textstyle\prod_{j=1}^{p} \frac{1}{m_j})^{\frac{1}{p}},\ K_{10} = k_{KM5} = (\prod_{j=1}^{p} m_j)^{\frac{1}{p}},$$

$$K_{11} = k_{KM6} = \mathrm{median}\left(\frac{1}{m_j}\right),\ K_{12} = k_{KM8} = \max\left(\frac{1}{q_j}\right),$$

$$K_{13} = k_{KM9} = \max\left(\sqrt{q_j}\right),\ K_{14} = k_{KM10} = (\textstyle\prod_{i=1}^{p} \frac{1}{\sqrt{q_j}})^{\frac{1}{p}};$$

$$K_{15} = k_{KM11} = (\prod_{i=1}^{p} \sqrt{q_j})^{\frac{1}{p}},\ K_{16} = k_{KM12} = median\left(\frac{1}{\sqrt{q_j}}\right).$$

where $q_j = \sqrt{\frac{t_{max}\hat{\sigma}^2}{(n-p)\hat{\sigma}^2 + t_{max}\hat{\alpha}^2{}_j}}$.

Khalaf, Mansson and Shukur Estimators

Khalaf et al. (2013) proposed modifications of all of the above estimators by multiplying them by a factor,

$$w_j = \frac{t_{max}}{\sum_{j=1}^{p} |\hat{\alpha}|_j}$$

t_{max} is the maximum eigenvalue of $\boldsymbol{X^t X}$ matrix. This modification was proposed on the basis that as degree of correlation increases initial eigenvalues are larger than others. Thus, factor w_j will also become larger as it is based on the maximum eigenvalue of $\boldsymbol{X^t X}$ matrix. This will lead to an increase of the estimated value of the ridge parameters 'k'. Hence, this modification leads to larger values of the ridge parameter especially when the degree of correlation is high. These modified estimators were denoted by K1M–K16M. The performance of these estimators was good for high collinearity level.

New Proposed Estimators

Khalaf et al. (2013) proposed modifications by multiplying K_1– K_{16}estimators by a factor w_j which is based on maximum eigenvalue.

The trend of eigenvalues associated with $\boldsymbol{X^t X}$ matrix was explored and it was identified that for very high multicollinearity (0.95, 0.99) the maximum eigenvalues is larger than the others, however, as the level of multicollinearity decreases from the said level (0.95, 0.99), the difference between the maximum eigenvalue and remaining eigenvalues decreases. Hence, it was expected that some other ridge estimators have potential to be explored in view of this phenomenon.

Thus, instead of using maximum eigenvalue of $\boldsymbol{X^t X}$ in the numerator of w_j, Geometric mean of eigenvalues of the $\boldsymbol{X^t X}$ matrix were used. When degree of correlation among regressors were not very high (0.95, 0.99) then there will not be see difference between maximum eigenvalue and the remaining eigenvalues so in this situation geometric mean were expected to give good results.

Thus, a new set of ridge estimators were developed. Set of estimators were proposed by using geometric mean of the eigenvalues (T_{GM}) associated with $\boldsymbol{X^t X}$ and defining,

$$v_{1j} = \frac{T_{GM}}{\sum_{j=1}^{p} |\hat{\alpha}|_j}$$

Finally, the new sets of estimators were defined by multiplying K_1– K_{16} by v_{1j} and the resulting estimators were denoted by K1G –K16G.

2.2 Mean Square Error (MSE)

The performances of Ridge regression estimators have long being compared making use of MSE. Thus, to explore the competitive performance of the new suggested estimators and existing estimators, MSE was used. MSE is defined as,

$$MSE = \sum_{i=}^{N} \frac{(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})_i{}^t (\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta})_i}{N}$$

$\hat{\beta}$ is the estimator of β obtained from RR or OLS and 'N' is number of replications used in Monte Carlo study.

3. THE MONTE CARLO SIMULATION

Theoretically the proposed and the existing estimators cannot be compared, so simulation studies were designed to explore the performance of the developed and already existing ridge estimators.

Kibria (2003), Gibbons (1981), Wichern and Churchill (1978), McDonald and Galarneau (1975), and many other researchers used the following method to simulate or generate the predictor variables that is;

$$x_{ij} = (1 - r^2)^{\frac{1}{2}} z_{ij} + r z_{ip};\ \text{i = 1, 2, 3...n; j =1, 2, 3...p} \quad (1.12)$$

z_{ij} are standard normal random variables, 'r' is level of collinearity between any two explanatories and 'n' is the number of observations. In this study the model that is used is,

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i;\ \text{i=1, 2, 3 ... n} \quad (1.13)$$

β_0 is taken to be zero and $\beta_1, \beta_2, \ldots\ldots, \beta_p$, regression coefficients, are considered so that $\sum_j^p \beta_j^2 = 1$. Simulations studies are carried out using programming language R version 3.1.0.

4. FACTORS AFFECTING RIDGE ESTIMATORS

Various factors can potentially affect the performance of ridge estimators. These include severity of multicollinearity, sample size, number of explanatory variables, error term variance (normal distribution case) and distributions of error terms. Four levels of multicollinearity between any two regressors were considered as high ($r = 0.80, 0.90$) and very high ($r = 0.95, 0.99$). The variation of sample size and number of explanatories considered as n = 30, 70, 100, 150, 200, 300 and p= 2, 4, 6, 12 respectively. In case of normal error term distribution, the variation of error term variance was considered as σ^2= 0.1. To explore the effects of error term distribution, normal and non-normal distributions were considered. For non-normal F- distribution with (4, 20) were considered.

5. SIMULATION RESULTS

In this study a simulation studies have been designed to explore the competitive performance of K1M–K16M and K1G–K16G. The comparisons of these two sets of estimators have been gauged considering different levels of sample size (n=30, 70, 100, 150, 200, 300), number of predictors (p=2, 4, 6, 12), correlation levels (r=0.80, 0.90, 0.95, 0.99) and error term distributions (N $(0, \sigma^2)$ with σ^2=0.1 & Non-Normal distributions F (4, 20)). The performances of these estimators have been evaluated making use of Mean Square Error (MSE). Study compares K1M–K16M and K1G–K16G. Studies addresses two cases; case–1 is pertaining to the combination of the levels of sample size, number of predictors and correlation levels with normal error terms distribution (σ^2=0.1). However, case-2 caters the combinations of sample size, number of predictors and correlation levels with non-normal error terms distributions (F (4, 20)). The results of the simulation studies are summarized by graphs to make comparative performance of all estimators visible in a particular scenario. The results in terms of the tables are maintained as well but are not included in this document to save space. The results of MSE for Case-1 are presented in figures 4.1 ($1-8$) and those of Case-2 are presented in figures 4.2 ($1-8$).

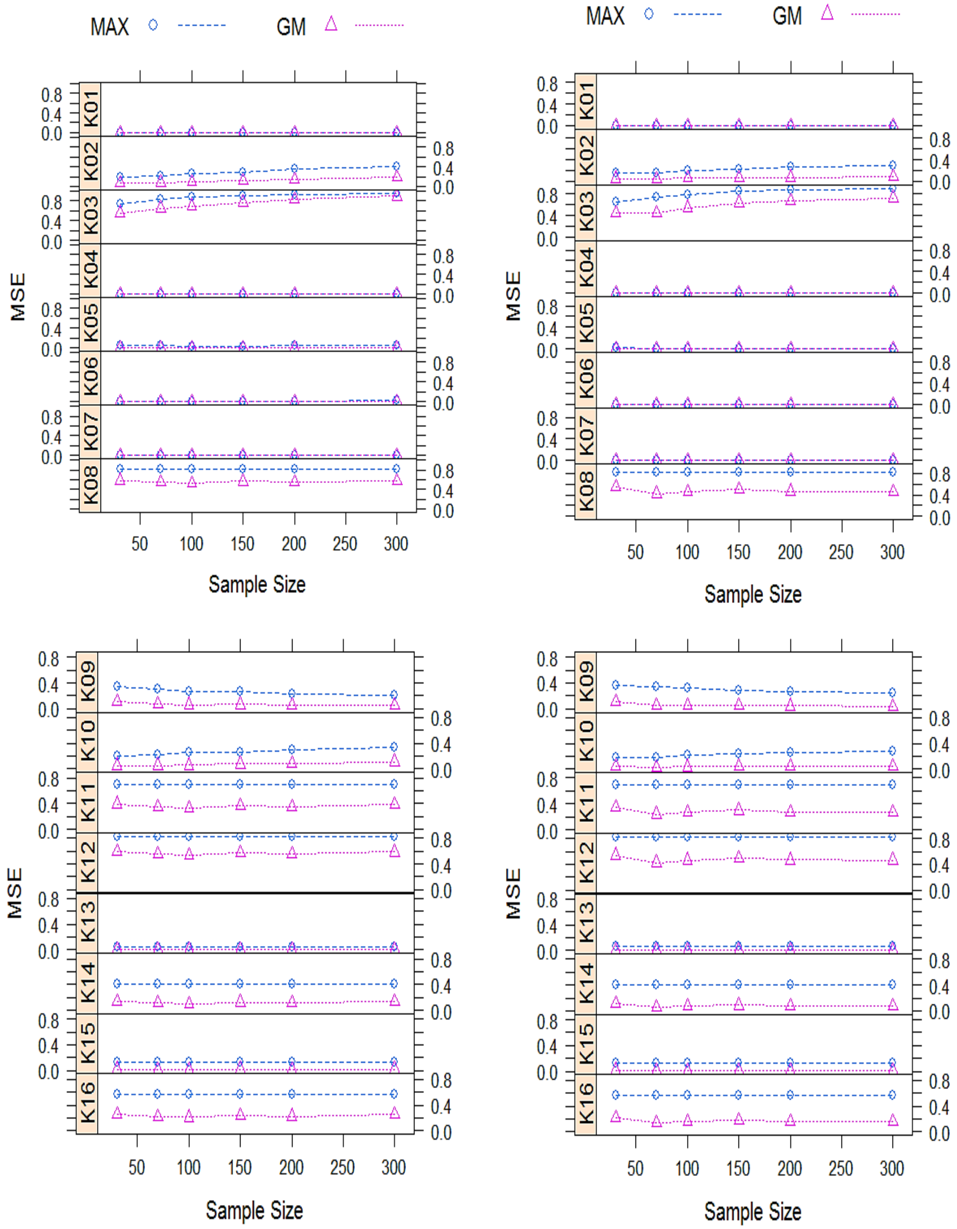


Figure 4.1(1): MSE at p=2 and $\varepsilon_i \sim N(0, 0.1)$. 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

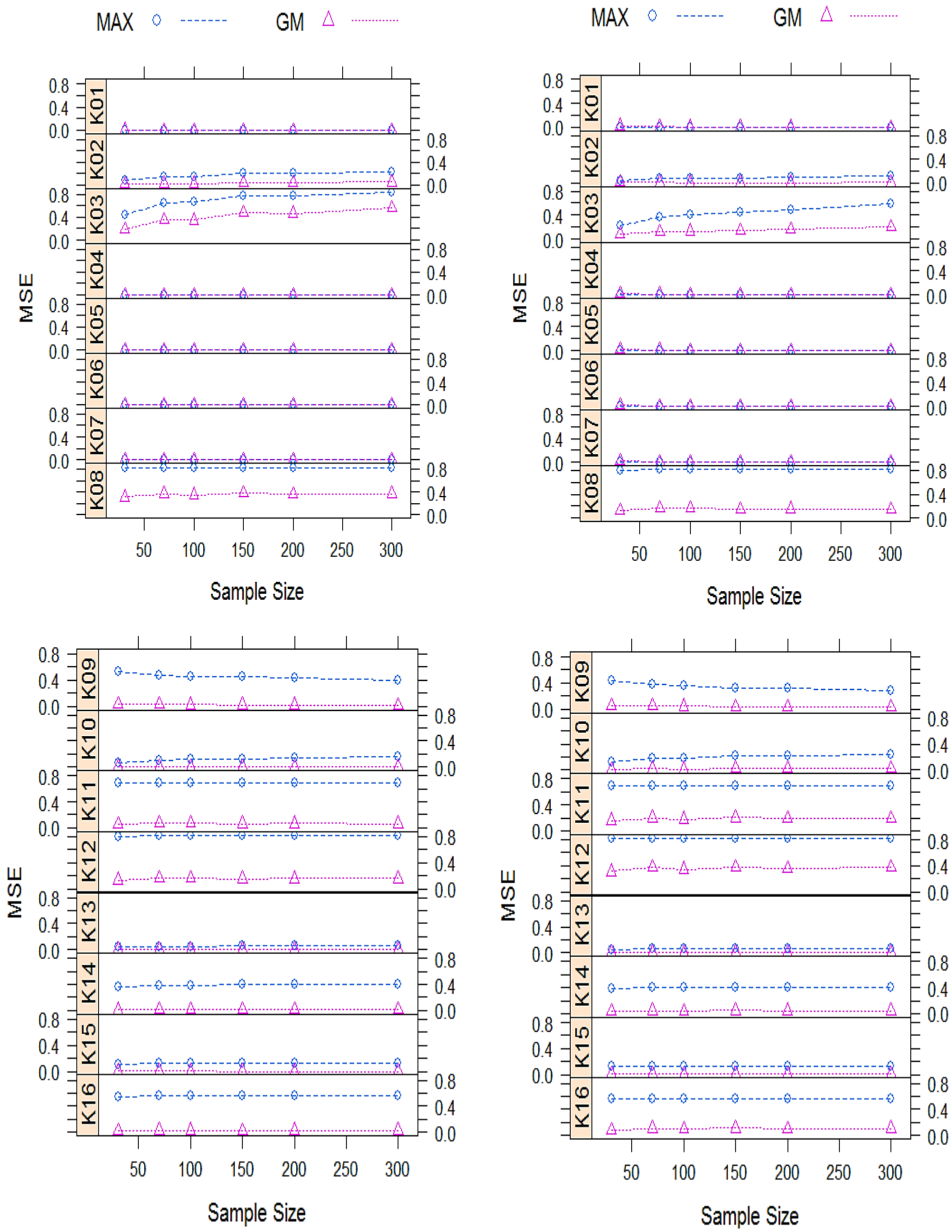


Figure 4.1(2): MSE at p=2 and ε_i~N (0, 0.1). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

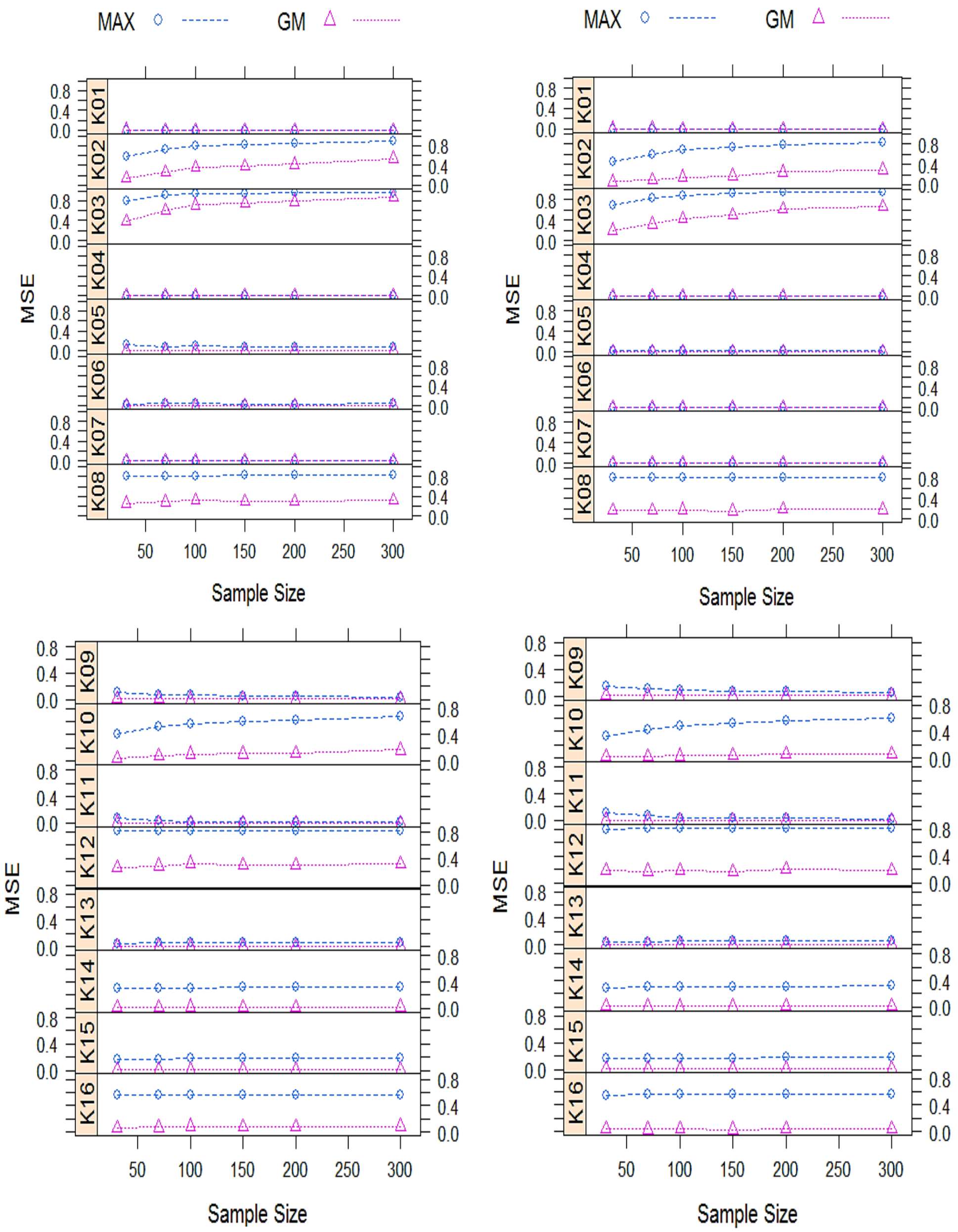


Figure 4.1(3): MSE at p=4 and $\varepsilon_i \sim N\ (0, 0.1)$. 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

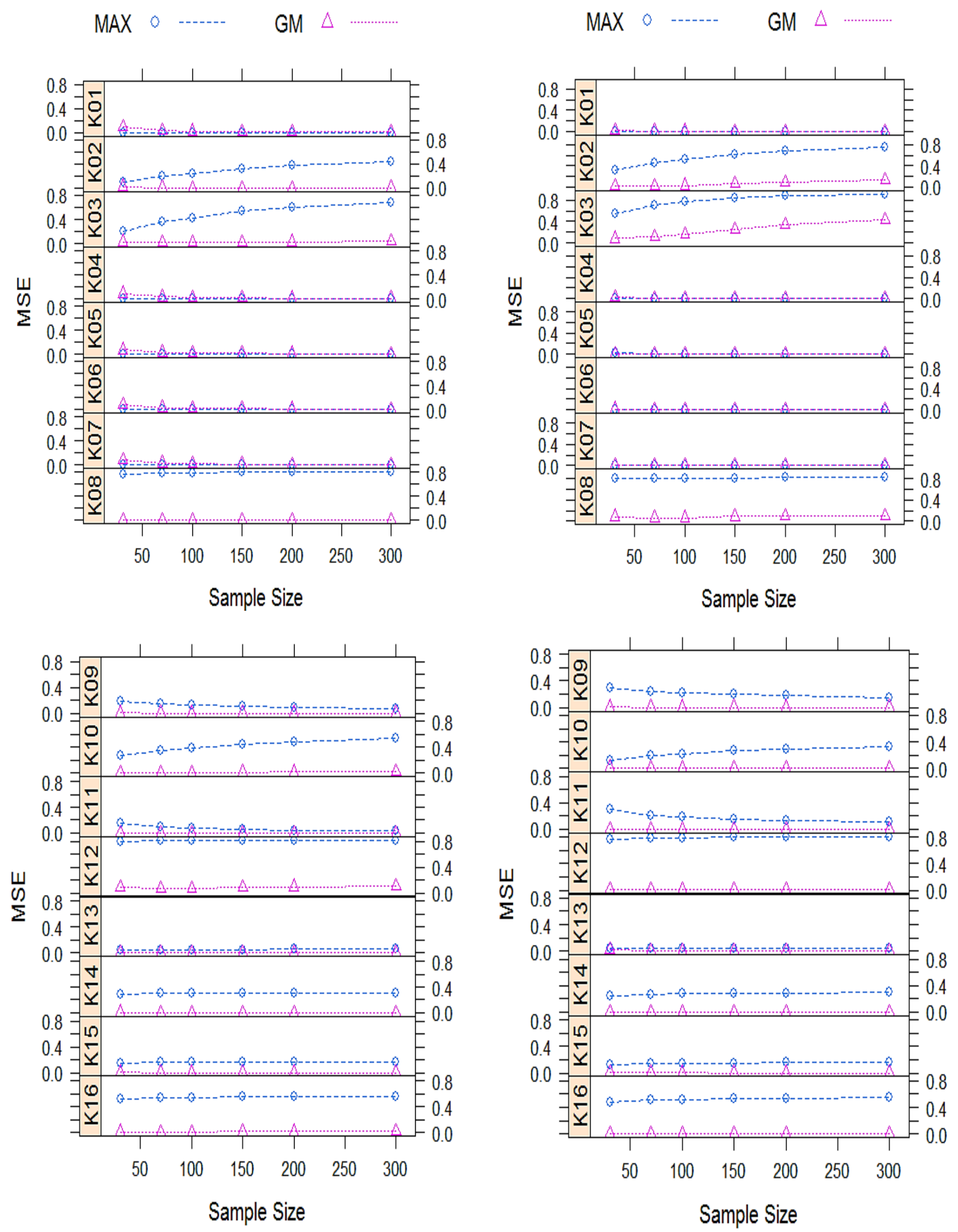


Figure 4.1(4): MSE at p=4 and $\varepsilon_i \sim N(0, 0.1)$. 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

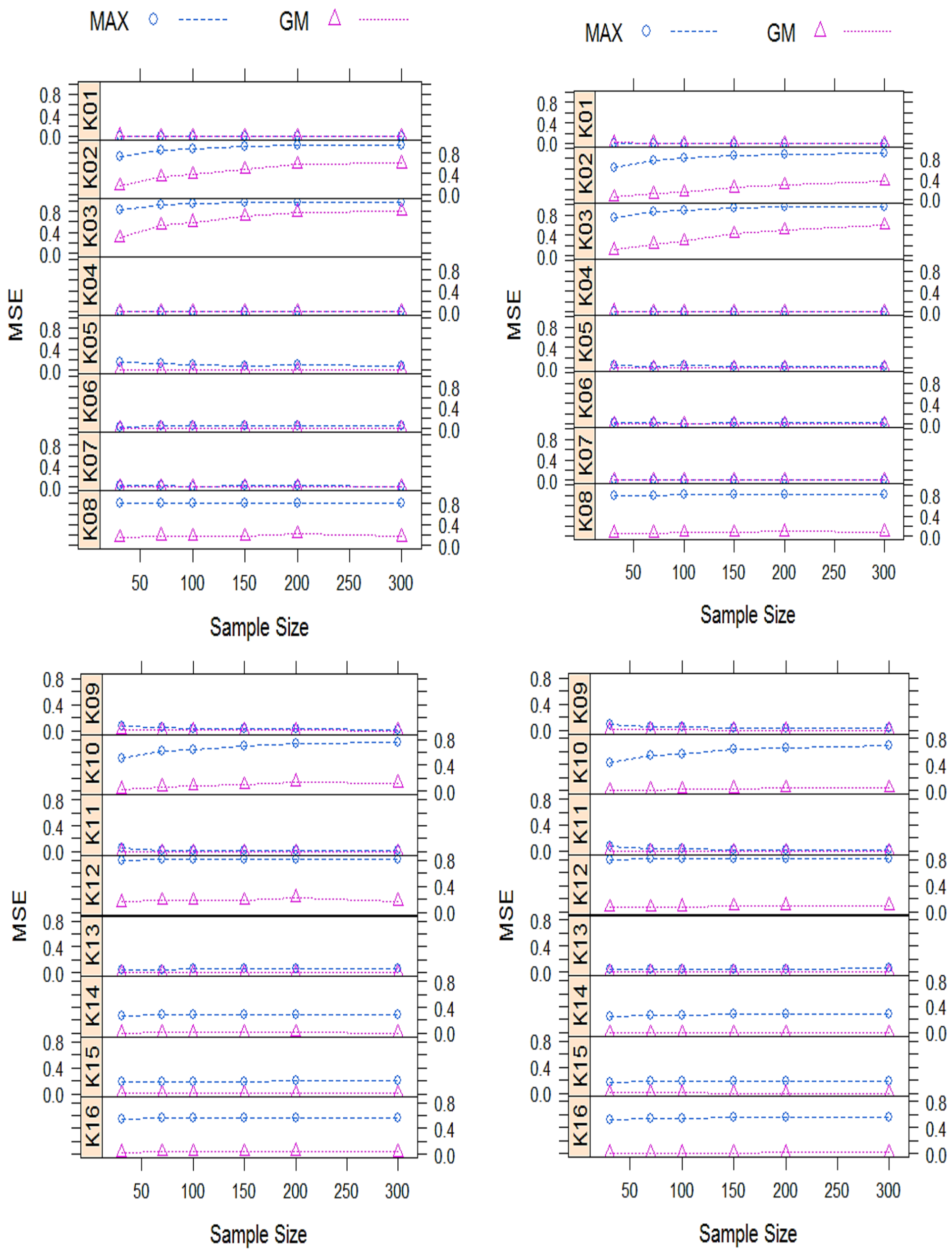


Figure 4.1(5): MSE at p=6 and ε_i~N (0, 0.1). 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

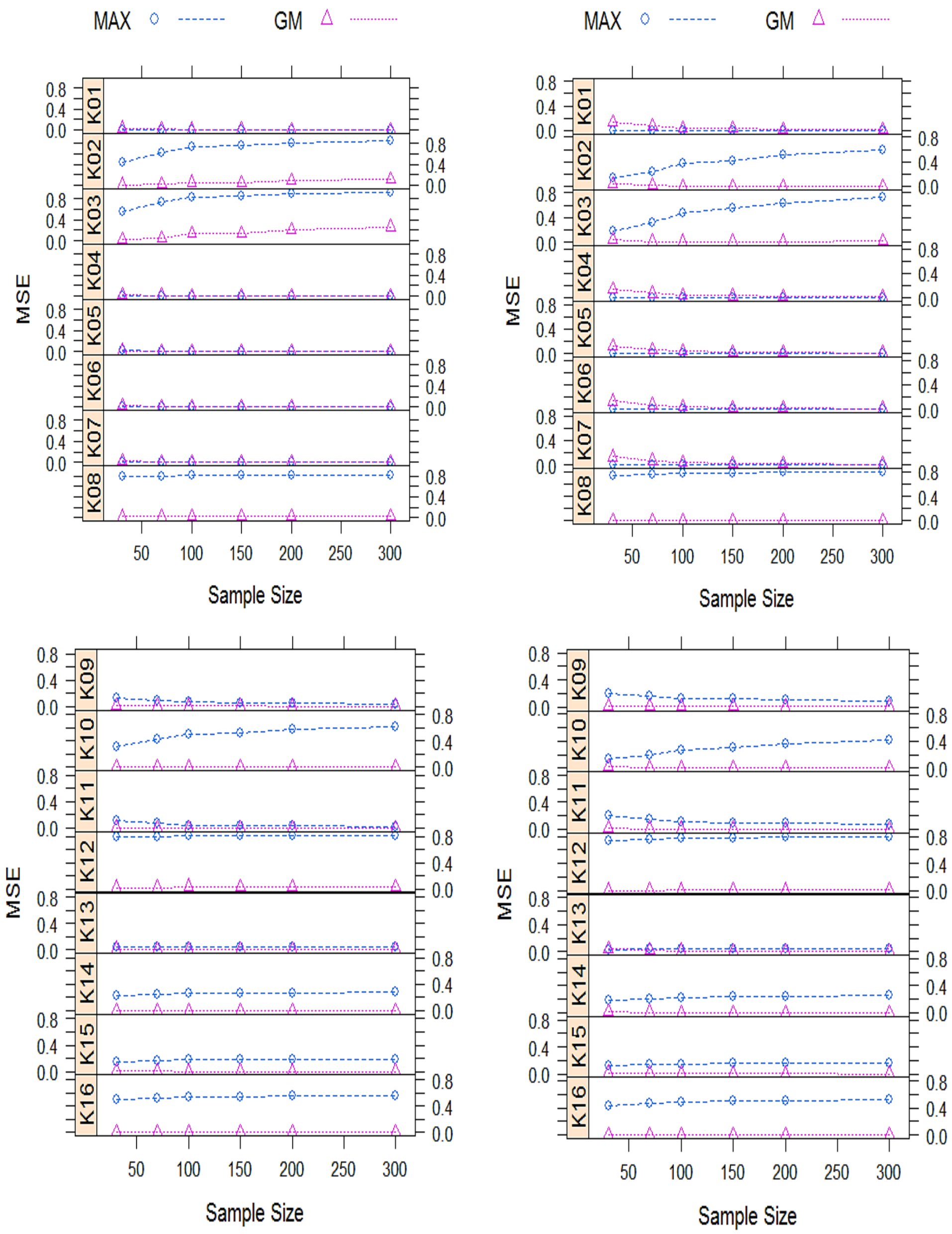


Figure 4.1(6): MSE at p=6 and ε_i~N (0, 0.1). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

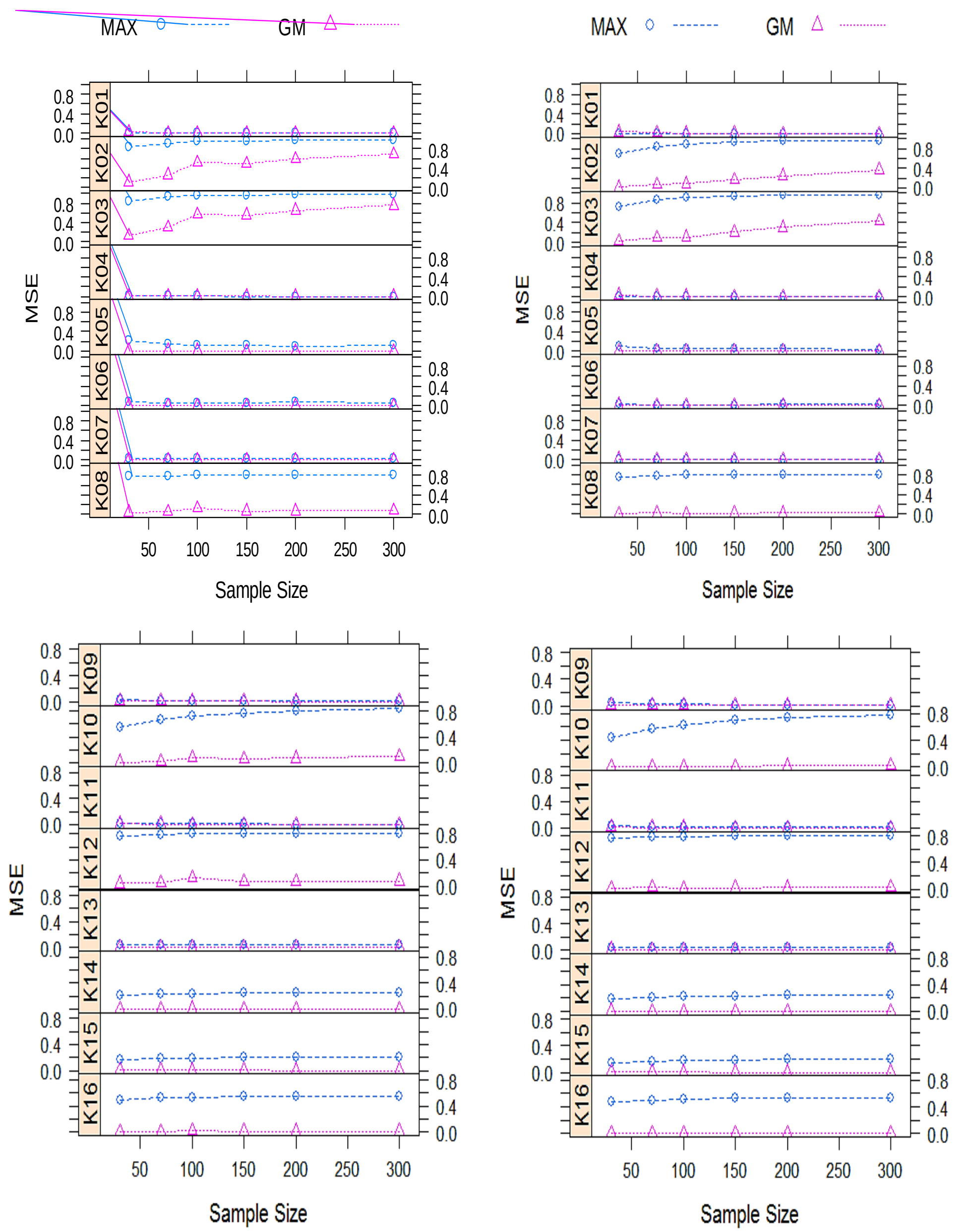


Figure 4.1(7): MSE at p=12 and ε_i~N (0, 0.1). 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

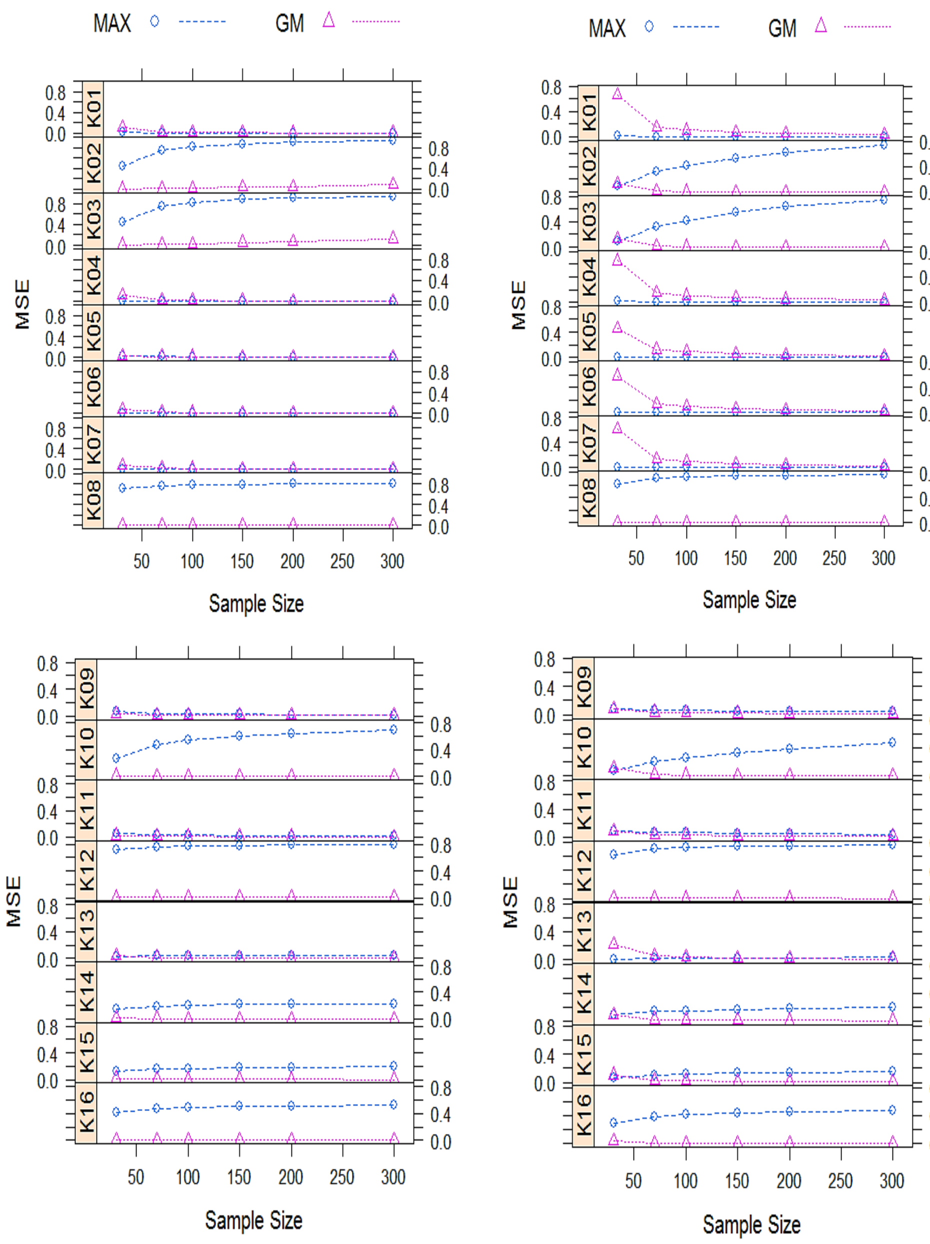


Figure 4.1(8): MSE at p=12 and ε_i~N (0, 0.1). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

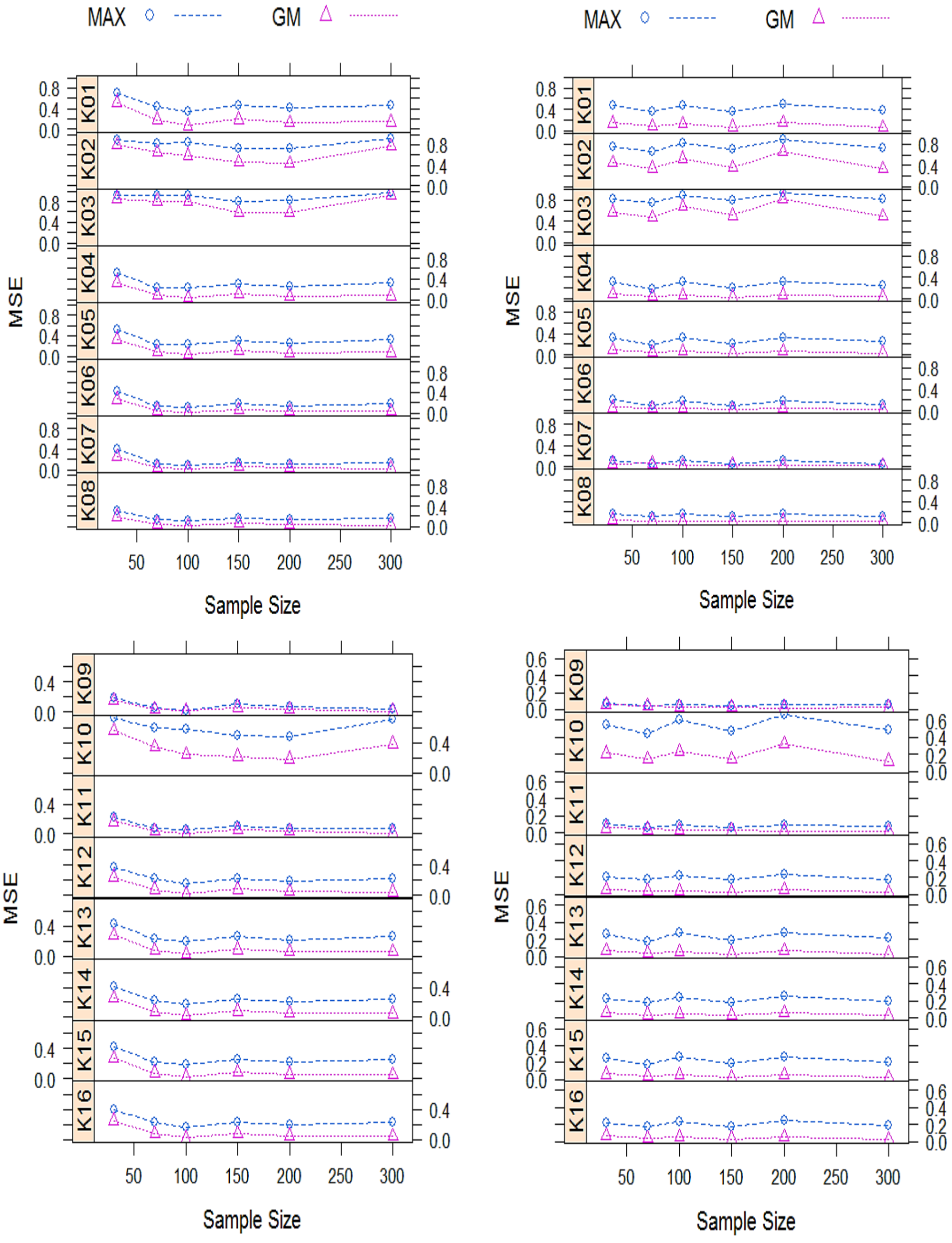


Figure 4.2(1): MSE at p=2 and ε_i~F (4, 20). 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

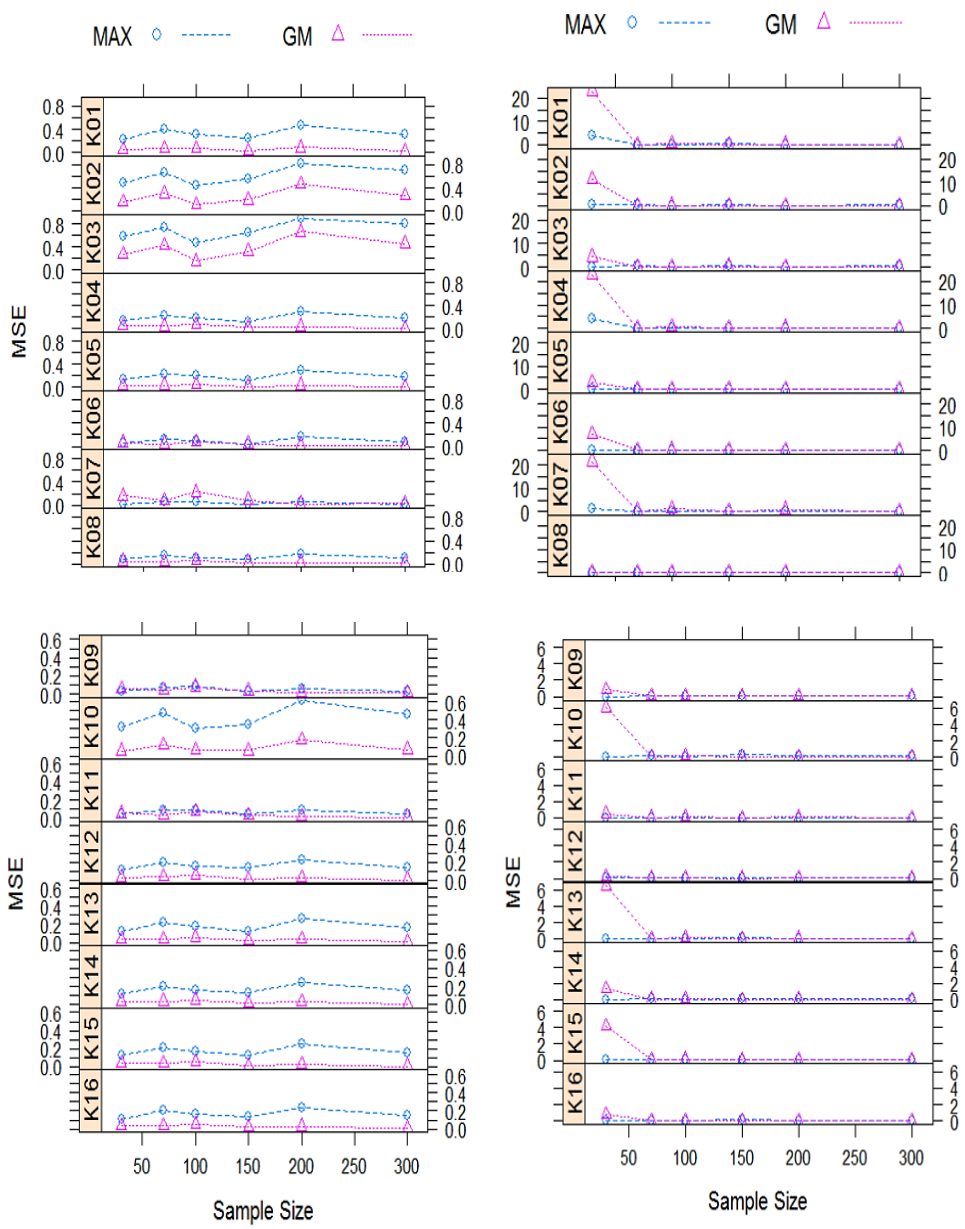


Figure 4.2(2): MSE at p=2 and ε_i~F (4, 20). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

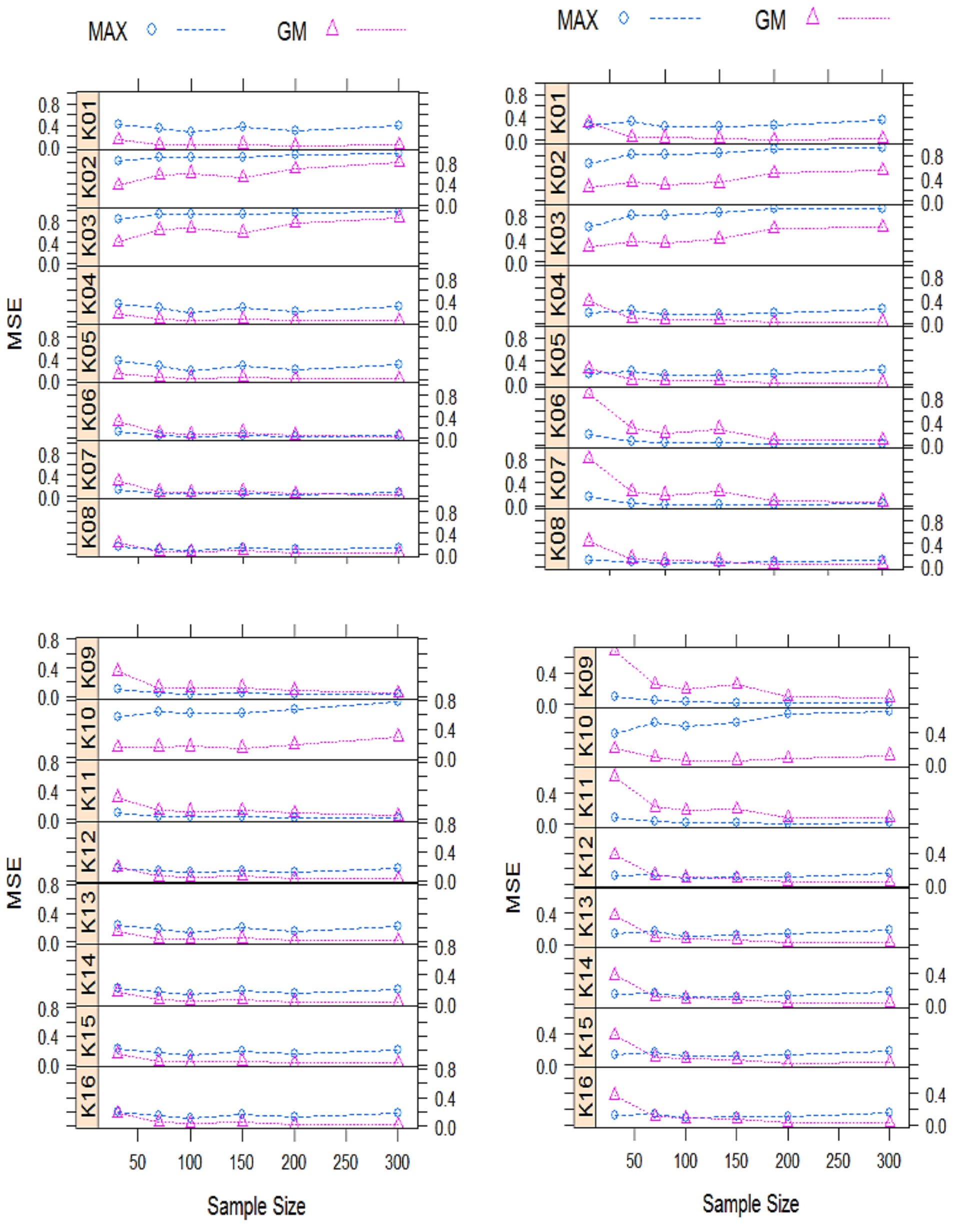


Figure 4.2(3): MSE at p=4 and ε_i~F (4, 20). 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

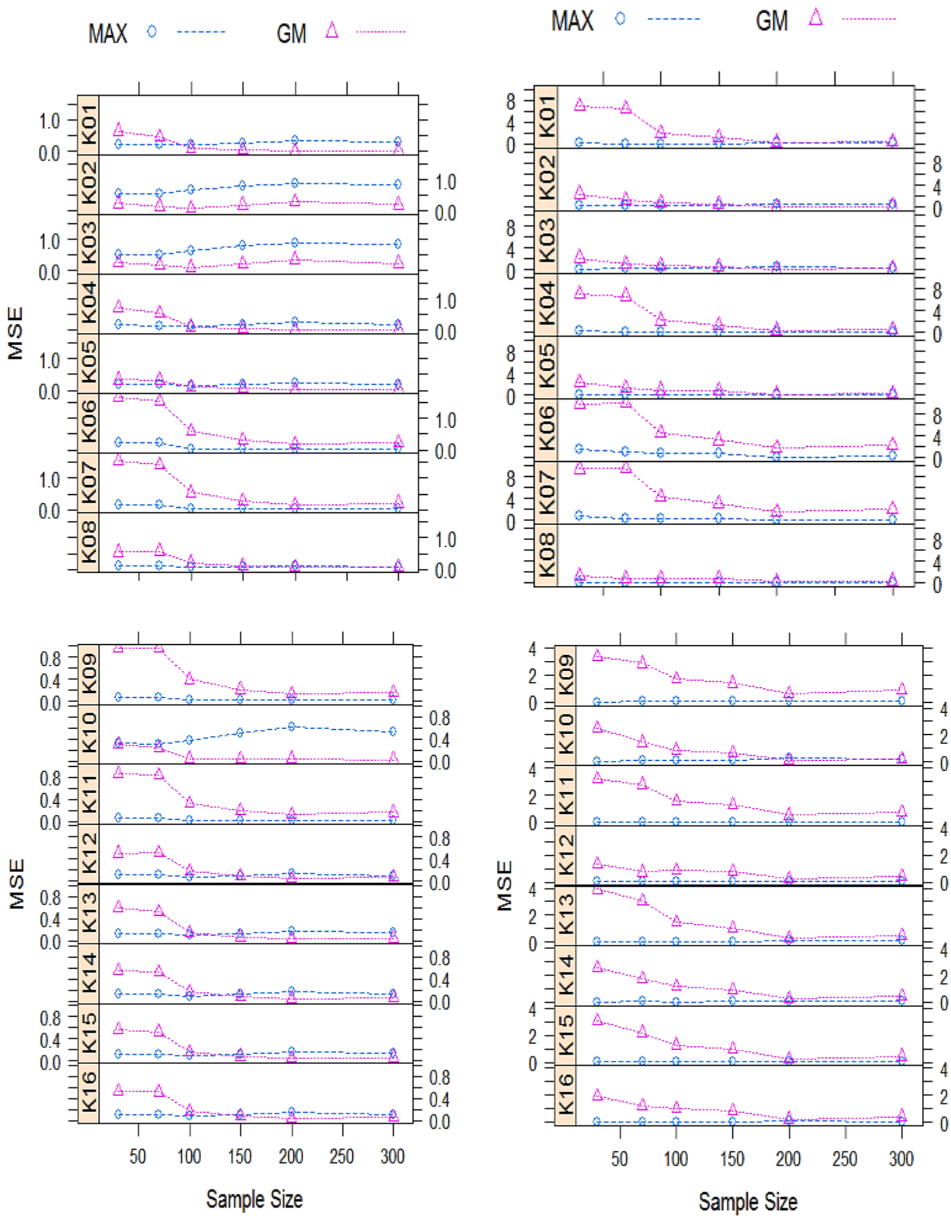


Figure 4.2(4): MSE at p=4 and ε_i~F (4, 20). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

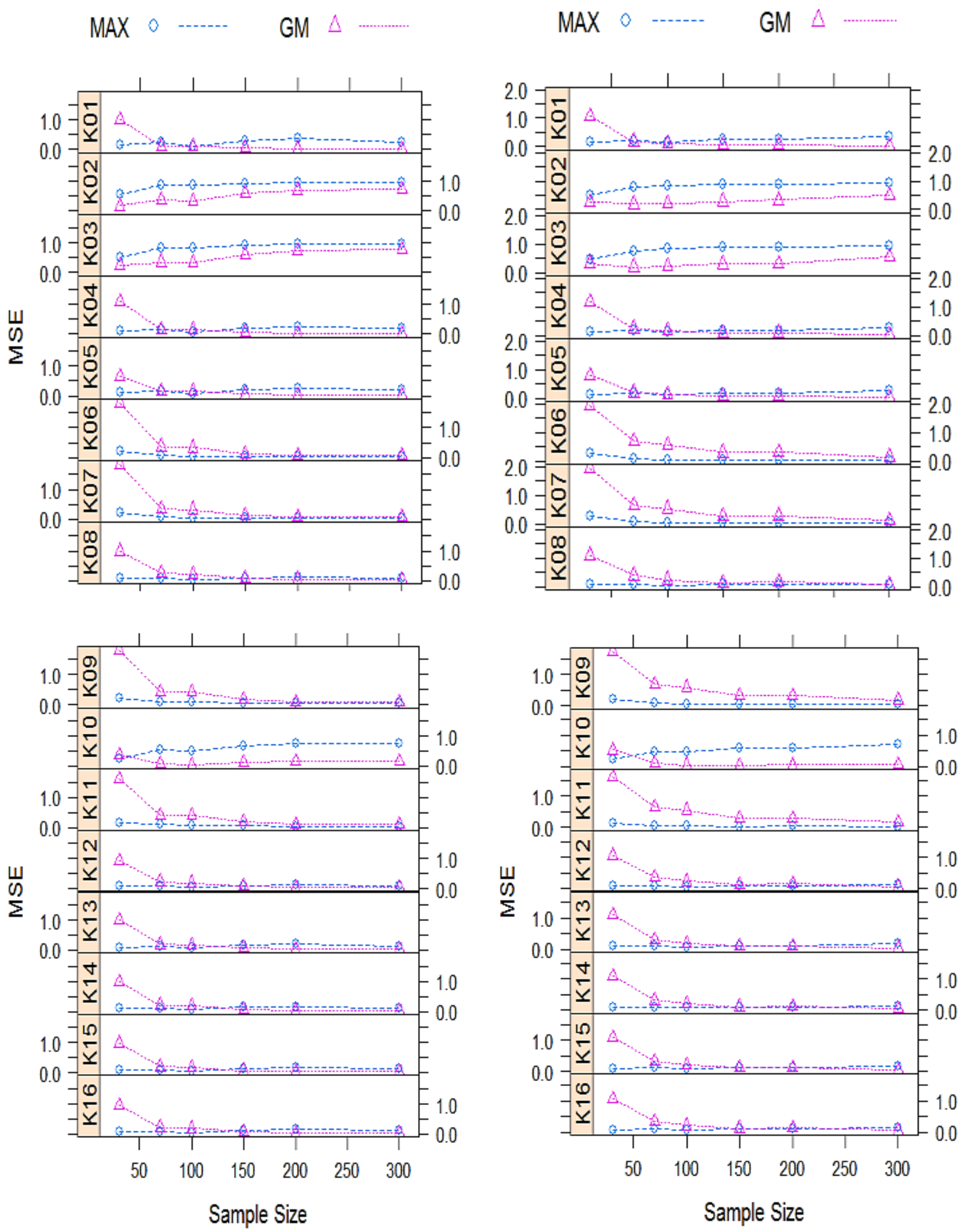


Figure 4.2(5): MSE at p=6 and ε_i~F (4, 20). 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

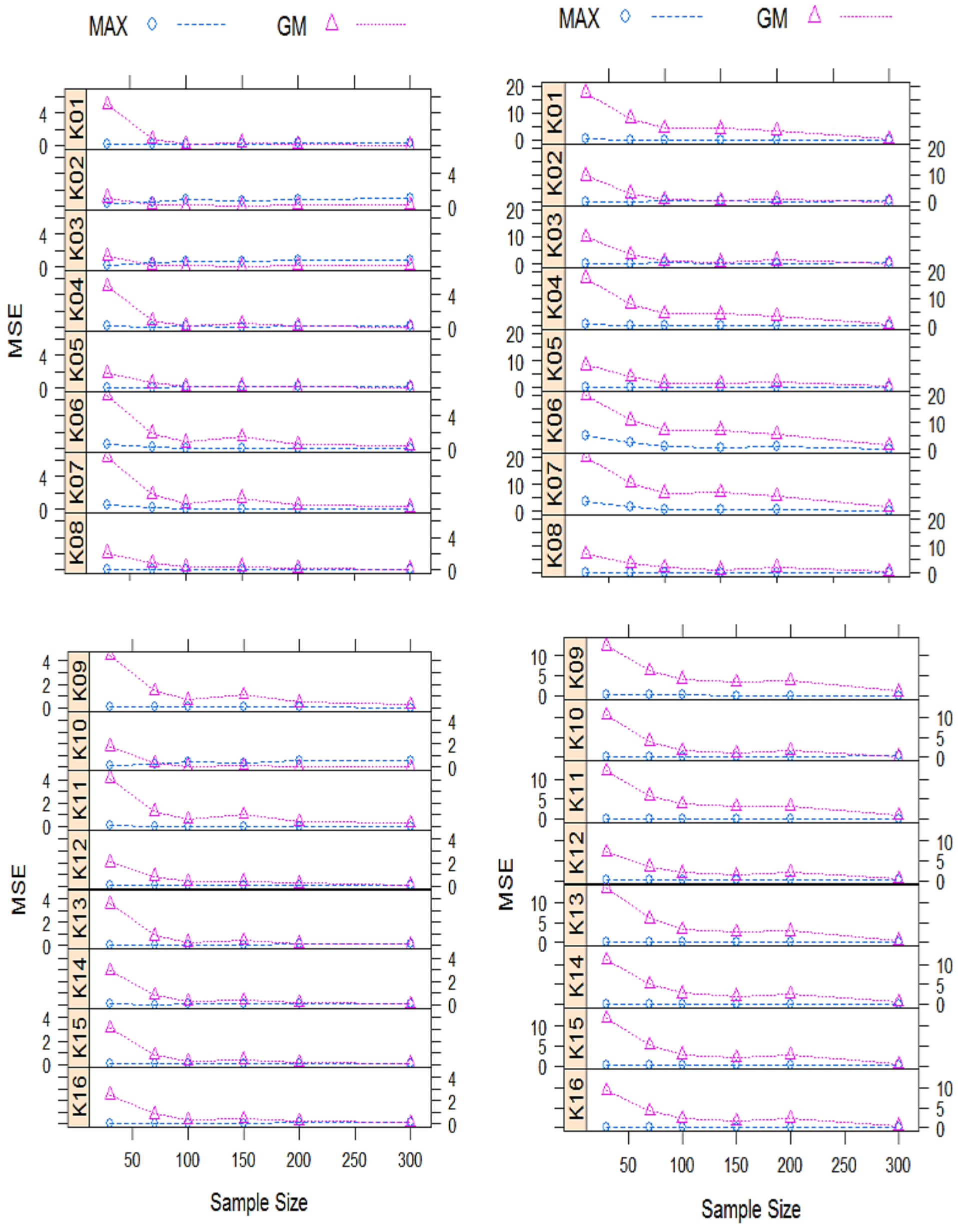


Figure 4.2(6): MSE at p=6 and ε_i~F (4, 20). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

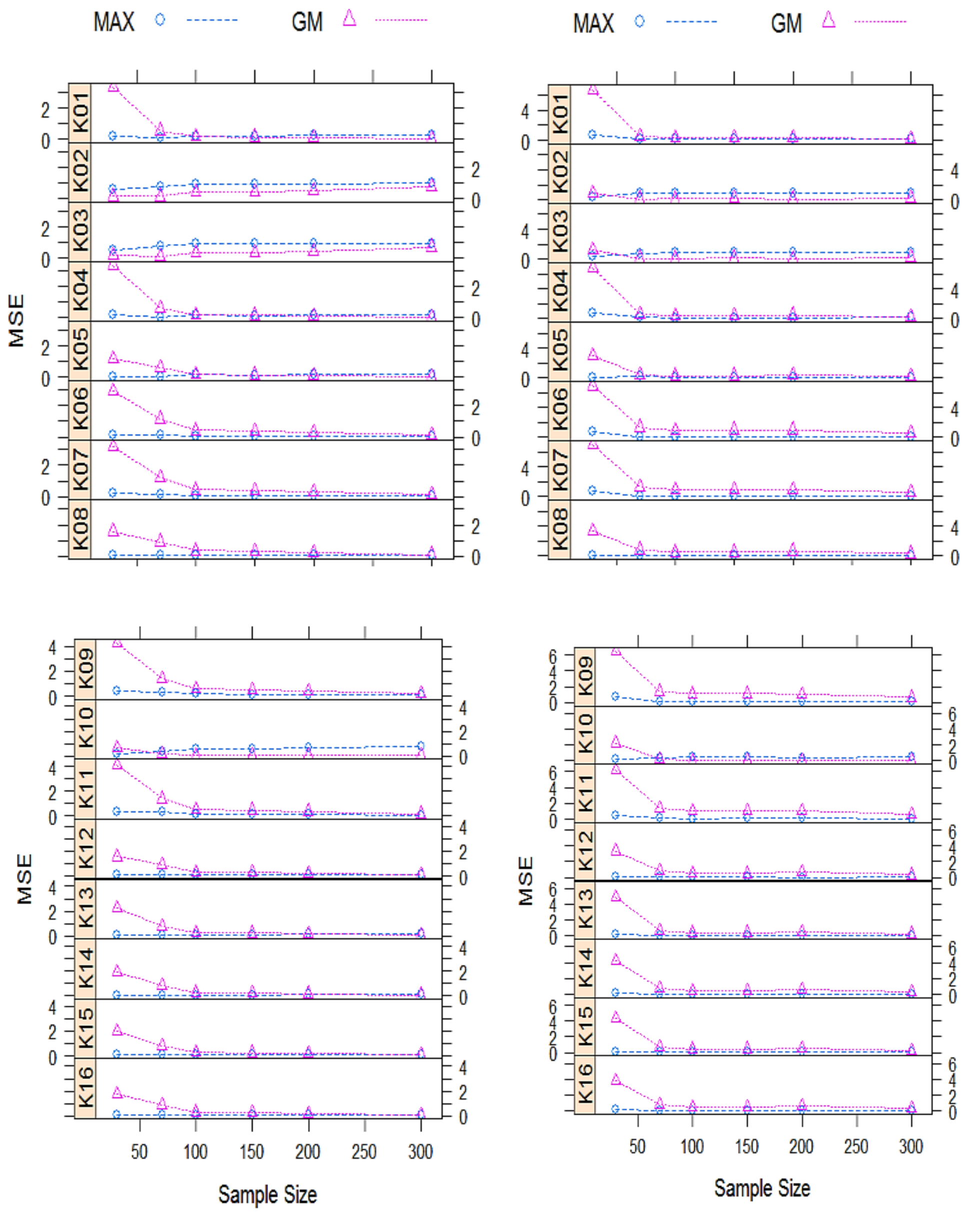


Figure 4.2(7): MSE at p=12 and ε_i~F (4, 20). 1st column is the case of r=0.80 and 2nd column is the case of r=0.90.

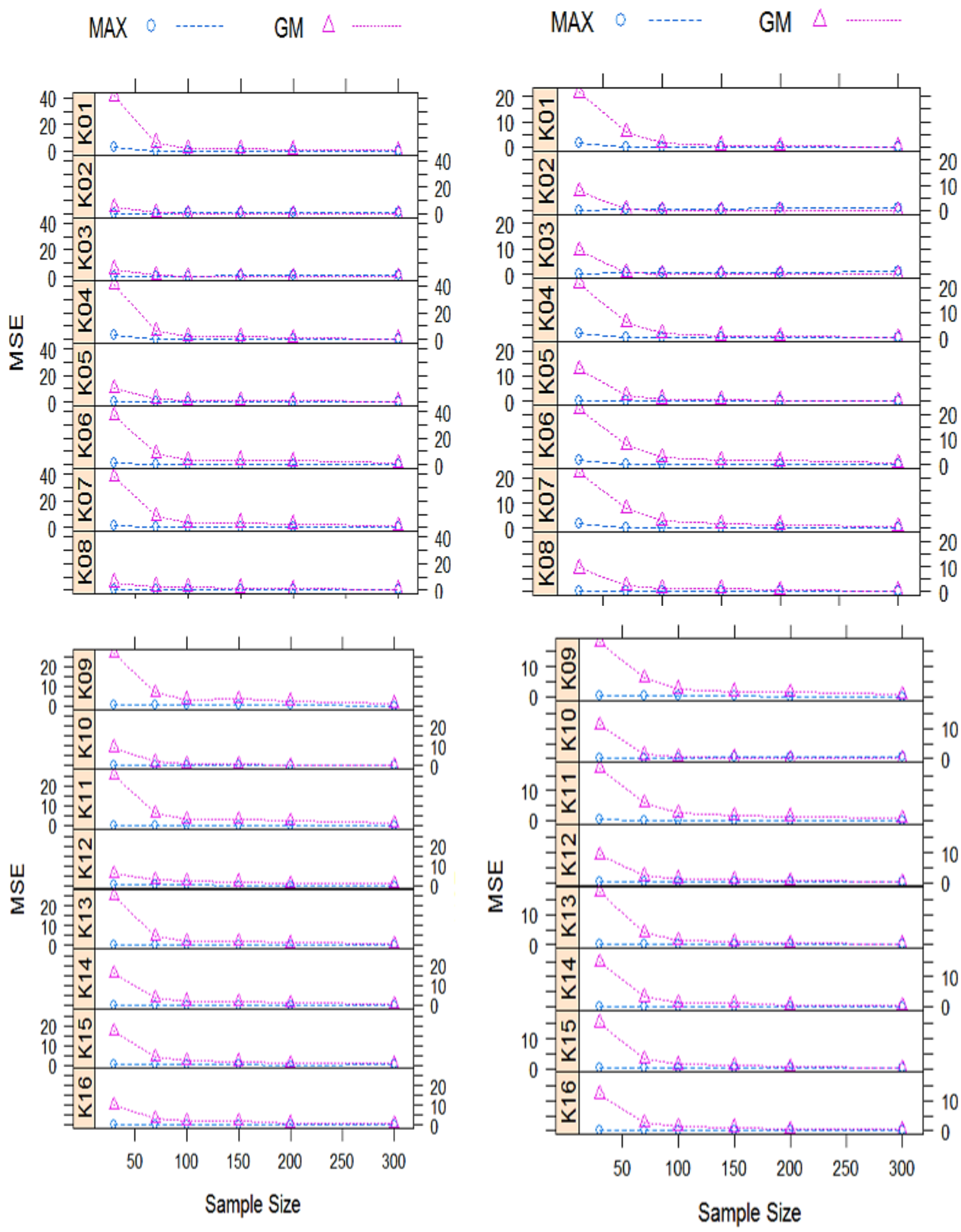


Figure 4.2(8): MSE at p=12 and ε_i~F (4, 20). 1st column is the case of r=0.95 and 2nd column is the case of r=0.99.

SUMMARY AND CONCLUSIONS

The results of simulation study comparing a set of proposed estimators K1G–K16G with K1M–K16M showed that when error terms follow normal distribution (see figures 4.1 (1-8)). It was observed that when σ^2= 0.1, at all levels of r and any numbers of p our modified method (estimators K1G–K16G) perform well. When p=2 MSE of estimators K1G, K4G, K5G, K6G, K7G and K13G, when p=4 MSE of estimators K1G, K4G, K5G, K6G, K7G, K11G and K13G and when p=6, 12 MSE of estimators K1G, K4G, K5G, K6G, K7G, K9G, K11G and K13G are minimum.

The MSE of (K1M–K16M) estimators are maximum when error terms follow F-distribution on the whole (see figures 4.2 (1–8)). When error terms follows F-distribution, p=2, at levels of r=0.80, 0.90, 0.95 our proposed estimators (K1G–K16G) perform well and at levels of r=0.99 our proposed estimators (K1G–K16G) not perform well, MSE of estimators K2M, K3M and K10M are maximum. When p=4, at r=0.80, 0.90 our proposed estimators (K1G–K16G) give good results, at r=0.95 proposed estimators (K1G–K16G) give miscellaneous pattern, at r=0.99 proposed estimators (K1G–K16G) not perform well and at high levels of collinearity (0.80, 0.90) MSE of estimators K2M, K3M and K10M are maximum. When p=6, at levels of r=0.80 our modified method perform well, at levels of r =0.90 our modified method gave mix pattern and at levels of r=0.95, 0.99 our modified method not perform well. For the highest number of predictors (p=12) and at all levels of ‘r’ estimators (K1G–K16G) do not perform well.

Thus, it can be inferred that set of our proposed estimators (K1G–K16G) clearly out performs K1M–K16M at all levels of collinearity and number of predictors when error terms follow normal distribution (σ^2=0.1). K1G–K16G were performing better than K1M–K16M at high collinearity levels (0.80, 0.90) and small to moderate number of predictors (p=2, 4) when error terms follow non-normal distribution F (4, 20)). However, miscellaneous pattern (i.e., some of the proposed estimators, K1G-K16G outperform the corresponding K1M–K16M) is observed for higher collinearity level (r=0.95).

REFERENCES

1. Al-Hassan, Y.M. (2008). A Monte Carlo evaluation of some ridge estimators. *Jordan Journal of Applied Science*, 10(2), 101-110.
2. Alkhamisi, M., Khalaf, G. and Shukur, G. (2006). Some modifications for choosing ridge parameters. *Communications in Statistics—Theory and Methods*, 35(11), 2005-2020.
3. Alkhamisi, M.A. and Shukur, G. (2007). A Monte Carlo study of recent ridge parameters. *Communications in Statistics—Simulation and Computation*, 36(3), 535-547.
4. Alkhamisi, M.A. and Shukur, G. (2008). Developing ridge parameters for SUR model. *Communications in Statistics—Theory and Methods*, 37(4), 544-564.
5. Chatterjee, S. and Hadi, A.S. (2006). *Analysis of collinear data: Regression analysis by example* (4th ed.). New Jersey: John Wiley.
6. Chatterjee, S. and Hadi, A.S. (2013). *Regression analysis by example*. New Jersey: John Wiley.

7. Clark, A. and Troskie, C. (2006). Ridge Regression–A Simulation Study. *Communications in Statistics—Simulation and Computation*, 35(3), 605-619.
8. Dempster, A.P., Schatzoff, M. and Wermuth, N. (1977). A simulation study of alternatives to ordinary least squares. *Journal of the American Statistical Association,* 72(357), 77-91.
9. Dorugade, A. and Kashid, D. (2010). Alternative method for choosing ridge parameter for regression. *Applied Mathematical Sciences*, 4(9), 447-456.
10. Epanechnikov, V.A. (1969). Non-parametric estimation of a multivariate probability density. *Theory of Probability & Its Applications,* 14(1), 153-158.
11. Greene, W.H. (2000). *Econometric analysis* (4th ed.). New Jersey: Macmillan.
12. Gujarati, D.N. (1978). *Basic Econometrics* (2nd ed.). New York: McGraw Hill.
13. Gujarati, D.N. (1995). *Basic econometrics* (3rd ed.). New York: McGraw Hill.
14. Gujarati, D.N. and Sangeetha. (2007). *Basic Econometrics*. New Delhi*:* Tata McGraw Hill Publishing Company Limited.
15. Hamed, R., Hefnawy, A.E. and Farag, A. (2013). Selection of the ridge parameter using mathematical programming. *Communications in Statistics-Simulation and Computation* 42(6), 1409-1432.
16. Hefnawy, A.E. and Farag, A. (2014). A Combined nonlinear programming model and Kibria method for choosing ridge parameter regression. *Communications in Statistics-Simulation and Computation* 43(6), 1442-1470.
17. Hoerl, A.E. (1975). Ridge regression: Some simulations. *Communications in Statistics-Theory and Methods* ,4(2), 105-123.
18. Hoerl, A.E. and Kennard, R.W. (1976). Ridge regression iterative estimation of the biasing parameter. *Communications in Statistics-Theory and Methods,* 5(1), 77-88.
19. Hosmer, D.W. and Lemeshow, S. (2004). *Applied logistic regression*. New Jersey: John Wiley.
20. Khalaf, G., Mansoon, K. and Shukur, G. (2013). Modified ridge regression estimators. *Communications in Statistics-Theory and Methods,* 42(8), 1476-1487.
21. Khalaf, G. and Shukur, G. (2005). Choosing ridge parameter for regression problems. *Communications in Statistics—Theory and Methods*, 34, 1177-1182.
22. Kibria, B.G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics-Simulation and Computation,* 32(2), 419-435.
23. Kibria, B.G. (2012). Performance of some logistic ridge regression estimators. *Computational Economics*, 40(4), 401-414.
24. Lawless, J.F. and Wang, P. (1976). A simulation study of ridge and other regression estimators. *Communications in Statistics-Theory and Methods*, 5(4), 122-143.
25. Locking, H, Mansson, G. and Shukur, G. (2013). Ridge estimators for probit regression with an application to labour martket data. *Bulletin of Economic Research,* 30(5), 501-513.
26. Månsson, K. and Shukur, G. (2011). On ridge parameters in logistic regression. *Communications in Statistics-Theory and Methods,* 40(18), 3366-3381.
27. Mansson, K. (2010). On some ridge regression estimators: A Monte Carlo simulation study under different error variances. *Journal of Statistics,* 17(1), 1-22.
28. McDonald, G.C. and Galarneau, D.I. (1975). A Monte Carlo evaluation of some ridge-type estimators. *Journal of the American Statistical Association*, 70(350), 407-416.

29. Montgomery, D.C. (2012). *Introduction to linear regression analysis*. New York: John Wiley.
30. Muniz, G. and Kibria, B.G. (2009). On some ridge regression estimators: An empirical comparisons. *Communications in Statistics—Simulation and Computation*, 38(3), 621-630.
31. Pasha, G. and Shah, M. (2004). Application of ridge regression to multicollinear data. *Journal of research (Science)*, 15(1), 97-106.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 145-156

GENDER INEQUALITY AND ECONOMIC GROWTH: A CROSS-COUNTRY ANALYSIS

Dilawar Kabir and **Syed Adil Hussain**
Department of Mathematics & Statistics, International Islamic University
Islamabad, Pakistan
Email: dilawarkabir@gmail.com; syed.adil@uog.edu.pk

ABSTRACT

This research attempts to analyze effect of gender inequality on economic growth a cross- country analysis for Pakistan, India and Bangladesh. This impact is investigated at the secondary panel data from 1990-2018 used in this study. The data for variables has been collected from World Development Indicators (World Bank). Estimation are done by Pooled OLS regression model, fixed effect and Random effect model on 3 cross-countries. Finally random effect model is selected by Hausman test and diagnostics are done by feasible generalized least square (FGLS), so this model having no auto-correlation and no heteroscasdicity present in random effect model. The result of this study leads us to demonstrate, that Gross fixed capital formation (GFCF) as a percentage of GDP has statistically significant and positive effect on GDP. Whereas Labor Force Participation Rate of Female (LFPF), Labor Force Participation Rate of Male (LFPM), Gender Parity Index (GPI), Openness of Trade (OP) is measured using trade as a percentage of GDP and Annual percentage of population growth (POP) have statistically insignificant and negative effect on GDP in three countries.

1. INTRODUCTION

Gender inequality refers to unequal treatment or perceptions of Individuals based on their gender. Gender-equality is not merely an issue of human rights, but an economic necessity. Worldwide, productivity growth and the pace of human development are slowing (ILO, 2017), thus women's full and effective participation in the workforce and decent work for all are indispensable to inclusive and sustainable economic growth. While women account for half of the total population, they remain an underutilized resource, constituting less than a third of the actual workforce (Lagarde, 2013). According to the a report of UN High-Level Panel on Women's Economic Empowerment, 700 million fewer women than men of working age were in paid employment in 2016, and even when women are paid, they tend to work in jobs with relatively low earnings, poor working conditions, and limited career prospects (UN, 2016). Implementing policies that remove labor market distortions and create a level playing field for all gives women the opportunity to develop their potential and to participate in economic life more visibly (IMF, 2013). Women are more likely to invest their resources in education and health of their children, building human capital to fuel future growth (Schultz, 2002). Helping women fully participate in the economy is not only growth-promoting, but it also diversifies the economies, reduces income inequality, and mitigates demographic shifts (Kochar et al., 2016). In many countries, constraints such as discriminatory laws, lack of legal protection, unfavorable social norms, and a lack of access to real and financial assets have held women back, which, in turn, have held back the economies.

Several literatures have been written with regard to the topic of interest in this paper. Most researcher have used time-series data on a specific country while others use cross-sectional data to examine the effect of gender inequality across several countries. This study, however, focuses on 3 countries. The main purpose is to investigate whether or not improvements in gender equality will cause economic growth and development in these regions.

A study by Ali (2015) investigates the relationship between gender inequality and economic growth in the case of Pakistan by using times series data from 1980-2009. The main purpose of his study is to prove women's important role in the development of the country. He specifically uses the difference in labor participation rates and education between males and females to do his research. Using a multiple regression model with Labor Force Participation Rate of Females (LFPF), Labor Force Participation Rate of Males (LFPM), Gender Parity Index (GPI), and Openness of Trade (OP) as the independent variables affecting the dependent variable GDP, Ali (2015) concludes that there is a positive relationship between gender equality and economic growth. Klasen and Lamana (2009) reach the same conclusion regarding gender inequality in education and the labor force using cross-country and panel data for the period from 1960-2000. The results are clear in the sense that gender inequality reduces economic growth in the countries of study. They concluded with the finding that gender gaps are the reason behind the growth differences across regions.

Nowadays in developed countries, new aspects of discrimination have emerged. In fact, Gender inequality in Education and Health has been superseded by gender inequality in Political Empowerment and Economic Participation (Modern discrimination). Lofstrom (2009) say Gender Inequality, Economic Growth, and Employment states a simple yet influential notion that gender inequality in the work opportunities leads to fair distribution of work and therefore higher productivity. However, traditional attitudes and cultural norms drive such differences between men and women, causing women to deviate from job opportunities due to the duties of society. This idea is valid to a great extent in these countries. Most differences and inequalities between genders are a result of cultural beliefs and religious norms that tend to cause the irrational utilization of a country’s human capital.

1.1 Organization of the Study

This paper is organized with a specific format; the study is divided in 5 chapters. In first chapter there is introduction to the study, in the second chapter there is a theoretical framework and literature review related to this study, in the third chapter is the research methodology and data collection and in this chapter there is brief explanation about the data, model and variables. In the next part is chapter four which included research methodology and design, there are some estimation are done in chapter four, the analysis and interpretation of the output after estimation also included in chapter four. And chapter five has conclusion and policy recommendation.

2. LITERATURE REVIEW

Many studies examine the effect of gender inequality on economic growth by various estimation techniques. Although the studies investigate different countries with different economic structure. The results are mostly skewed to having a negative correlation between gender inequality and economic growth.

A study by Shammari, et al. (2017) examines empirically the influence of gender inequality on the economic growth across Arab region. The data sample includes 19 countries for the period from 1990 to 2014. The model is estimated using Pooled OLS and Fixed Effect techniques with Gross Fixed Capital Formation (GFCF), Labor Force Participation Rate of Females (LFPF), Labor Force Participation Rate of Males (LFPM), Gender Parity Index (GPI), Trade refers to total trade as a percentage of GDP, population represents population growth and oil is for oil price as the independent variables affecting the dependent variable GDP.

Interestingly, findings show that there is no evidence of both gender inequality in education and gender inequality in labor force to hinder economic growth across Arab region. In fact, the findings of this study show that the main factors driving the economic growth across the Arab region are capital accumulation as well as population growth.

Other studies have been made to yield similar results, including Yumusak, et al. (2013) and their paper to see the influences of gender inequality in education on the Turkish economic growth. Their paper explains how women participation in the economy is a major resource that should be used more in Muslim countries in order to reach economic development targets. It focuses on the case of Turkey for the time frame 1968-2005 using a co-integration approach, yielding the empirical result that women's low level of education has a negative effect on economic growth and there is a positive long-run relationship between closing the gender gap in education and economic development. Similarly, Pervaiz, et al. (2011) analyze the impact of gender inequality on economic growth of Pakistan. An annual time series data for the period of 1972-2009 has been used in this study. They regressed growth rate of real gross domestic product (GDP) per capita on labour force growth, investment, trade openness and a composite index of gender inequality. They findings the results reveal that labour force growth, investment and trade openness have statistically significant and positive impact whereas gender inequality has a significant and negative effect on economic growth of Pakistan.

Khayria, et al. (2015) analyze the effect of gender inequality on economic growth using the GMM dynamic panel for the five countries of the Great Maghreb for the period 1985-2011. The results of this study show that, one hand, a positive and significant at the 5% of the population effect on economic growth and on the other hand a statistically significant and negative in 10% of the investment on economic growth effect. In contrast, gender inequality has a significantly negative effect of 5%.

3. RESEARCH METHODOLOGY

The objective of this paper is to measure the effect of gender inequality on economic growth a cross- country analysis for Pakistan, India and Bangladesh. This impact is investigated at the secondary time series data from 1990-2018 used in this study. The data for variables has been collected from World Development Indicators (World Bank).

This study have 6 independent variables and 1 dependent variable, in this model relationship between independent and dependent variable are obtained by using the Pooled OLS regression model, Fixed effect model, Random effect model as:

$$GDP = \beta_0 + \beta_1 GFCF + \beta_2 LFPF + \beta_3 LFPM + \beta_4 GPI + \beta_5 OP + \beta_6 POP + \varepsilon_i$$

where dependent Variable GDP measure the country economic growth, Gross fixed capital formation (GFCF) as a percentage of GDP, Labor Force Participation Rate of Female age 15+ (LFPF) and Labor Force Participation Rate of Male age 15+ (LFPM) are used as to refers to female and male labor force. Openness of Trade (OP) is measured using trade as a percentage of GDP as an indicator of total export and import as a share of gross domestic product. Annual percentage of population growth (POP) of each country is used across the time frame of the study. Gender Parity Index (GPI) for gross enrollment ratio in primary and secondary education is the ratio of girls to boys enrolled at primary and secondary levels in public and private schools. And ε_i is the error term in the above regression model. It means that if there is no change occur in dependent variable due to independent variables then there are some other factors which causes the change in dependent variable which we take as an error term in our regression model.

3.1 Analysis is used at STATA14 Software

Estimation for panel data using following Pooled OLS regression model, fixed effect model and random effect model.

Here we have three countries such as Pakistan, India and Bangladesh. And have seven variables such as GDP, GFCF, LFPF, LFPM, GPI, OP and POP. We want to check the relationship between GDP and other 6 explanatory variables such as GFCF, LFPF, LFPM, GPI, OP and POP. Our data is from 1990 to 2018 with number of observation is 87.

3.1.1 Pooled Regression

Here we pool all 87 observations together and run the regression model, neglecting the cross-section and time series nature of data.

The major problem with this model is that it does not distinguish between the various countries that we have. In other words, by combining three countries by pooling we deny the heterogeneity or individuality that may exist among 3 countries.

3.1.2 Fixed Effect Model

The fixed effect model allows for heterogeneity or individuality among three countries by allowing having its own intercept value.

The term fixed effect is due to the fact that although the intercept may differ across countries, but intercept does not vary over time, that is it is time invariant.

3.1.3 Random Effect Model

Here countries have a common mean value for the intercept. Now, apply Hausman Test to check which model (Fixed effect or Random effect) is suitable to accept. Hausman Test have this hypothesis that Null hypothesis: Random effect model appropriate and Alternative hypothesis: fixed effect model is appropriate.

If we get a statistically significant P-value, we shall use fixed effect model, otherwise random effect model.

3.1.4 Diagnostic

Diagnostic test on random effect and feasible generalized least square (FGLS) test are applied.

4. DATA DESCRIPTION

The data sample is composed of 3 Asia countries throughout the period from 1990-2018. Gross Domestic Product Growth, Gross Fixed Capital Formation, Labor Force Participation Rate of Females, Labor Force Participation Rate of Males, Gender Parity Index, Trade as a Percentage of GDP, and Annual Population Growth secondary data are obtained from the World Bank Database *World Development Indicat*ors (WDI).

Empirical Results and Analysis

Table 4.1. Displays the results of the estimated model using a Pooled Ordinary Least Square regression. The estimation gives a very insignificant relationship between the variables of this study, Labor Force Participation of Females (LFPF) and the Gender Parity Index (GPI) are presented with insignificant P-values of 0.609 and 0.953 respectively.

Table 4.1
Descriptive result of Ordinary Least Square Regression

Pooled OLS Model				**Number of obs.=**	**87**
Source	SS	df	MS	F(6,77)=	6.56
Model	95.276987	6	15.8794978	Prob>F=	0.0000
Residual	186.380676	77	2.42052825	R-squared	0.3383
Total	281.657663	83	3.39346581	Adj R-squared	0.2867

However, in Table 4.2 Gross Fixed Capital Formation (GFCF) resulted in significant P-values of 0.05. Meanwhile, Labor Force Participation of Males and female are shown insignificant with a P-Value of 0.771 and 0.609 respectively. This could be an indicator of the main factors that are actually driving economic growth in 3 countries.

Considering the fitness of the model as a whole, Table 4.1 shows R-squared of 0.3383 and Adjusted R-Squared of 0.2867.

Table 4.3
Ordinary Least Square Regression

GDP	Coeff.	Std. Err.	t	P>|t|	[95% conf. Interval]	
GFCF	.1533932	.0792307	1.94	0.05	-.0043753	.3111616
LFPF	-.0350475	.0682756	-0.51	0.609	-.1710015	.1009065
LFPM	-.031094	.1065123	-0.29	0.771	-.2431872	.1809992
GPI	-.1873419	3.158626	-0.06	0.953	-6.476969	6.102285
OP	-.0350113	.0461656	-0.76	0.451	-.1269388	.0569161
POP	-1.136606	1.501373	-0.76	0.451	-4.126221	1.853009
Constant	8.779172	8.449484	1.04	0.302	-8.045899	25.60424

The results are estimated further as shown in the following Fixed Effect Regression in Table 4.3. The previous observation is confirmed that one significant variable in this model are Gross Fixed Capital Formation (GFCF).

Table 4.3
Descriptive result of Fixed-effects Regression

Fixed-effects(within) Regression		**Number of obs.=**		**87**
Group variable: id		Number of groups=		3
R-Squared:				
	Within=	0.1769	F(6,75)=	2.69
	Between=	0.7095	Prob > F=	0.0204
	Overall=	0.2559	Corr (u_i, Xb)=	-0.8565

Table 4.4
Fixed-effects Regression

GDP	Coeff.	Std. Err.	T	P>|t|	[95% conf. Interval]	
GFCF	0.1959106	0.10922	1.79	0.077	-0.021667	0.413488
LFPF	-0.0270035	0.070612	-0.38	0.703	-0.16767	0.113663
LFPM	0.2357232	0.186357	1.26	0.21	-0.13552	0.606966
GPI	5.982354	4.738874	1.26	0.211	-3.457967	15.42268
OP	-0.0525124	0.05026	-1.04	0.299	-0.152635	0.04761
POP	-0.6712157	1.515764	-0.44	0.659	-3.690772	2.34834
Constant	-20.18951	18.72144	-1.08	0.284	-57.48454	17.10551
Sigma_u	1.4279511					
Sigma_e	1.5455435					
Rho	0.4605149	(fraction of variance due to u_i)				

The fixed effect model allows for heterogeneity or individuality among three countries by allowing having its own intercept value.

The term fixed effect is due to the fact that although the intercept may differ across countries, but intercept does not vary over time, which is it time invariant.

Here we assume that 3 countries are having different intercepts. In the data number of observation is 87 and number of groups is 3 meaning that there are 3 countries such as Pakistan, India and Bangladesh. F-statistic is 2.69and p-value is 0.0204 which is smaller than 5% which means that all the coefficients of this model are not equal to zero, because p-value is significant. That means all the coefficients of independent variables are not equal to zero. So the model is significant. It is easily see that all the independent variables have an insignificant result with the GDP dependent variable because its corresponding p-valve is greater than 5%.

Table 4.5
Descriptive result of Random-effects GLS Regression

Random-effects GLS regression	**Number of obs.=**		**87**	
Group variable: id	Number of groups=		3	
R-Squared:				
	Within=	0.1467	Wald chi2=	39.36
	Between=	0.9904	Prob > chi2=	0.0000
	Overall=	0.3383	Corr (u_i, X)=	0 (assumed)

Table 4.6
Random-effects GLS Regression

GDP	**Coeff.**	**Std. Err.**	**Z**	**P>|z|**	**[95% conf. Interval]**	
GFCF	.1533932	.0792307	1.94	0.053	-.0018961	.3086825
LFPF	-.0350475	.0682756	-0.51	0.608	-.1688651	.0987701
LFPM	-.031094	.1065123	-0.29	0.770	-.2398544	.1776664
GPI	-.1873419	3.158626	-0.06	0.953	-6.378134	6.003451
OP	-.0350113	.0461656	-0.76	0.448	-.1254942	.0554715
POP	-1.136606	1.501373	-0.76	0.449	-4.079242	1.80603
Constant	8.779172	8.449484	1.04	0.299	-7.781511	25.33986
Sigma_u	0					
Sigma_e	1.5455435					
rho	0	(fraction of variance due to u_i)				

Here 3 countries have a common mean value for the intercept. In run random-effect GLS regression Wald-statistic is 39.36 and the probability value is 0.000 which is less than 5% which is meaning that the all the coefficients are not equal to zero. And it is easily see that one significant variable in this model are Gross Fixed Capital Formation (GFCF) and all other independent variables have an insignificant results with the GDP dependent variable because its corresponding p-valve is greater than 5%.

Now, we shall to check which one model is appropriate fixed effect or random effect model? For this criteria apply the Hausman test which tells as which one model is appropriate one.

Hausman Test Hypothesis:

Null hypothesis: Random effect model appropriate

Alternative hypothesis: fixed effect model is appropriate

If we get a statistically significant P-value, then we shall use fixed effect model, otherwise random effect model is the appropriate.

Table 4.7
Hausman Fixed random

	(b)	**(B)**	**(b-B)**	**Sqrt(diag(V_b-V_B))**
	Fixed	**Random**	**Difference**	**S.E.**
GFCF	.1959106	.1533932	.0425174	.0751767
LFPF	-.0270035	-.0350475	.008044	.0180149
LFPM	.2357232	-.031094	.2668172	.1529188
GPI	5.982354	-.1873419	6.169696	3.532706
OP	-.0525124	-.0350113	-.017501	.0198689
POP	-.6712157	-1.136606	.4653904	.2083731

Chi2 = 3.03 Prob > chi2 = 0.805

In Hausman test the Chi-square value is 3.03 and probability value 0.805 is very high from 5% which is meaning that we do not reject null hypothesis. Which conclude that random effect model is the appropriate model.

Now, we check this model have a serial correlation or not? For this checking we run the Pesaran CD (cross- sectional dependence) test in the final model.

Null hypothesis: there is no serial correlation

Alternative hypothesis: there is serial correlation

In the test p-value is less than 5% so we reject the null hypothesis and conclude that accept the alternative hypothesis: there is serial correlation in the model.

Diagnostic checking

Diagnostic test for serial correlation and auto-correlation:

Null hypothesis: There is no auto-correlation and serial correlation in error terms

Alternative: There is auto-correlation and serial correlation in error terms

Table 4.8
Autocorrelation in panel data

Wooldridge test for autocorrelation in panel data	**F (1,2) =**	**14.229**
Ho: no first-order autocorrelation	Prob > F =	0.0636

Thus P-value > 5% which means that there is no auto-correlation.

Diagnostic test for heteroscedasticity:

Table 4.9
Heterosedasticity Test

Panel Group-wise Heterosedasticity Test			
Ho: panel Homoscedasticity **Ha: panel Group-wise heteroscedasticity**			
Langrange multiplier LM test=	898.1811	P-Value > Chi2	0.0000
Likelihood Ratio LR test =	22.946	P-Value > Chi2	0.0000
Wald test =	1.15E+04	P-Value > Chi2	0.0000

Here Wald test =1.15e+04 and P-value= 0.0000 which is less than 5%, so the model have group-wise heteroscedasticity in the model.

Table 4.10
Cross-sectional time-series FGLS regression

Coefficients: generalized least squares **Panels: homoscedastic** **Correlation: no autocorrelation**	**Number of obs. =** **Number of groups =** **Time periods =**	**87** **3** **28**
Estimated covariance =1 **Estimated autocorrelation =0** **Estimated coefficients =7**	Log likelihood = Wald chi2= Prob > chi2=	-152.6638 42.94 0.0000

Table 4.11
Cross-sectional time-series FGLS regression

GDP	Coef.	Std. Err.	z	P>|z|	[95% conf. Interval]	
GFCF	.1533932	.0758576	2.02	0.043	.004715	.3020713
LFPF	-.0350475	.0653689	-0.54	0.592	-.1631681	.0930731
LFPM	-.031094	.1019778	-0.30	0.760	-.2309668	.1687789
GPI	-.1873419	3.024154	-0.06	0.951	-6.114575	5.739891
OP	-.0350113	.0442002	-0.79	0.428	-.1216421	.0516194
POP	-1.136606	1.437455	-0.79	0.429	-3.953966	1.680754
Constant	8.779172	8.089765	1.09	0.278	-7.076475	24.63482

When we use FGLS (feasible generalized least square) in such case then this model is not suffering from heteroscedasticity which means homoscedasticity and second this model is not suffering from auto-correlation means there is no serial correlation in the model, so this is an appropriate model in the data set. Thus GFCF has statistically significant and positive effect on GDP whereas LFPE, LFPM, GPI, OP and POP have statistically insignificant and negative effect on GDP in three countries.

5. CONCLUSION AND POLICY RECOMMENDATION

This study provides evidence of the impact of gender inequality in education and the labor market on economic growth in the cross-country analysis. This impact is investigated at the secondary time series data from 1990-2018 used in this study. Using the Pooled OLS regression model, fixed effect model and Random effect model for panel data and Husman test technique are used for model selection and finally, make its diagnostic on 3 countries is used to discover that gender inequality has statistically insignificant and negative effect on economic growth.

In the Pooled regression model it easily can see that Gross Fixed Capital Formation (GFCF) resulted in significant P-values of 0.05. Meanwhile, Labor Force Participation of Males and female are shown insignificant with a P-Value of 0.771 and 0.609 respectively. This could be an indicator of the main factors that are actually driving economic growth in 3 countries.

In the Fixed effect model it assume that 3 countries have different intercepts and is also see that all the independent variables have an insignificant results with the GDP dependent variable because its corresponding p-valve is greater than 5%. In the data number of observation is 84 and number of groups is 3 meaning that there are 3 countries such as Pakistan, India and Bangladesh have different intercepts, so F-statistic is 2.69 and p-value is 0.0204 which is smaller than 5% which means that all the coefficients of this model are not equal to zero, because p-value is significant.

Finally Random effect GLS model that is the appropriate model which is selected according to Hausman test. The probability value is 0.000 (significant) and Wald chi2=39.36 so, happy about this model because probability value is significant. In this model three countries have a common mean value for the intercept. Wald-statistic is 39.36 and the probability value is 0.000 which is less than 5% which is meaning that the all the coefficients are not equal to zero. And it is easily see that one significant variable in this model are Gross Fixed Capital Formation (GFCF) and all other independent variables have an insignificant results with the GDP dependent variable because its corresponding p-valve is greater than 5%.

After Diagnostic checking by using FGLS (feasible generalized least square) test we conclude that this model is not suffering from heteroscedasticity which means homoscedasticity and second this model is not suffering from auto-correlation means there is no serial correlation in the model, so this is an appropriate model in the data set. Thus GFCF has statistically significant and positive effect on GDP whereas LFPE, LFPM, GPI, OP and POP have statistically insignificant and negative effect on GDP in three countries.

REFERENCE

1. Ali, M. (2015). Effect of Gender Inequality on Economic Growth. Case of Pakistan. *Journal of Economics and Sustainable Development,* 6(9), 125-133.
2. Ahang, M. (2014). The impact of Gender inequality on Economic Growth in Developed Countries. *Advances in Environmental Biology*, 8(17), 508-513.
3. Berik, G. (2000). Mature export -led growth and gender wage inequality in Taiwan. *Feminist Economics*, 6(3), 1-26.
4. Berik, G., Rodgers, Y.V.D.M and Zveglich, J. (2004). International Trade and Gender Wage Discrimination: Evidence from East Asia. *Review of Development Economics,* 8(2), 237-254.
5. Busse, M. and Spielmann, C. (2006). Gender Inequality and Trade", *Review of International Economics*, 14(3), 362-379.
6. Khayria, K. and Feki, R. (2015). Gender Inequality and Economic Development. *Business and Economics Journal*, 6(4), 1-3.
7. Klasen, S. (2002). Low schooling for girls, slower growth for all? Cross-country evidence on the effect of gender inequality in education on economic development. *The World Bank Economic Review*, 16(3), 345-373.
8. Kochhar, K., Jain-Chandra, S. and Newiak, M. (2016). Women, Work, and Economic Growth: Leveling the Playing Field. *International Monetary Fund*.
9. Lamanna, F. and Klasen, S. (2009) "the impact of gender inequality in education and employment on economic growth: new evidence for a panel of countries. *Feminist Economics*, 15(3), 91-132.
10. Natali, L and Kabeer, N. (2013). *Gender Equality and Economic Growth: Is there a Win-Win*? Institute of Development Studies, ISSN: 2040-0209, ISBN: 978-1-78118-108-9.
11. Shammari, N and Rakhis, M. (2017). Impact of Gender Inequality on Economic Growth in the Arab Region. *Research in Applied Economics*, 9(2), 18-31.

12. Seguino, S. (2000). Gender Inequality and Economic Growth: A Cross-Country Analysis. *World Development*, 28(7), 1211-1230.
13. Seguino, S. (2000). Accounting for gender in Asian economic growth. *Feminist Economics*, 6(3), 27-58.
14. Teignier, M. and Cuberes, D. (2014). Gender inequality and economic growth: A critical review. *Journal of International Development*, 26, 260-276.
15. Weichselbaumer, D. and Ebmer, R. (2005). A meta-analysis of the international gender wage gap. Journal of Economic Surveys, 19(3), 479-510.
16. Yumusak, I.G., Bilen, M. and Ates, H. (2013). The Impacts of Gender Inequality in Education on Economic Growth in Turkey. *Procedia - Social and Behavioral Sciences*, 103, 1093-1103.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 157-168

INVERSE EXPONENTIAL LOMAX DISTRIBUTION: PROPERTIES AND APPLICATION

Ayesha Riaz, Ahmad Saeed Akhter and **Maha Shabbir**
College of Statistical and Actuarial Sciences
University of the Punjab, Lahore, Pakistan
Email: ayeshariaz21@yahoo.com
asakhter@yahoo.com
maha.sh010@gmail.com

ABSTRACT

A continuous distribution known as (IELD) Inverse Exponential-Lomax distribution is introduced which is an extension of Exponential Lomax distribution. New proposed model has three unknown parameters which includes one shape and the two scale parameters. Different statistical measures of the new uni-model distribution are derived which involves failure rate, moments, mean deviation (MD) and many other properties. The graphical presentation of the pdf of the IELD and its cumulative distribution function is represented. In this article, the impact of each parameter on the properties of the statistical measure of distribution are explored. Moreover, order statistics expression with the L-moments is also provided for the new model. The Maximum Likelihood method of estimation is used to estimate the values of the parameters of IELD. A real life data set is used in this paper to highlight the potentiality of the Inverse Exponential-Lomax distribution. As for this data set, new model performs better than many other distributions used for comparison on the basis of different model selection criterions.

KEYWORDS

Inverse Exponential, Lomax, Hazard, Maximum likelihood and mean deviation

1. INTRODUCTION

Lomax, Pareto, Rayleigh, Exponential, Weibull and many other distributions are helpful in modeling. K.S Lomax introduced the Lomax distribution in 1954 to resolve the business failure issue (Lomax, 1954). Lomax distribution is mostly used in business failure, economics, wealth inequalities, medical sciences, life time, reliability modeling and many others. Moreover, it is a generalization of the Pareto distribution that starts from zero and instead of using gamma, Weibull and exponential, Lomax is recommended. On the other hand, Keller and Kamath presented the extension of exponential model known as the inverse exponential model in 1982 to access CNC machine reliability. (IED) Inverse exponential distribution has the inverse bathtub hazard rate (Keller, et. al., 1982).

2. METHODOLOGY

Various methods are available in literature to develop the mixture of distributions. The IELD is formed by merging the Inverse-Exponential distribution with the Lomax distribution by utilizing the same method suggested by Al-Kadim and Boshi (2013). Also using the approach of Al-Kadim and Boshi (2013), Shabbir et al. (2017) proposed the (NWLD) New Weibull Lomax distribution. On similar approach Shabbir et al. (2018) introduced the model of Rayleigh Lomax distribution. El-Bassiouny et al. (2015) suggested (EL) exponential-Lomax distribution by using the approach the new distribution function of the distribution can be generalized using equation 1.

$$F(x) = \int_{0}^{\frac{1}{S(x)}} f_1(x)dx; \tag{1}$$

where S(x) indicates the Lomax distribution's survival function and $f_1(x)$ indicates the pdf of the (IE) Inverse Exponential distribution. Lomax distribution's survival function is

$$S(x) = (\frac{\beta}{x+\beta})^{\gamma}; \tag{2}$$

whereas the pdf of the (IE)

$$f_1(x) = \frac{\alpha}{x^2} e^{-(\frac{\alpha}{x})} ; \alpha, x > 0. \tag{3}$$

By substituting the values of the S(x) and $f_1(x)$ in equation 1, the (CDF) of the newly proposed (IELD) is

$$F(x) = e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)} ; x > -\beta. \tag{4}$$

The probability density function of the IELD is obtained by differentiation of the F(x)

$$f(x) = \alpha\beta\gamma\left(\frac{\beta}{x+\beta}\right)^{\gamma-1} (\frac{1}{x+\beta})^2 \, e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}, ; x > -\beta\,, \alpha, \beta, \gamma > 0. \tag{5}$$

where α, β are scale parameters and γ is the shape parameter in model. The various shapes of the pdf of the IELD at different values of the parameters. Figure 1 to Figure 3 shows the pdf plots for the Inverse Exponential Lomax distribution for different set of values. It is indicated that the IELD holds positively skewed shape.

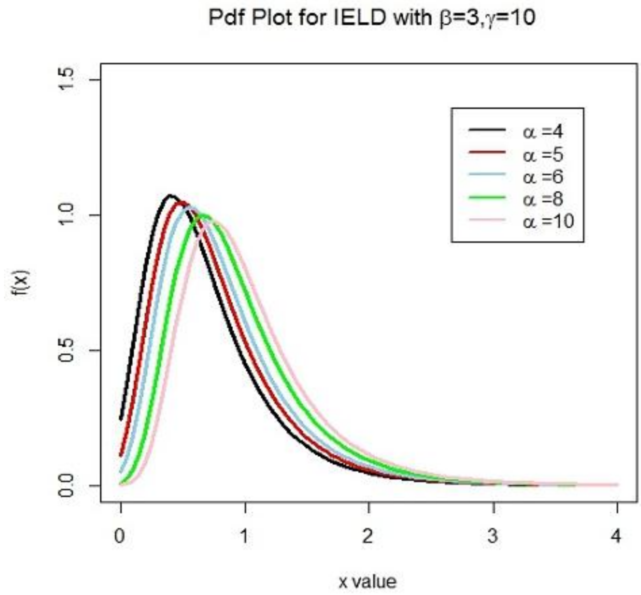


Figure 1: IELD pdf for different choices of α

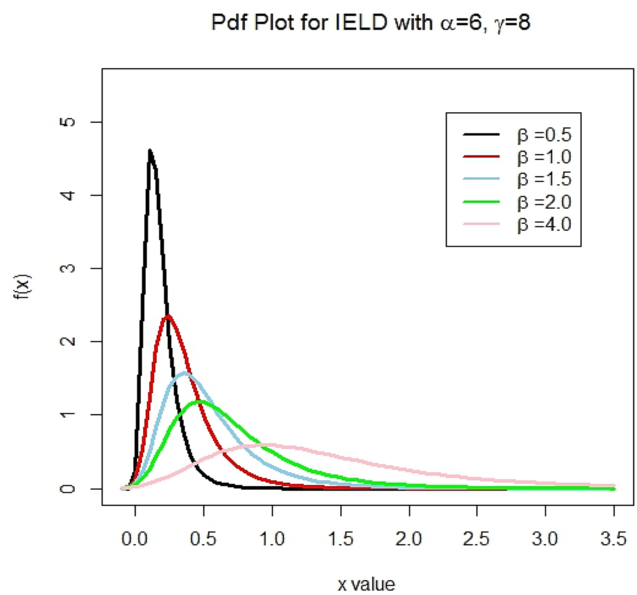


Figure 2: IELD pdf for Different Choices of β

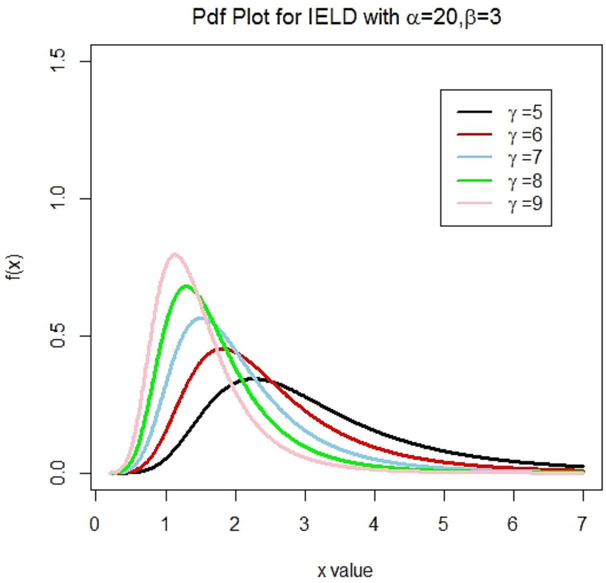


Figure 3: IELD pdf for Different Choices of γ

3. STATISTICAL PROPERTIES

3.1 Moments

The R^th moment for the IELD are

$$E(X)^r = \sum_{j=0}^{r} \binom{r}{j} \left(\beta \alpha^{\frac{1}{\gamma}}\right)^j (-\beta)^{r-j}\, \Gamma\left(1-\frac{j}{\gamma}\right), \tag{6}$$

for $r = 1$,

$$\mu_1{}' = \beta[(\alpha)^{\frac{1}{\gamma}} \Gamma\left(1-\frac{1}{\gamma}\right) - 1] \tag{7}$$

for $r = 2$,

$$\mu_2{}' = (\beta)^2 \left[1 - 2(\alpha)^{\frac{1}{\gamma}} \Gamma\left(1-\frac{1}{\gamma}\right) + (\alpha)^{\frac{2}{\gamma}} \Gamma\left(1-\frac{2}{\gamma}\right)\right], \tag{8}$$

if $r = 3$,

$$\mu_3{}' = (-\beta)^3 [1 - 3(\alpha)^{\frac{1}{\gamma}} \Gamma\left(1-\frac{1}{\gamma}\right) + 3\alpha^{\frac{2}{\gamma}} \Gamma\left(1-\frac{2}{\gamma}\right) - \alpha^{\frac{3}{\gamma}} \Gamma\left(1-\frac{3}{\gamma}\right), \tag{9}$$

$if\ r = 4$,

$$\begin{aligned} \mu_4{}' = (-\beta)^4 [1 + 4(\alpha)^{\frac{1}{\gamma}} \Gamma\left(1-\frac{1}{\gamma}\right) - 6\left((\alpha)^{\frac{1}{\gamma}}\right)^2 \Gamma\left(1-\frac{2}{\gamma}\right) \\ + 4\left((\alpha)^{\frac{1}{\gamma}}\right)^3 \Gamma\left(1-\frac{3}{\gamma}\right) - \left((\alpha)^{\frac{1}{\gamma}}\right)^4 \Gamma\left(1\frac{4}{\gamma}\right). \end{aligned} \tag{10}$$

3.1 Moment Generating Function

Moment generating function of IELD is as follows

$$M_0(\mathrm{t}) = \sum_{i=0}^{\infty} \frac{t^{\iota}}{\iota!}\left[\sum_{j=0}^{t} \binom{\iota}{j} \left(\beta\alpha^{\frac{1}{\gamma}}\right)^{j} (-\beta)^{i-j}\, \Gamma\left(1-\frac{j}{\gamma}\right)\right] \tag{11}$$

3.2 Mean deviation (MD)

MD from mean as well as median of IELD respectively is given as follows

$$\mathrm{D}(\mu) = (\mu+\beta)2\mathrm{F}(\mu) - 2\beta\alpha^{\frac{1}{\gamma}}\Gamma\left(\left(1-\frac{1}{\gamma}\right), \alpha\left(\frac{\beta}{\mu+\beta}\right)^{\gamma}\right) \tag{12}$$

$$\mathrm{D}(\mathrm{med}) = \mu + 2\beta e^{-\left(\alpha\left(\frac{\beta}{m+\beta}\right)^{\gamma}\right)} - 2\beta\alpha^{\frac{1}{\gamma}}\Gamma\left(1-\frac{1}{\gamma}\right), \alpha\left(\frac{\beta}{m+\beta}\right)^{\gamma} \tag{13}$$

3.3 Quantile Function (QF)

IELD’s Quantile function is given as

$$F(x) = e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)} \tag{14}$$

3.4 Random Number Generator

We consider a uniform variable U having the unit interval 0,1 that is U~u (0,1). Thus, the inverse transformation mean method, the value of the random variable x for IELD is given by

$$x = \beta\left[\frac{1}{\left[\frac{1}{\alpha}\ln\left(\frac{1}{u}\right)\right]^{\frac{1}{\gamma}}} - 1\right] \tag{15}$$

4. RELIABILITY

4.1 Survival Function

The survival function for IELD is

$$S\,(x) = 1 - e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}. \tag{16}$$

4.2 Failure Rate Function

The Failure rate function for IELD is

$$h(x) = \frac{\alpha\beta\gamma\left(\frac{\beta}{x+\beta}\right)^{\gamma-1}\left(\frac{1}{x+\beta}\right)^{2} e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}}{1 - e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}}. \tag{17}$$

The shapes of the hazard rate function at various values of parameters are shown in Figure 4 to Figure 6. The Hazard rate of IELD is increasing and decreasing for different set of parameter values.

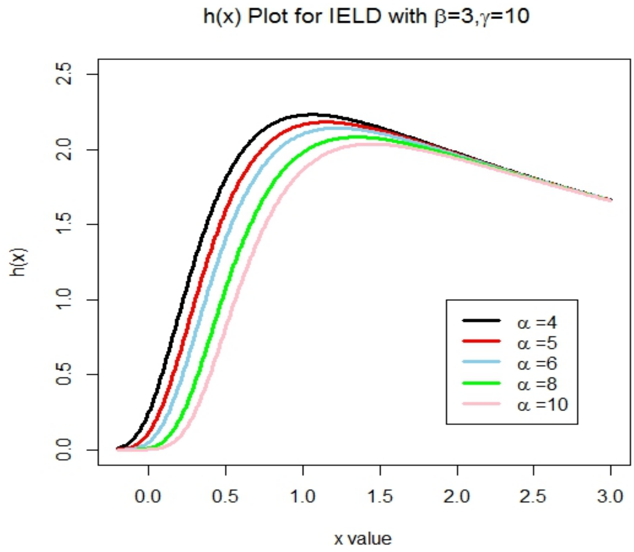


Figure 4: Hazard Rate Plot for different choices of α

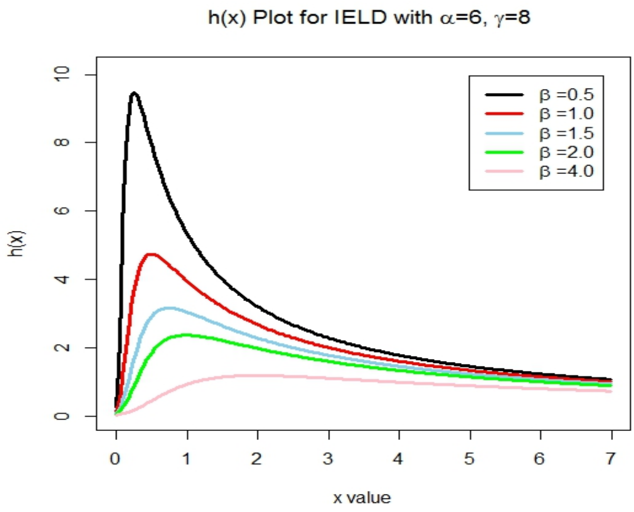


Figure 5: Hazard Rate Plot for different choices of β

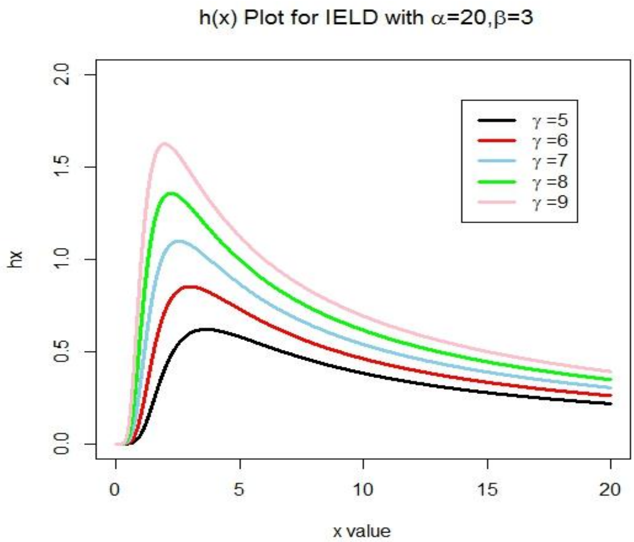


Figure 6: Hazard Rate Plot for different choices of γ

5. ORDER STATISTICS

The probability density function of the r^th order statistics is derived by

$$gX_{(r)}(x) = \frac{n!}{(r-1)!(n-r)!} g_x(X)[G_X(X)]^{r-1}[1-G(X)]^{n-r}\,; \quad r = 1,2,3,4,\ldots\ldots n; \quad -\infty < X_{(1)}\,X_{(2)}X_{(3)}X_{(4)\ldots\ldots\ldots}X_{(n)} < \infty. \tag{18}$$

The r^th order statistics for IELD is

$$gX_{(r)}(x) = \frac{n!}{(r-1)!(n-r)!}\left[\alpha\beta\gamma\left(\frac{\beta}{x+\beta}\right)^{\gamma-1}\left(\frac{1}{x+\beta}\right)^{2} e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}\right] \left[e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}\right]^{r-1}\left[1-e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}\right]^{n-r}. \tag{19}$$

The smallest order statistics for IELD is

The largest order statistics for IELD is

$$gX_{(n)}{}^{(x)} = n\left[\alpha\beta\gamma\left(\frac{\beta}{x+\beta}\right)^{\gamma-1}\left(\frac{1}{x+\beta}\right)^{2} e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}\right]\left[e^{-\left(\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right)}\right]^{n-1}\,; \quad 0 < X_{(n)} < \infty. \tag{21}$$

6. MAXIMUM LIKELIHOOD ESTIMATION

Maximum likelihood estimation method is used to estimate the unknown parameters of the IELD. The log likelihood function of the IELD based on n samples is

$$lnL = \mathrm{n}\ln\alpha + \mathrm{n}\ln\gamma + \mathrm{n}\gamma\ln\beta - (\gamma+1)\sum_{i=1}^{n}\ln(x+\beta) - \alpha\sum_{i=1}^{n}\left(\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right) \tag{22}$$

Differentiating the log likelihood w.r.t parameters

$$\frac{\partial lnL}{\partial\alpha} = \frac{n}{\alpha} - \sum_{i=1}^{n}\left(\frac{\beta}{x+\beta}\right)^{\gamma}, \tag{23}$$

$$\frac{\partial lnL}{\partial\beta} = \frac{n\gamma}{\beta} - (\gamma+1)\sum_{i=1}^{n}\left(\frac{1}{x+\beta}\right) - \alpha\sum_{i=1}^{n}\frac{\gamma[\beta^{\gamma-1}-\beta^{\gamma}]}{(x+\beta)^{\gamma}} \tag{24}$$

$$\frac{\partial lnL}{\partial\gamma} = \frac{n}{\gamma} + \sum_{i=1}^{n}\ln\left(\frac{\beta}{x+\beta}\right)\left[1-\alpha\left(\frac{\beta}{x+\beta}\right)^{\gamma}\right]. \tag{25}$$

The maximum likelihood (ML) estimates can be obtained by solving the equations through the non-linear system of equations. The most common method is used to solve the equation known as Newton Raphson.

7. REAL LIFE APPLICATION

Real life data set of 128 bladder cancer patients is used to support the IELD. The data set have 128 observations which was initially used by Yousof et al. (2015), Nofal et al. (2015) and many others to highlight the feasibility of the probability models.

0.08, 0.20, 0.40, 0.50, 0.51, 0.81, 0.90, 1.05, 1.19, 1.26, 1.35, 1.40, 1.46, 1.76, 2.02, 2.02,2.07, 2.09,2.23, 2.26, 2.46, 2.54, 2.62, 2.64, 2.69, 2.69, 2.75, 2.83, 2.87, 3.02, 3.25, 3.31, 3.36, 3.36, 3.48,3.52,3.57, 3.64, 3.70, 3.82, 3.88, 4.18, 4.23, 4.26, 4.33, 4.34, 4.40, 4.50, 4.51, 4.87, 4.98, 5.06, 5.09, 5.17, 5.32,5.32,5.34, 5.41, 5.41, 5.49, 5.62, 5.71, 5.85, 6.25, 6.54, 6.76, 6.93, 6.94, 6.97, 7.09, 7.26, 7.28, 7.32, 7.39,7.59,7.62,7.63, 7.66, 7.87, 7.93, 8.26, 8.37, 8.53, 8.65, 8.66, 9.02, 9.22, 9.47, 9.74, 10.06, 10.34, 10.66, 10.75,11.25,11.64, 11.79, 11.98, 12.02, 12.03, 12.07, 12.63, 13.11, 13.29, 13.80, 14.24, 14.76, 14.77, 14.83, 15.96, 16.62,17.12, 17.14, 17.36, 18.10, 19.13, 20.28, 21.73, 22.69, 23.63, 25.74, 25.82, 26.31, 32.15, 34.26, 36.66, 43.01,46.12, 79.05.

We compare the IELD with many other distributions including

Exponential-Lomax Distribution (ELD)

$$f_{ELD}(x) = \frac{\alpha\lambda}{\beta}\left(\frac{\beta}{x+\beta}\right)^{-\alpha+1} e^{-\lambda\left(\frac{\beta}{x+\beta}\right)^{-\alpha}}, \qquad \alpha, \beta, \lambda > 0, \tag{26}$$

Exponential Pareto Distribution (EPD)

$$f_{EPD}(x) = \frac{\lambda\theta}{p}\left(\frac{x}{p}\right)^{\theta-1} e^{-\lambda\left(\frac{x}{p}\right)^{\theta}}, \qquad \mathrm{P}, \lambda, \theta > 0, \tag{27}$$

Lomax Distribution (LD)

$$f_{LD}(x) = \frac{\alpha\beta^{\alpha}}{(\beta+x)^{\alpha+1}}, \qquad \alpha, \beta > 0, \tag{28}$$

Inverse-exponential Distribution (IED)

$$f_{IED}(x) = \frac{\alpha}{x^2} e^{-\frac{\alpha}{x}}, \qquad \alpha > 0. \tag{29}$$

The ML estimates of these distributions are compared by using four criteria's to measure the relative efficiency or quality of these statistical models. The criteria's being used are

$$-LL = -Log - likelihood \tag{30}$$

$$AIC = -2LL + 2q \tag{31}$$

$$BIC = -2LL + qlog(n) \tag{32}$$

The goodness of fit of model is evaluated using the two statistics

$$\text{W(Cramer –von Mises)}=\left(\frac{1}{2n}+1\right)\left[jn\ = 1\left(zi-\frac{2j-1}{2n}\right)^2+\frac{1}{12n}\right] \tag{33}$$

$$A(\text{Anderson} - \text{Darling}) = \left(\frac{9}{4n^2}+\frac{3}{4n}+1\right) \left[n+\frac{1}{n}j^n\ = 1(2j-1)\log\left(z_i\left(1-z_{n-j+1}\right)\right)\right] \tag{34}$$

q= Total Parameters, n= Observations,
$z_i = F(y_j)$ where y_j indicates the values used for order observation

Table 1
ML Estimates and Various Model Selection Criteria's

Models	*MLE*	*-LL*	*AIC*	*BIC*	*W*	*A*
ELD	α=1.04 β=0.07 λ=0.01	-415.02	836.04	844.59	0.13	0.81
EPD	λ=0.11 p=1.15 θ=1.05	-414.08	834.17	842.72	0.13	0.78
IED	α=2.48	-460.38	922.76	925.61	1.11	6.60
LD	λ=0.01 θ=10.37	-413.87	831.75	837.46	0.07	0.43
IELD	**α=4.76** **β=4.61** **γ=2.28**	**-411.31**	**828.63**	**837.18**	**0.02**	**0.11**

Table 1 indicates that the proposed IELD is the best fitted model for this data set as the values of the goodness of fit measures is smaller than the other distributions used for the comparison. The empirical CDF graph also suggests that the IELD is the best fitted model as compared to the based distributions for this data set. Figure 7 shows the empirical cumulative distribution plot for the data set indicating that IELD is a good fit on the data.

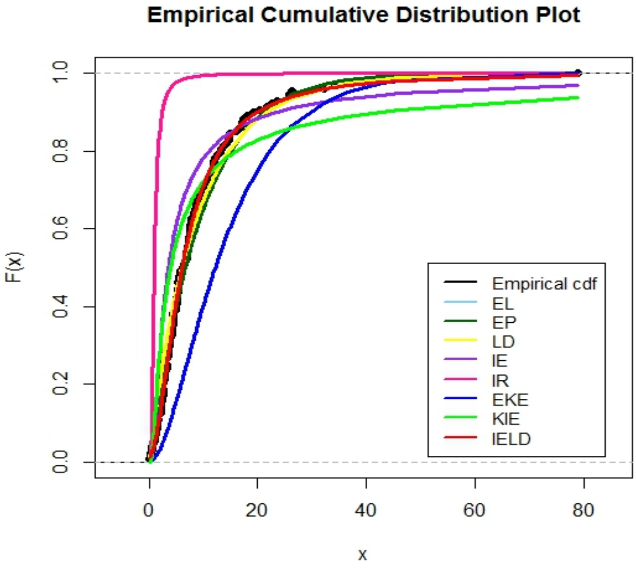


Figure 7: The Empirical CDF Plot for Data

8. COMMENTS AND CONCLUSION

The new probability model has been proposed known as inverse exponential Lomax distribution (IELD). It was observed that the new model is much more flexible then the parent distributions. The IELD holds a positively skewed shape in the model. Various structural properties have been derived in this paper. Also, the reliability analysis of new model is also represented. Moreover, various expression of order statistics has also proposed. Most importantly, the real-life application has been provided which suggests that the newly proposed distribution is much better than the other distributions based on the various model selection criterions.

REFERENCES

1. Al-Kadim, K.A. and Boshi, M.A. (2013). Exponential Pareto Distribution. *Mathematical Theory and Modeling*, 3(1), 135-146.
2. El-Bassiouny, A.H., Abdo, N.F. and Shahen, H.S. (2015). Exponential Lomax Distribution. *International Journal of Computer Applications*, 121(13), 24-29.
3. Keller, A.Z., Kamath, A.R.R. and Perera, U. D. (1982). Reliability analysis of CNC machine tools. *Reliability Engineering*, 3(6), 449-473.
4. Lomax, K.S. (1954). Business failures: Another example of the analysis of failure data. *Journal of the American Statistical Association*, 49(268), 847-852.
5. Nofal, Z.M., Afify, A.Z., Yousof, H.M. and Cordeiro, G. M. (2016). The generalized transmuted-G family of distributions. *Communications in Statistics-Theory and Methods*. 46(8), 4119-4136.

6. Shabbir, M., Noor, N., Riaz, A. and Gull, H. (2017). The New Weibull Lomax Distribution. *Imperial Journal of Interdisciplinary Research*, 3(1), 1881-1885.
7. Shabbir, M., Riaz, A. and Gull, H. (2018). Rayleigh Lomax Distribution. *The Journal of Middle East and North Africa Sciences*, 4(12), 1-4.
8. Yousof, H.M., Afify, A.Z., Alizadeh, M., Butt, N.S., Hamedani, G.G. and Ali, M.M. (2015). The transmuted exponentiated generalized-G family of distributions. *Pakistan Journal of Statistics and Operation Research*, 11(4), 441-464.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 169-178

IMPACTS OF TECHNOLOGY ON YOUTH: A CASE STUDY IN GC WOMEN UNIVERSITY SIALKOT, PAKISTAN

Tehmeena Yousaf, Mahwish Rabia[§], Ayesha Shajhan and **Azab Liaqat**
GC Women University Sialkot, Pakistan.
[§]Corresponding author Email: mahwish.rabia@gcwus.edu.pk

ABSTRACT

This paper attempts to investigate the effects of technology on youth. Technology can have large impact on user mental and physical health. Technology has many different effects on education one of them being enhancing the students learning. The aim of this study is therefore to analyze the casual effects of technology on youth. The data is collected using survey method sample of 373 students are taken from GC Women University Sialkot. 52% students are strongly agreed that technology is necessary for education. On the other hand, 48.8% students think that technology addiction is a bigger problem now a day's and 45.8% agree that people are connected to each other through social media. The significant result shows that technology has positive effect as well as have negative effects on youth. Moreover, the relationship between variables i.e. technology addiction and sleep disorder is calculated by using correlation coefficient. The fitted regression model demonstrates that technology addiction is a bigger problem now days. It is concluded that technology has positive effects as well as have negative effects on youth.

KEYWORDS

Technology, technology addiction, Effects on Health, Awareness.

1. INTRODUCTION

Technology can be the knowledge of techniques processes and the like or it can be embedded in machine to allow for operation without detailed knowledge of their working the simplest form of technology is the development and use of basic tools. Technology has revolutionized the lives of people. Technology has played a prominent role in the development of various industries; it has changed the banking sector, education, agricultural industry, entertainment world, it has restructured many businesses. Technology is continuously changing. These growing changes bring social and economic consequences on different aspects of our daily life. In the future, these changes will happen faster, with less time to prepare. Our approach is to avoid wrong choice otherwise our technology will destroy us. Automobiles provided quicker and convenient mode of transportation caused freedom from family bonds and mixing with different cultures and new communities. Television became source of entertainment, learning and information, with social impacts like family interaction, leisure time, aggressiveness, mixing up of different cultures and fashion.

1.1 Impact of Technology

The impacts of technology cannot be measured because it is still changing the way we do everything. However, technology also has some adverse effects. Below I have listed some impacts of technology on our lives both positive and negative. Thus two aspects of this changing technology emerged that is “Positive and Negative impact” on the users.

1.2 Positive Impacts of Technology

1.2.1 Globalization of Education

Technology has made education remote. With the help of internet technology, online education has become a strong force in the education sector. Now students can study courses which are provided in other countries without having any boundary limitations.

1.2.2 Improved Transportation Sector

Technology has made the movement so easy and cheap, technologies like automobiles, airplanes, speedboats, electronic trains, have made moving from one place to another so easily.

1.2.3 Improved on Human Relationships and Connection

Technology has improved on how we connect or discover new relationships. With things like mobile phones and internet social Medias, people can connect with each so easily than before. Confidence Boosters: Lastly, technology can benefit students by boosting their confidence. Many students come into school already well-versed in technology; sometimes, students are more knowledgeable than their teachers

1.2.4 Role of Technology in Education

The role of computers and Internet in education cannot be ignored. The use of computers in teaching has made learning more interesting. Computers making the process of teaching easier and more effective. Educational CDs are available everywhere. Lectures/lessons can be uploaded on websites, as PDFs or even as videos. This has made information more easily accessible for students. Online degrees and training courses have led to spread of education. Today, people can continue learning even while managing their jobs, because online education.

1.3. Negative Impacts

1.3.1 Exposure to Wrong Data

Though technology has made learning so easy and cheap, students get exposed to data which is not approved and many times they just copy and pastes this information they acquire online without any deep research on whether the data is correct. This has increased the level of misinformation and failure in exam.

1.3.2 Lack of Real Life and Friends

Yes, social networks and text messaging technologies have played a very big role in connecting people more than before, but what is funny is that the number of people who spend time alone is increasing. You will find a person spends more time in the virtual world chatting with strangers ‘’virtual friends’’, a user will have over 500 friends on a specific social network, but they have no real friends in the real world, WHY? Because they spend most their time in the virtual world and give up on getting real friends, which results in loneliness and cyber-sickness.

1.4 Is Technology Important for us or not

Have you asked yourself why technology is so important today? Just look around and you'll know why. Literally, at every instant of time, you are surrounded by technology. Whether you are working or resting, you are always using technology. It is used everywhere and all the time. Places far away have come closer, thanks to technology. Communication is rapid, travel is fast, movement is easy, action is quick, interaction has sped up and so has life. Things that once took hours to complete, can be done in seconds today. The world is smaller and life is fast.

1.4.1 Why Technology is Important

Owing to the application of technology, our standard of living has increased. Our needs are met with greater ease. Owing to the developments in engineering and architecture, building robust civil structures has been possible. Technology has boosted every industry. Businesses have grown, creating more employment opportunities. Advancements in technology have led to the evolution of newer and faster modes of transport and communication. They have changed every sector; be it medicine, tourism, education, entertainment or any other. Technology has changed living.

Significance of the Study

- To determine whether technology is necessary for students.
- To examine the positive and negative impacts of technology on students.
- To examine whether the technology is helpful for students in their study.
- To examine whether the relationship is affected by the technology or not.
- To determine how much the health of students is affected by technology.

2. LITERATURE REVIEW

Farlie and Kalil (2016) performed a trial in California on technology and education**.** In California where over thousand computers were distributed randomly for free to children attending 6th -7thgrade to use at home they observe that the children who are given computers are more associated to have a social networking site, but also spend more time interacting with friend in person. There are no casual effects found in educational outcomes of the students their studies are not effected by using computers and found a small positive benefit to children's social development by using computers. Pater Osharive (2015) examined the relationship between social media and academic performance of students. The total population of 2466 they collect a sample of 378 students. Testing the research hypothesis, the results show that the students are addicted to social media. A balance creates between social media and academic activities of students to avoid problem/issues/difficulties in the academic performance of the students. Floyd and Judge (2012) conducted a study on the micro level, following the progress of 6 students who had some form of a learning disability. The study was completed through the use of piece of technology called Classmate Reader. A reading and comprehension passage was given to all students. Students were then asked to test using traditional pen and paper methods, followed by a second assignment completed uses the classmate Reader. This result showed that the use of assistive technology is an effective support and accommodation for students with learning disabilities. Heafner (2004) examined the effects of technology on student motivation in a classroom. As a part of their assignment students were required to make a power point slides. Although standard classroom and hallway behavior was noted, once students arrived at the computer lab, it was noted that students exhibited a marked change in behavior. Students began to get excited about learning, and showed pride in their work. All students reported enjoying the assignment and stated that they felt more motivated. Wellman et al. (2001) conducted a survey on the National Geographic Society Website. Came out with the suggestion that the internet increase connectivity in involvement not only can expose people to more connect and information, it can reduce commitment new technologies may try to create a new life. Some technologies had greater social impacts than others. Katz and Aspden (1997) conducted a survey using 2500 respondents, out of which 8 percent were Internet users. Comparing users with non-users, they found no evidence of Internet use reducing people's membership in social and religious organizations. Among users, greater use of the Internet was associated with increased contact with family members and an increased participation in online communities. Flanagan (2008) made a research on one eighth grade math class of nineteen students from an urban middle school was taught a unit on Solving Systems of Equations by Graphing. Students were first introduced to the unit through the use of pencils and graph paper. All the graphing was done by hand and students had to determine the solutions of the systems of equations by proper graphing techniques. The students were then tested at the end of the unit. Research has shown both positive and negative effects on student achievement through the implementation of technology.

3. RESEARCH METHODOLOGY

In this study total population size is 5447 is considered taking BS students of Government College Women University Sialkot. The representative part of population is called sample. Data is collected taking a sample of 373 which is calculated using the formula provided by Yamane (1967) given in equation (1). Where 'N' is the population size and 'e' is the margin of error. The procedure is given below:

$$n = \frac{N}{1 + Ne^2} \tag{1}$$

$$n = \frac{5447}{1 + 5447\,(0.05)^2} = 373 \tag{2}$$

3.1 Sampling Technique:

There are 17 homogeneous groups that consisting of different departments. Therefore, sample is selected using stratified random sampling with proportional allocation. The basic formula of stratified random sampling with proportional allocation is denoted as:

$$n_h = n * \frac{N_h}{N} \tag{3}$$

where h denotes number of strata (Cochran, 1997).

Using equation (2) required sample size is calculated as:

Department	Stratum Size	Selected Sample	Department	Stratum Size	Selected Sample
Political Science	237	16	Fine arts	80	5
Statistics	173	12	Psychology	301	21
English	346	24	Sociology	200	14
Islamiyat	295	20	BBA	350	24
Chemistry	670	46	IT	227	15
Math's	537	37	CS	217	15
Zoology	530	36	E. Science	243	17
Botany	520	36	Urdu	286	19
Economics	235	16			

3.2 Analysis of Data

The following statistical techniques are used to analyze the data:

1) Frequency distribution;
2) Reliability
3) Correlation Coefficient;
4) Regression Analysis;
5) ANOVA.

4. RESULTS AND DISCUSSION

4.1. Descriptive Analysis

Table 4.1: Illustrates that 2.7% people are strongly disagreeing, 2.4% are disagree, 6.2% are neutral, 51.5% are agree and 37.3% are strongly agree that technology is necessary for education. 1.3% peoples are strongly disagreeing, 2.9% are disagree, 12.9% are neutral, 43.2% are agree and 39.7% are strongly agree that technology device makes daily tasks more efficient. Similarly,1.6% peoples are strongly disagreeing, 2.9% are disagree, 10.5% are neutral, 49.6% are agree and 35.4% are strongly agree that technology helpful to cure different diseases accurately in hospitals. 2.7% peoples are strongly disagreeing, 3.2% are disagree, 9.1% are neutral, 45.8% are agree and 39.1% are strongly agree that now a day's people are connected to each other through social media. 2.1% are strongly disagree, 3.2% are disagree, 14.7% are neutral, 39.1% are agree and 40.8% are strongly agree that peoples think that by using modern technology we can earn money at home. 0.8% peoples are strongly disagreeing, 3.2% are disagree, 11.0% are neutral, 46.9% are agree and 38.1% are strongly agree that by using internet we get sufficient material to get knowledge. 2.1% peoples are strongly disagreeing, 1.1% are disagree, 8.8% are neutral, 57.1% are agree and 30.8% are strongly agree that heavy use of cell phones increase sleep disorder and indicates depression. 1.9% peoples are strongly disagreeing, 3.2% are disagree, 8.6% are neutral, 48.8% are agree and 37.5% are strongly agree that technology addiction is a bigger problem now-a-days. 2.4% peoples are strongly disagreeing, 4.0% are disagree, 12.9% are neutral, 35.1% are agree and 45.6% are strongly agree that they got awareness about global issues due to internet. 2.1% are strongly disagree, 4.3% are disagree, 4.6% are neutral, 57.9% are agree and 31.1% are strongly agree that too much use of technology (e.g. mobile phones computer, tabs etc.) damaging our mental health.

Table 4.1
Univariate Data Analysis

Variable	Categories	Number	Percent
Age	18-20	207	55.5
	21-23	166	44.5
Area	Rural	170	45.6
	Urban	203	54.6
Semester	3rd	120	32.2
	5th	124	33.2
	7th	129	34.6
No. of hours spent on using technology (mobile, laptop, iPad)	1-4	100	26.8
	5-10	130	34.85
	11-15	143	38.34
Do you think that technology is necessary for education?	S.D	10	2.7
	D	9	2.4
	N	23	6.2
	A	192	51.5
	S.A	139	37.3
Do you think technology device makes daily tasks more efficient?	S.D	5	1.3
	D	11	2.9
	N	48	12.9
	A	161	43.23
	S.A	148	39.7
Is technology helpful to cure different diseases accurately in hospitals?	S.D	6	1.6
	D	11	2.9
	N	39	10.5
	A	185	49.6
	S.A	132	35.4
Do you think that now a day's people are connected to each other through social media?	S.D	10	2.7
	D	12	3.2
	N	34	9.1
	A	171	45.8
	S.A	146	39.1
Do you think that by using modern technology we can earn money at home?	S.D	8	2.1
	D	12	3.2
	N	55	14.7
	A	146	39.1
	S.A	152	40.8
Do you think that by using internet we get sufficient material to get knowledge?	S.D	3	0.8
	D	12	3.2
	N	41	11.0
	A	175	46.9
	S.A	142	38.1

Variable	Categories	Number	Percent
Do you feel that heavy use of cell phones increase sleep disorder and indicates depression?	S.D	8	2.1
	D	4	1.1
	N	33	8.8
	A	213	57.1
	S.A	115	30.8
Do you think that addiction of using technology is a bigger problem now a day?	S.D	7	1.9
	D	12	3.2
	N	32	8.6
	A	182	48.8
	S.A	140	37.54
Do you think you have got awareness about global issues due to internet?	S.D	9	2.4
	D	15	4.0
	N	48	12.9
	A	131	35.1
	S.A	170	45.6
Do you think that too much use of technology (e.g. mobile phones computer, tabs etc.) damaging our mental health.	S.D	8	2.1
	D	16	4.3
	N	17	4.6
	A	216	57.9
	S.A	116	31.1

4.2 Advantages and Disadvantages of Technology

The present study significantly determines the advantages (Figure-1) and disadvantages (Figure-2) of technology. Technology has positive effects on youth as well as has negative effects. The most powerful uses of technology include: (i) it can cure diseases accurately in hospitals. (ii) People can connect through social media. (iii) It is helpful to earn money at home through online jobs. (iv) It creates awareness among people. Similarly, the negative effects of technology are (i) Technology addiction is a bigger problem now a day's. (ii) It increases sleep disorder. (iii) It causes depression and damage mental health.

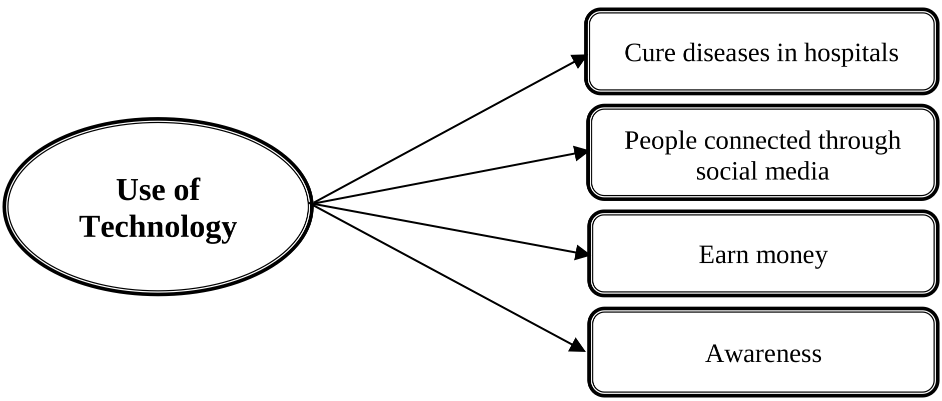


Figure 1: Advantages of Technology

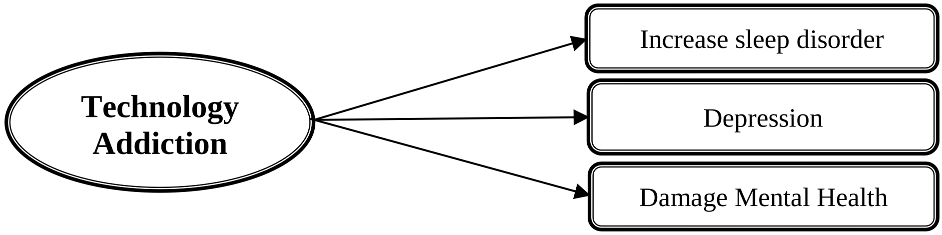


Figure 2: Disadvantages of Technology

4.3 Reliability

"Reliability" is how well a test measures what it should. Table 2 represents the reliability analysis. A value of Cronbach's Alpha which will greater than 0.7 is acceptable. In the present study the value of Cronbach's Alpha is 0.676 which I approximately near to 0.7 which shows that the data is reliable for further statistical analysis.

Table 2
Reliability Statistics

Cronbach's Alpha	N of Items
.676	10

4.4 Inferential Analysis

The process of drawing inferences about population on the basis of sample data is known as inferential statistics.

4.4.1 Correlation Coefficient

The main purpose of the correlation analysis is to measure the strength or degree of linear relationship between the two variables.

Table 3
Correlation Co-efficient between different attributes.

Attributes	Correlation	p-value
Technology addiction and increase sleep disorder	.299**	0.000
Technology is necessary for education and using modern technology people can earn money.	.175**	0.001

**. Correlation is significant at the 0.01 level (2-tailed)

Table 3 demonstrates that the value of correlation co-efficient (r=0.299) shows that there is positive relationship between technology addiction and sleep disorder. Studies from literature approved that technology addiction has also been increasingly linked to insomnia and disturbed sleep.

Moreover, the value of correlation co-efficient (r=0.175) shows that there is positive relationship between technology is necessary for education and using modern technology people can earn more money. Since, modern technology is inspiring students to become creative and innovative. Creativity and innovation will make students successful in their career and life.

4.4.2 ANOVA

H_0: On the average sleeping hours are equal according to the time spend on technology.
H_1: On the average sleeping hours are not equal according to the time spend on technology.

Table 4
Results for ANOVA Test

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	7.639	4	1.910	2.881	.002
Within Groups	738.361	1114	.663		
Total	746.000	1118			
Interpretation: Table 4 shows the significant result as F=2.881 and p-value is .002. The results revealed that null hypothesis is rejected because the average sleeping hours are not equal according to the time spend on technology. It is examined that sleeping disorder depends on heavy use of technology.					

4.4.3 Model Analysis

The simple linear regression model is defined as:

$$y = \beta_0 + \beta_1 x + e$$

where βo is intercept and $\beta_1$is the slope of the model

y = Health Problems (Sleep disorder, Mental Health)
x = Technology Addiction (Tech. addiction, Communication through social media).

Table 6
Analysis of Correlation Co-efficient

Predictor variable	B	S.E	t-Statistic	p-value
(Constant)	3.967	0.206	19.382	0.000
Technology Addiction	0.102	0.049	2.144	0.033
R^2 =0.72 S.E=0.056		F= 4.597 , p-value=0.033		
$\hat{Y} = 3.967 + 0.102X$				

In Table 6 the value of R^2 = 0.72 it shows that there is 72% variation is explained by the independent variable. The simple linear regression model is based on two variables where dependent variable is health problems while independent variable is technology addiction. In addition, both parameters (βo =3.96, β_1=0.102) are significant that is p-value is less than significant level ($\alpha < 0.05$). The fitted regression model shows that although people communicate to each other through social media but technology addiction is harmful for the health of people because too much use of technology damage mental health because technology addiction is a bigger problem nowadays.

5. COMMENTS AND CONCLUSION

Technology has become a part and parcel of almost everyone’s live, especially for the youth. While technology has done wonders for the world, the debate is still on as to how good or bad technology is for us. Technology has connected the world like messaging applications, social media, broadcasting systems. This study attempts to investigate the impact of technology on youth. Data is collected through questionnaire. The results demonstrate that, the point of view of mostly people is that, technology is necessary for education and it makes our daily task more efficient. It is also observed that technology is helpful to cure diseases and peoples are connected through social media. On the other hand, most of the respondents are strongly agree, that technology addiction is a big problem now a day's, it increases sleep disorder and depression. The significant results obtained through simple linear regression show that technology addiction can damage mental health. Because the excessive use of technology not only destroys student's future but also produces stress, anxiety or fear which leads to depression. It can be concluded that technology has positive as well as negative effects on youth.

Suggestions:

1) Technology has become inspirable part of our life we need to adopt it in a way that we don't harm ourselves.
2) It is good to embrace new technologies but completely dependent on them without overcoming their draw backs is harmful.
3) It is indeed the need of an hour for humans to keep thinking and formulating a sustainable solution for all the technological problem.
4) Government should introduce new technologies in education sector.

REFFERENCES

1. Farlie, R.W. and Kalil, A. (2016). The Effect of Computers on Children's Social Development and School Participation: Evidence from a Randomized Control Experiment. *National Bureau of Economic Research, NBER Working Paper* No. 22907.
2. Osharive, P. (2015). Social media and academic performance of students. *Research project submitted to Department of Educational Administration*, 100302125.
3. Floyd, K.K. and Judge, S.L. (2012). The efficacy of assistive technology on reading comprehension for postsecondary students with learning disabilities. *Assistive Technology Outcomes and Benefits*, 8, 48-64.
4. Heafner, T. (2004). Using technology to motivate students to learn social studies. *Contemporary Issues in Technology and Teacher Education*, 4, 42-53.
5. Wellman, & K. Frank, (2001). Network capital in a multilevel world: Getting support from personal communities, in N. Lin, R. Burt, & K. Cook (Eds.), *Social capital: Theory and research*, Chicago, MI.
6. McAdam, D. (2010). Political process and the development of black insurgency, 1930-1970. University of Chicago Press, 2010.
7. Katz, J. and Aspden, P. (1997). Motivations for and Barriers to Internet Usage: Results of a national Public. Opinion survey. *Internet Research-Electronic Networking Applications and Policy*, 7(3), 170-188.
8. Flanagan, J.L. (2008). *Technology: The positive and negative effects on student achievement*, Master Thesis, State University of New York College at Brockport
9. Yamane, T. (1967). Statistics: *An Introductory Analysis*. Harper and Row, New York.
10. Cochran, W.G. (1977). *Sampling Techniques*. John Wiley and Sons, New York.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 179-187

STUDY OF ENVIRONMENTAL FACTORS AFFECTING CLIMATE CHANGE IN LAHORE OVER TIME BY USING MULTIPLE REGRESSION

Ruqayya Farooq, Naima Mubarak and **Humma Nawaz**
Department of Statistics, Govt. College Women University, Sialkot, Pakistan
Email: ruqiafarooq77@gmail.com
naima.mubarak@gcwus.edu.pk
nawazhuma75@gmail.com

ABSTRACT

This study investigated the relationship between temperature, sea pressure, humidity, precipitation, visibility, wind speed and rain. The relationship between the variables is modeled by using multiple linear regression model. The assumptions of the regression model have been checked using normal probability plot, residual vs. fitted plot, variance inflation factor and Durbin Watson statistic. The results showed that all the assumptions of the model are fullfilled. The adjusted R square showed that model is best fitted on our data. The overall significance of the model and significance of the slope coefficients have been checked on the basis of p-value. The prediction for the year 2001 to 2018 has been done using multiple linear regression model.

KEYWORDS

Climate, Temperature, Multiple Regression, Prediction, Variance Inflation Factor

1. INTRODUCTION

Parameters of Climate Change

Climate

It is the statistics of weather over long periods of time. It is measured by assessing the patterns of variation in temperature, humidity, atmospheric pressure, wind, precipitation, atmospheric particle count and other meteorological variables in a given region over long periods of time.

Weather

Weather describes what is happening around us & in the atmosphere. The difference between weather and climate is a measure of time. Weather is what conditions of the atmosphere are over a short period of time, and climate is how the atmosphere "behaves" over relatively long periods of time.

Wind

Wind is expresses with direction and speed. The oversight of wind is described as direction from which the wind is blowing. Wind is a great cause of transportation for tiny seed and small birds with time things can flying thousands of miles in the globe.

Temperature

Temperature is calculating how hot or cold specially calculating the average of kinetic energy of tiny particular.

Humidity

It is an indicator used to describe the amount of water vapor present in air. Too much increase of humidity or too decrease in humidity can be harmful. If humidity combined with hot temperature than health risk can occur in young or old age people. Humidity plays essential role, without water vapor our weather becomes weather of mars.

Precipitation

All liquid and solid water drop that collapse from clouds and arrived at ground. These particles may contains drizzle, rain, snow, and hails. Precipitation is one of three leading process evaporation, condensation, precipitation, that initiate the hydrological cycle the endless interchange of water between atmosphere and earth's surface. Water dehydrate from ocean, land, and freshwater surfaces is sustain as cloud. The necessary difference between precipitation drop and a cloud particle is one of size, because of their huge magnitude precipitation particles have notable falling speed and are skilful to survive the fall to the ground.

Reasons of Climate Change

Natural processes of solar and volcanic activities can cause of climate changing. Land, sea and atmosphere relations can cause of usual cyclic variations in weather. Anthropogenic activities are mostly responsible for global. Anthropogenic causes are manageable but they been controlling by natural due to this balance of the atmospheric heat has been disturbed and more amount of heat has been stuck in the biosphere than usually required to control the life processes (Peter et al., 1997). The atmospheric absorption of all the GHGs has been increased and ratio is increase from gas to gas. A particle of sulphur Hexafluoride (SF6) can occur in atmosphere 3200 times longer time than a particle of Carbon Dioxide due to the longer life time of these gases, the scientists say that ongoing warming tendency will remain for the afterward 50 years at the normal rate. Similarly a kilogram of methane can warmness up atmosphere 72 times further than a kilogram of carbon. Increasing absorption of GHGs in the atmosphere is not only creating problems related to climate variability but also signals severe consequences in future.

Jonathan et al. (2000) explained the effects of changing weather on public health. He conducted study on environmental health concerns have traditionally focused on toxicological infectious risk to human health from local factors.

Anjum et al. (2005) studied Climate change perspective in Pakistan. They assessed the past climate changes and compute the projected changes in different Agro-climatic regions of country for next half century using Regional and Global Climate models. Preceding record and projections by GCMs and RCMs depicts that extreme events drought and flooding would become more frequent and of greater magnitude in different parts of the country.

Ramanathan et al. (2008) conducted study on Air pollution, greenhouse gases and climate change: Global and regional perspectives. He suggested that Greenhouse gases

(GHGs) warm the surface and the atmosphere with significant implications for rainfall, retreat of glaciers and sea ice, sea level, among other factors. About 30 years ago, it was recognized that the increase in tropospheric ozone from air pollution (NOx, CO and others) is an important greenhouse forcing term. The dimming has a surface cooling effect and decreases evaporation of moisture from the surface, thus slows down the hydrological cycle.

Rizwana et al. (2012) Predicted impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. They conducted study on Maximum entropy (Maxent) modeling was used to predict the potential climatic niches of three medicinally important Asclepiad species: Pentatropis spiralis, Tylophora hirsuta, and Vincetoxicum arnottianum. They differ in ecological requirements, biogeographic importance, and conservation value. The Maxent model predicted habitat gains for P. spiralis in southern Punjab and Balochistan, and loss of habitat in south-eastern Sindh. Vincetoxicum arnottianum as well as T. hirsuta would gain habitat in upper Peaks of northern parts of Pakistan.

Aslam et al. (2018) conducted study on integrated climate risk assessment & evaluation of adaptation perspective in southern Punjab Pakistan. He proposed climate change risks are continuously increasing. The main issue in climate change integrated risk assessment & management are unseen interdependencies between risks which are attributed in physical, geographical, economic policies & social mechanism of that specific area.

2. MATERIALS AND METHODS

Multiple Linear Regression Analysis

In this technique we measures cause and effect relationship between given variables. Also we used to predict the value of a variable founded on the value of two or more explanatory variables. The prediction always based on the regressed variables. The universal determination of multiple regression (the term was first used by Pearson, 1908) is to study more about the relationship between numerous independent or predictor variables and a dependent variable.

Regression Model

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 ... + \beta_j x_j + \varepsilon_i$$

Assumptions of the model

Linearity of the Data

Linear regression essentials the relationship between the independent and dependent variables to be linear. It is also significant to check for outliers mean while linear regression is sensitive to outlier effects. The linearity assumption can best be tested with scatter plots. Linear regression adopts that there is little or no multicollinearity in the data. Also independence assumption is that the error of the mean has to be independent from the independent variables. Linear regression analysis requires that there is little or no autocorrelation in the data. Autocorrelation occurs when the residuals are not independent from each other.

Normal Probability Plot

It is a graphical technique for evaluating whether or not a facts are almost normally distributed. The data are plotted against a theoretical normal distribution in such a way that the points should form an estimated straight line. Departures from this straight line indicate departures from normality.

Homoscedasticity

The assumption of homoscedasticity meaning “same variance” is essential to linear regression models. Homoscedasticity defines a condition in which the error term in the relationship between with the independent variables and the dependent. The influence of violating the assumption of homoscedasticity is a matter of degree, increasing as heteroscedasticity increases. A homoscedasticity plot is a graphical data analysis method for measuring the assumption of constant variance across subsets of the data.

Autocorrelation

The autocorrelation (***Box and Jenkins, 1976***) function can be used for the following two purposes:

1. To detect non-randomness in data.
2. To identify an appropriate time series model if the data are not random.

Autocorrelation is a characteristic of data in which the correlation between the values of the same variables is based on related objects. It violates the assumption of instance independence, which underlies most of the conventional models. It generally exists in those types of data-sets in which the data, instead of being randomly selected, is from the same source.

Variance Inflation Factor (VIF)

Variance inflation factor (*VIF*) quantifies how much the variance is inflated. Variances of the estimated coefficients are inflated when multicollinearity exists. Multicollinearity is the relationship between explanatory variables. There should be no linear relationship among explanatory variables, and we detect multicollinearity through VIF.

Durbin Watson Statistic

The Durbin Watson (DW) statistic is a test for autocorrelation in the residuals from a statistical regression analysis. The Durbin-Watson statistic will always have a value between 0 and 4. A value of 2.0 means that there is no autocorrelation detected in the sample. Values from 0 to less than 2 indicate positive autocorrelation and values from 2 to 4 indicate negative autocorrelation.

3. RESULTS AND DISCUSSION

Data Description

We take yearly data from source ***https://en.tutiempo.net/climate*** of climatic variables Temperature, Maximum Temp, Minimum Temp, Sea Pressure, Humidity, Precipitation, Visibility, Wind Speed, Maximum Wind Speed, Rain, Snow, Storm, and Fog from 2001 to 2018. For future climate prediction we apply Multiple Regression Analysis technique and some variable shows insignificant results for the regression model and we omit irrelevant variables. And we get this equation:

$$Temp = 931.7 - 0.8950 \text{ Sea Pressure} - 0.1025 \text{ Humidity} - 0.00067 \text{ Precipitation} + 0.965 \text{ Visibility} - 0.599 \text{ Wind Speed} - 0.1594 \text{ Rain}$$

As we can see that our variables has inverse relationship except visibility if we change one unit in sea pressure than there will be 0.8 unit change occur in temperature. Similarly if we change one unit in humidity, precipitation, wind speed and rain then there will be 0.1025, 0.00067, 0.599 and 0.1594 unit change in temperature respectively. If we change one unit in visibility than there will be 0.9 unit change occur in temperature.

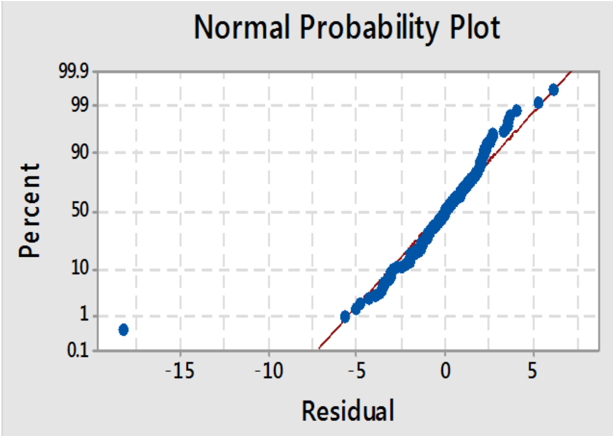


Fig. 1: Normality Probability Plot

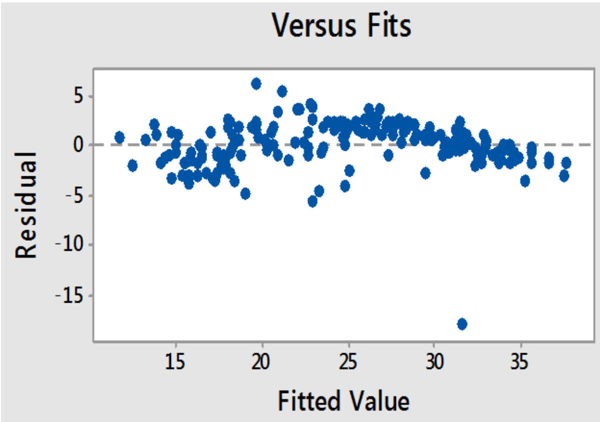


Fig. 2: Residual verses Fitted Graph

From normal probability plot of residuals, we can observe that all our data values lies around a standard normal line. It means residuals are normally distributed. From the residuals vs. fitted diagram, we conclude that residuals are randomly distributed and have constant variance.so assumption of homoscedasticity has been full filled.

Pearson's Correlation Coefficient

Table 1.6
Correlation Coefficient

	Temp	Humidity	Precipitation	Visibility	Wind Speed	Rain	Sea Pressure
Pearson Correlation	1	-.355*	.267**	.283**	.373**	.267**	-.917**
Sig.(2-tailed)		.000	.000	.000	.000	.000	.000
N	210	210	210	210	210	210	204

From the provided evidence as we see that our P-value < α, so we conclude that our results are significant at prescribed level of significance.

Variance inflation factor (VIF)

Table 3.1
Collinearity Statistics

Model	95.0% Confidence Interval for B		Collinearity Statistics	
	Lower Bound	Upper Bound	Tolerance	VIF
(Constant)	876.042	987.361		
Sea Pressure	-.950	-.840	.609	1.641
Humidity	-.133	-.072	.620	1.614
Precipitation	-.003	.001	.624	1.604
Visibility	.451	1.479	.532	1.879
Wind Speed	-.894	-.304	.411	2.432
Rain	-.297	-.021	.521	1.920

From above table, we can conclude that there is no multicollinearity in the variables because value of VIF is close to one for all variables.

Autocorrelation

Table 3.2
Durbin Watson Statistic

Test Statistic	Value
Durbin Watson	1.6053

As the value of Durbin Watson statistic is close to 2 which shows that there is no autocorrelation between residuals.

Overall significance of the model

Table 3.3
ANOVA Table

Model		Sum of Squares	df	Mean Square	F	Sig.
1	**Regression**	9054.182	6	1509.030	273.778	.000***
	Residual	1085.839	197	5.512		
	Total	10140.022	203			

Conclusion:
From the provided evidence, as our p value 0.00< α. so we can say that the slope coefficients are jointly significant at 5% level of significance.

Individual Testing of hypothesis

Table 3.4
Individual Testing of Slope Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	931.702	28.224		33.011	.000
Sea Pressure	-.895	.028	-.954	-31.954	.000
Humidity	-.102	.015	-.198	-6.676	.000
Precipitation	-.001	.001	-.019	-.640	.523
Visibility	.965	.260	.118	3.705	.000
Wind Speed	-.599	.150	-.145	-4.002	.000
Rain	-.159	.070	-.074	-2.278	.024

Conclusion:

From the provided evidence, we can observe that our regressor sea pressure, humidity, visibility, windspeed and rain shows significant results except precipitation at 5% prescribed level of significance.

Goodness of Fit of Model

Table 3.5
Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.945[a]	.893	.890	2.34774

As value of adjusted R-Square is 0.890 which shows that independent variables explain 89% variation of dependent variable. So our model is best fitted on our data.

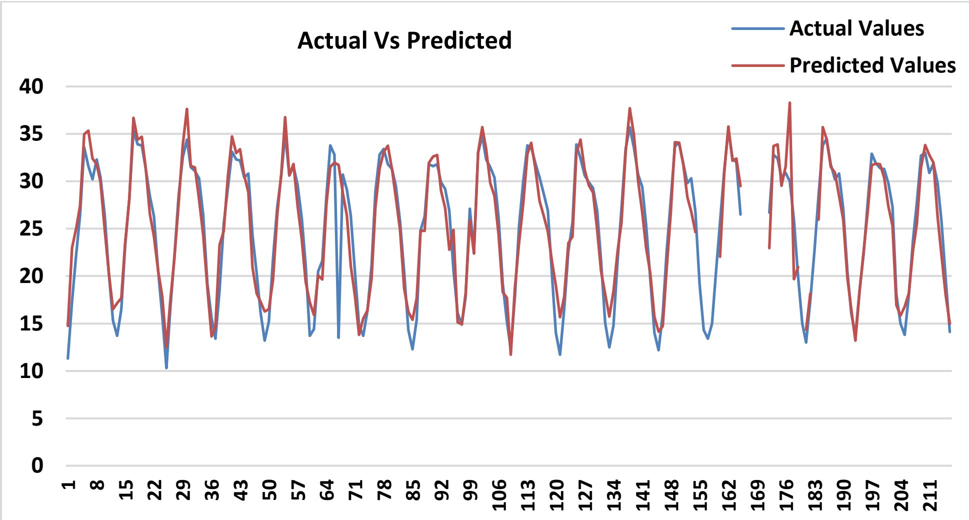


Fig. 3: Actual and Predicted Graph for Temperature

From the Figure 3, we can see that our actual and predicted graph is approximately same which means that our predicted values are accurate. So we conclude that multiple linear regression model gives accurate prediction for the temperature.

4. CONCLUSION

In this paper, we have fitted the multiple linear regression model on our data. Firstly, we have checked the assumptions of normality through normal probability plot of residuals. Our data values lies around a standard normal line. So the assumption of normality has been fulfilled. We checked the assumption of homoscedasticity through residuals vs. fitted graph. We can observe that our residual are random and doesn't shows any pattern which means that error variance is constant. We also checked the assumption of multicollinearity and detect the multicollinearity through Variance Inflation Factor. On the basis of VIF, it is concluded that value of VIF is greater than 1 which indicated that regressors are not correlated. Also we have checked the assumption of autocorrelation through Durbin Watson statistics. Its value is close to 2 so we conclude that there is no autocorrelation. The Pearson's correlation coefficient have been calculated between the dependent variable and the independent variables. On the basis of P-value, it is concluded that correlation among the variables is highly significant.

We checked the overall significance of the model through p-value which shows that all the variables are jointly significant. All the slope coefficients are tested using t-statistic. On the basis of P-value, it is concluded that humidity, visibility, sea pressure, wind speed, rain shows significant effect and precipitation shows insignificant effect. The prediction of dependent variable i.e., Temperature has been done by using regression model. The graph of actual and predicted values showed that our prediction is accurate. Also the goodness of fit of the model is decided on the basis of adjusted R square. The value of adjusted R square is 0.890 shows that our model is best fitted.

REFERENCES

1. Patz, J.A., Engelberg, D. andf Last, J. (2000). The effects of changing weather on public health. *Annual Review of Public Health*, 21(1), 271-307.
2. Farooqi, A.B., Khan, A.H. and Mir, H. (2005). Climate change perspective in Pakistan. *Pakistan Journal of Meteorology*, 2(3), 11-21.
3. Ramanathan, V. (2007). Air pollution, greenhouse gases and climate change. In *AGU Fall Meeting Abstracts*.
4. Chaudhry, Q.U.Z. (2017). *Climate Change Profile of Pakistan*. Asian Development Bank.
5. Aslam, A.Q., Ahmad, I., Ahmad, S.R., Hussain, Y., Hussain, M.S., Shamshad, J. and Zaidi, S.J.A. (2018). Integrated climate change risk assessment and evaluation of adaptation perspective in southern Punjab, Pakistan. *Science of the Total Environment*, 628, 1422-1436.
6. Paquette, A., Vayreda, J., Coll, L., Messier, C. and Retana, J. (2018). Climate change could negate positive tree diversity effects on forest productivity: a study across five climate types in Spain and Canada. Ecosystems, 21(5), 960-970.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 189-198

EXPONENTIATED GENERALIZED FRÉCHET GEOMETRIC DISTRIBUTION

Aqsa Rafique and **Naz Saud**
Department of Statistics, Lahore College for Women University
Lahore, Pakistan
Email: aqsa.aim@gmail.com
naz.saud17@gmail.com

ABSTRACT

A five-parameter lifetime model, called Exponentiated Generalized Fréchet Geometric distribution, is introduced and studied. Four sub-models of the newly derived model are presented. The probability function of the EGFG distribution can be unimodal and reserve j-shaped. Descriptive measures, including mean, median, variance and moments are derived. The reliability measures such as survival, hazard rate and reverse hazard rate function are studied. Estimate the parameters of the EGFG distribution using maximum likelihood method. To illustrate the appropriateness of the proposed distribution, a real data set is used.

1. INTRODUCTION

Statistical distributions have become propitious in describing and predicting the real world application. The probability distributions have widely used in modeling lifetime data. The motivation is to suggest a lifetime distribution which has a modified upside-down-bathtub, increasing, decreasing and reverse J-shaped of the mortality rate function.

A French mathematician Kendall in 1977 developed the fréchet distribution. The Fréchet distribution that is employed in extreme value theory. The aim of EVT is to predict probabilities for maximum or minimum of the extreme values. Extreme Value Theory measuring events occurring with very small probability (da Silva et al., 2013). Many lifetimes models such as the Gompertz Fréchet distribution (Oguntunde et al., 2019), the Beta Exponentiated Fréchet distribution (Mead et al., 2017) are introduced. The distribution function and density function of Fréchet distribution are defined by

$$G(x) = exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\} \quad and \quad g(x) = \phi\gamma^{\phi}x^{-(\phi+1)}exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}$$

$$x > 0, \gamma > 0, \phi > 0$$

where γ is the scale parameter and shape parameter is ϕ.

In this paper, we develop the Exponentiated Generalized Fréchet Geometric (EGFG) distribution which is more flexible than Exponentiated Generalized Fréchet distribution. EGFG distribution contains the GFG, FG, GIRG and IRG distributions as sub-models.

The distribution function and density function of Exponentiated Generalized Fréchet distribution is defined respectively,

$$G(x) = \left[1 - \left(1 - \exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}$$

$$x > 0, \gamma > 0, \kappa > 0, \delta > 0, \phi > 0 \tag{1}$$

$$g(x) = \kappa\delta\phi\gamma^{\phi}x^{-(\phi+1)}exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa-1}$$
$$\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta-1} \tag{2}$$

where κ, δ and ϕ are shape parameters and scale parameter is γ.
(Cordeiro et al., 2013).

By using the power series expansion for w>0, a real non-integer

$$(1-s)^{w-1} = \sum_{h=0}^{\infty}\binom{w-1}{h}(-1)^h s^h \tag{3}$$

under the condition ($|s| < 1$)
(Elbatal and Muhammed, 2014).

We can write expansion of probability distribution and density function of EGF distribution as,

$$G(x) = \sum_{v=0}^{\infty}\sum_{t=0}^{\infty}(-1)^{v+t}\binom{\delta-1}{v}\binom{\kappa v}{t}exp\left\{-t\left(\frac{\gamma}{x}\right)^{\phi}\right\} \tag{4}$$

$$g(x) = \sum_{v=0}^{\infty}\sum_{t=0}^{\infty}(-1)^{v+t}\binom{\delta-1}{v}\binom{\kappa(v+1)}{t}\kappa\delta\phi\gamma^{\phi}x^{-(\phi+1)}$$
$$exp\left\{-(t+1)\left(\frac{\gamma}{x}\right)^{\phi}\right\} \tag{5}$$

The Exponentiated Generalized Fréchet distribution are useful in the field of risk management, insurance, hydrology, management and engineering. Exponentiated Generalized Fréchet distribution contain two extra parameters that can manipulate both tail weights.

Given U, Let $\{X_i\}_{i=1}^{u}$, be independent identically distributed (iid)) random variables from EGF distribution. Consider U be a zero truncated geometric variable which is given by

$$P(U = u;\tau) = (1-\tau)\tau^{u-1} \quad 0 < \tau < 1 \tag{6}$$

(Barreto-Souza et al., 2011)

Let $\min\{X_i\}_{i=1}^{u}$, then the marginal density function of X is as follows

Then EGFG dist. has the cdf given by

$$F(x;\gamma,\kappa,\phi,\delta,\tau) = 1 - \frac{(1-\tau)\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)}{1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)} \tag{8}$$

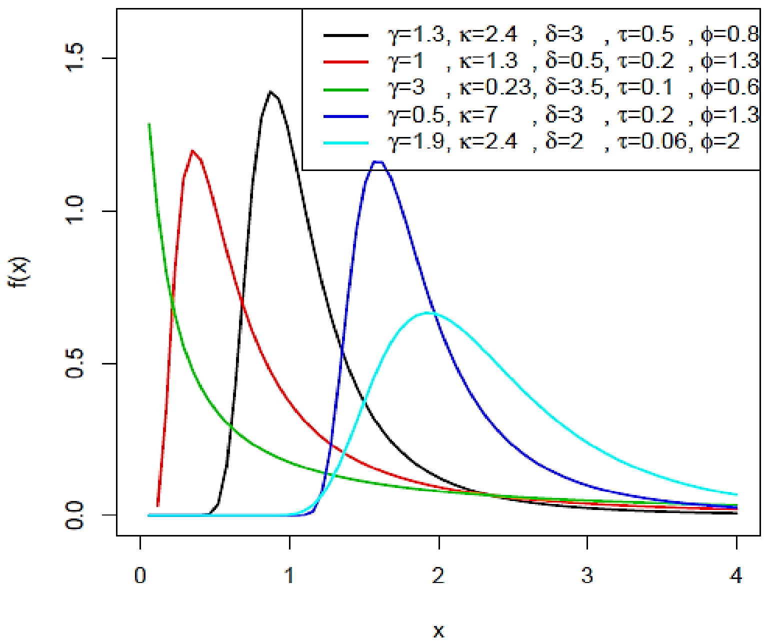


Figure 1: Probability Density Function of Exponentiated Generalized Fréchet Geometric Distribution

The basic inspirations of the suggested model are

- to enhance the properties of the proposed distribution.
- to study shape of the proposed distribution which is more flexible as compared to EGF distribution.
- to state closed expression of pdf using binomial expression.
- to get the maximum likelihood estimators of the newly derive model.
- to examine the goodness of fit of the proposed model by using R statistical software.

2. RELIABILITY ANALYSIS OF EXPONENTIATED GENERALIZED FRÉCHET GEOMETRIC DISTRIBUTION

In this section, we derived the survival function, hazard rate and reverse hazard rate function.

2.1. Survival Function and Hazard Rate Function

The survival and hazard rate function of EGFG dist. is obtained as

$$S(x;\gamma,\kappa,\phi,\delta,\tau)=\frac{(1-\tau)\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)}{1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)} \tag{9}$$

$$h(x;\gamma,\kappa,\phi,\delta,\tau)=\frac{(1-\tau)\kappa\delta\phi\gamma^{\phi}x^{-(\phi+1)}exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa-1}\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta-1}}{\left(1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)\right)(1-\tau)\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)} \tag{10}$$

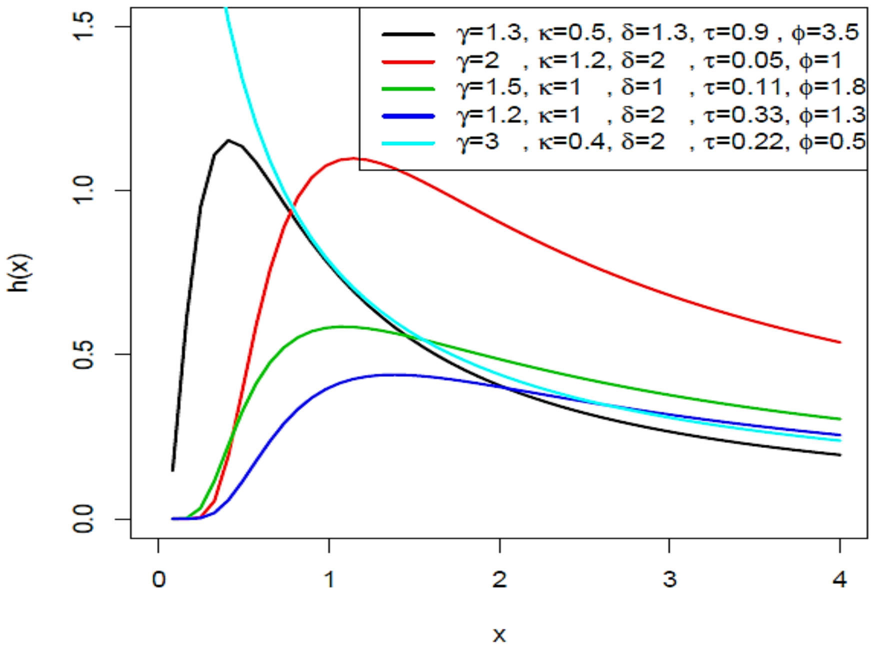


Figure 2: Hazard Rate Function of the Exponentiated Generalized Fréchet Geometric (EGFG) distribution

2.2. **Reverse Hazard Rate Function.** The reverse hazard rate function of the EGFG distribution is expressed as

$$R(x)=\frac{(1-\tau)\kappa\delta\phi\gamma^{\phi}x^{-(\phi+1)}exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa-1}\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta-1}}{\left(1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)\right)\left[1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)-(1-\tau)\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)\right]} \tag{11}$$

3. LIMITING BEHAVIOUR

3.1. Proposition 1

The density of EGFG model can be shown as infinite representation of EGF model with parameter κ, $a\delta$, ϕ, γ. Using pdf from (7)

$$f(x;\gamma,\kappa,\phi,\delta,\tau)=(1-\tau)\kappa\delta\phi\gamma^{\phi}x^{-(\phi+1)}exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa-1}$$
$$\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta-1}\left(1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)\right)^{-2}$$

The expression $(1-r(1-[1-(1-\exp\{-(\frac{\gamma}{x})^{\phi}\})^{\kappa}]^{\delta}))^{-2}$ can be rewritten by using negative binomial series representation i.e.

$$(1-w)^{-b}=\sum_{s=0}^{\infty}\frac{\Gamma(s+b)}{\Gamma(b)s!}=\sum_{s=0}^{\infty}\binom{s+b-1}{b-1}w^{s} \tag{12}$$

under the condition ($|b|>0$) is the real number (Barreto-Souza et al., 2011).

$$\left(1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)\right)^{-2}=\sum_{z=1}^{\infty}\sum_{a=1}^{\infty}\tau^{z-1}a(1)^{a-1}\binom{z}{a}\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta(a-1)} \tag{13}$$

Substituting (13) in (7), we have

$$f(x)=\sum_{z=1}^{\infty}\sum_{a=1}^{\infty}\binom{z}{a}\tau^{z-1}(1)^{a-1}(1-\tau)\kappa\delta a\phi\gamma^{\phi}x^{-(\phi+1)}$$
$$exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa-1}$$
$$\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta a-1}$$

again using (3), we obtained;

$$f(x)=\sum_{b=0}^{\infty}\sum_{m=0}^{\infty}\sum_{z=1}^{\infty}\sum_{a=1}^{\infty}\binom{z}{a}\binom{\delta a-1}{m}\binom{\kappa(m+1)-1}{b}$$
$$(1-\tau)\tau^{z-1}(1)^{a-1+m+b}\kappa\delta a\phi\gamma^{\phi}x^{-(\phi+1)}$$
$$exp\left\{-(b+1)\left(\frac{\gamma}{x}\right)^{\phi}\right\} \tag{14}$$
$$f(x)=\sum_{a=1}^{\infty}\sum_{z=1}^{\infty}(1-\tau)\tau^{z-1}(-1)^{a+b+m-1}\binom{z}{a}g(x,\kappa,a\delta,\phi,\gamma)$$

where,

$$g(x,\kappa,a\delta,\phi,\gamma)=\sum_{b=0}^{\infty}\sum_{m=0}^{\infty}\binom{\delta(a)-1}{m}\binom{\kappa(m+1)-1}{b}\kappa(a\delta)\phi\gamma^{\phi}x^{-(\phi+1)}$$
$$exp\left\{-(b+1)\left(\frac{\gamma}{x}\right)^{\phi}\right\}$$

g(x,$\kappa,a\delta,\phi,\gamma$) is the pdf of EGF distribution. (Elbatal et al., 2018)

$$f(x)=C_{bmaz}\kappa(a\delta)\phi\gamma^{\phi}x^{-(\phi+1)}exp\left\{-(b+1)\left(\frac{\gamma}{x}\right)^{\phi}\right\} \tag{15}$$

where

$$C_{bmaz}=\sum_{b=0}^{\infty}\sum_{m=0}^{\infty}\sum_{a=1}^{\infty}\sum_{z=1}^{\infty}(1-\tau)\tau^{z-1}\binom{z}{a}\binom{\delta(a)-1}{m}$$
$$\binom{\kappa(m+1)-1}{b}(-1)^{a-1+m+b}$$

3.2 Proposition 2

The EGF model is the limiting distribution of EGFG($\kappa,a\delta,\phi,\gamma$) when $\tau^{+}\rightarrow 0$.

$$\lim_{\tau\rightarrow 0^{+}}F(x;\gamma,\kappa,\phi,\delta,\tau)$$
$$=1-\lim_{\alpha\rightarrow 0^{+}}\frac{(1-\tau)\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)}{1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)}$$
$$\lim_{\tau\rightarrow 0^{+}}F(x;\gamma,\kappa,\phi,\delta,\tau)=\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}$$

which is the cdf of EGF distribution.

3.3 Quantile Function

Quantile function is obtained by taking the inverse of the corresponding distribution function. It gives an entire description of the random variable. The d^{th} quantile of the EGFG distribution is obtained by (8)

$$x_d = \gamma\Bigg(-ln\Bigg(1-\Bigg(1-\Bigg(\frac{(1-\tau)q}{1-\tau q}\Bigg)^{\frac{1}{\delta}}\Bigg)^{\frac{1}{\kappa}}\Bigg)\Bigg)^{\frac{-1}{\phi}} \quad 0<q<1 \tag{16}$$

The random number generator of EGFG model is

$$X = \gamma\Bigg(-ln\Bigg(1-\Bigg(1-\Bigg(\frac{(1-\tau)R}{1-\tau R}\Bigg)^{\frac{1}{\delta}}\Bigg)^{\frac{1}{\kappa}}\Bigg)\Bigg)^{\frac{-1}{\phi}} \tag{17}$$

where R follows uniform distribution it is denoted as $R\sim u(0,1)$.

In particular, put q=0.5 in (16) we defined the median of the proposed distribution.

$$x_{\frac{1}{2}} = \gamma\Bigg(-ln\Bigg(1-\Bigg(1-\Bigg(\frac{(1-\tau)\frac{1}{2}}{1-\frac{1}{2}\tau}\Bigg)^{\frac{1}{\delta}}\Bigg)^{\frac{1}{\kappa}}\Bigg)^{\frac{-1}{\phi}} \tag{18}$$

3.4 Moments

Let the random variable X with pdf (15). The K^{th} moments is given by

$$E(X^k) = \int_0^\infty x^k f(x)dx \tag{19}$$

$$E(X^k) = \int_0^\infty x^k C_{bmaz}\kappa(a\delta)\phi\gamma^\phi x^{-(\phi+1)} exp\left\{-(b+1)\left(\frac{\gamma}{x}\right)^\phi\right\}dx$$

$$E(X^k) = C_{bmaz}(b+1)^{\frac{s}{\phi}-1}\kappa a\delta\gamma^s\Gamma\left(1-\frac{s}{\phi}\right) \tag{20}$$

Under the condition $\frac{s}{\phi}<1$.

The mean and variance of the EGFG are defined by respectively

$$E(X) = C_{bmaz}\kappa a\delta(b+1)^{\frac{1}{\phi}-1}\gamma^1\Gamma\left(\frac{-1}{\phi}+1\right) \tag{21}$$

$$variance = \gamma^2\Bigg[C_{bmaz}\kappa a\delta(b+1)^{\frac{2}{\phi}-1}\Gamma\left(\frac{-2}{\phi}+1\right) - \left(C_{bmaz}\kappa a\delta(b+1)^{\frac{1}{\phi}-1}\Gamma\left(\frac{-1}{\phi}+1\right)\right)^2\Bigg] \tag{22}$$

Table 1
Sub-Models from the Exponentiated Generalized Fréchet Geometric (EGFG) Distribution

S#	***γ***	***ϕ***	***κ***	***δ***	***T***	***Models***	***Cdf***
1	$\gamma > 0$	$\phi > 0$	$\kappa = 1$	$\delta > 0$	$0 < \tau < 1$	*GFG*	$1-\frac{(1-\tau)\left(1-\exp\left\{-\delta\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)}{1-\tau\left(1-\exp\left\{-\delta\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)}$
2	$\gamma > 0$	$\phi > 0$	$\kappa = 1$	$\delta = 1$	$0 < \tau < 1$	*FG*	$1-\frac{(1-\tau)\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)}{1-\tau\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{\phi}\right\}\right)}$
3	$\gamma > 0$	$\phi = 2$	$\kappa = 1$	$\delta = 1$	$0 < \tau < 1$	*IRG*	$1-\frac{(1-\tau)\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{2}\right\}\right)}{1-\tau\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{2}\right\}\right)}$
4	$\gamma > 0$	$\phi = 1$	$\kappa = 1$	$\delta = 1$	$0 < \tau < 1$	*IEG*	$\frac{e^{\tau}-exp\left(\tau\left(1-\exp\left\{-\left(\frac{\gamma}{x}\right)^{1}\right\}\right)\right)}{e^{\tau}\;1}$

4. ESTIMATION AND INFERENCE

In this section, we estimate the unknown model parameters of the EGFG model by using maximum likelihood estimation method. Let $x_1, x_2, ...,x_n$ be a random sample from new model with parameters $\underline{\alpha}=\{\kappa, \delta, \phi, \gamma, \tau\}$. The log-likelihood function is given by

$$\begin{aligned}\ell(x,\theta) &= n\ln(1-\tau)+n\ln\kappa+n\ln\delta \\ &+n\ln\phi-(\phi+1)\sum_{i=1}^{n}\ln x_i+n\phi\ln\gamma-\sum_{i=1}^{n}\left(\frac{\gamma}{x}\right)^{\phi} \\ &+(\kappa-1)\sum_{i=1}^{n}\ln\left(1-\exp\left\{-\left(\frac{\gamma}{x_i}\right)^{\phi}\right\}\right) \\ &+(\delta-1)\sum_{i=1}^{n}\ln\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x_i}\right)^{\phi}\right\}\right)^{\kappa}\right] \\ &-2\sum_{i=1}^{n}\ln\left(1-\tau\left(1-\left[1-\left(1-\exp\left\{-\left(\frac{\gamma}{x_i}\right)^{\phi}\right\}\right)^{\kappa}\right]^{\delta}\right)\right)\end{aligned} \tag{23}$$

We obtained the maximum likelihood estimates by taking partially differentiating (23) w.r.t κ, δ, ϕ, γ, τ and equate to zero. Solution of these nonlinear equations cannot possible analytically so we use the fitdisrplus package in R to attain the numerical estimates of the unknown model parameters of EGFG distribution.

5. DATA ANALYSIS

This section deals with a real data application to investigated the model compatibility of EGFG distribution as compared to the other distributions; namely, Exponentiated Generalized Fréchet (EGF), Generalized Fréchet (GF) and Fréchet (F) distribution. The data set is obtained from (Murthy et al., 2004) and represent time between failures for repairable products.

Table 2
Estimates and their Standard Error for a Data Set

Models	κ	δ	Φ	Γ	τ
EGFG	42.5002 (0.01530)	0.2211 (0.0593)	0.6997 (0.0027)	18.6780 (0.0028)	0.1020 (0.3862)
EGF	37.3334 (12.0781)	0.2550 (0.04932)	0.7755 (0.0028)	11.3041 (0.0028)	-
GF	-	1.0242 (10.87)	1.0730 (0.1314)	0.7495 (7.4132)	-
F	-	-	1.0731 (0.1314)	0.7666 (0.1388)	-

Table 2 and 3 give the MLE's along with standard errors, test statistics and corresponding p-values. Figure 3 represent the graphs of estimated pdf's and cdf's for current data. Thus, according to all these graphs and tests, the EGFG model is good fit for this data set.

Table 3
The Goodness-of-Fit Statistics for the Repairing Items

Distribution	AIC	BIC	$\hat{l}$	$W*$	$A*$	KS	p-value
EGFG	**91.3152**	**98.32116**	**-40.6575**	**0.074**	**0.5308**	**0.1177**	**0.8005**
EGF	97.3097	102.9145	-44.6128	0.2621	2.1024	0.1689	0.3584
GF	98.75122	102.9548	-46.3756	0.1830	1.2297	0.1339	0.6559
F	96.75122	99.55361	-46.3756	0.1829	1.2291	0.1339	0.6556

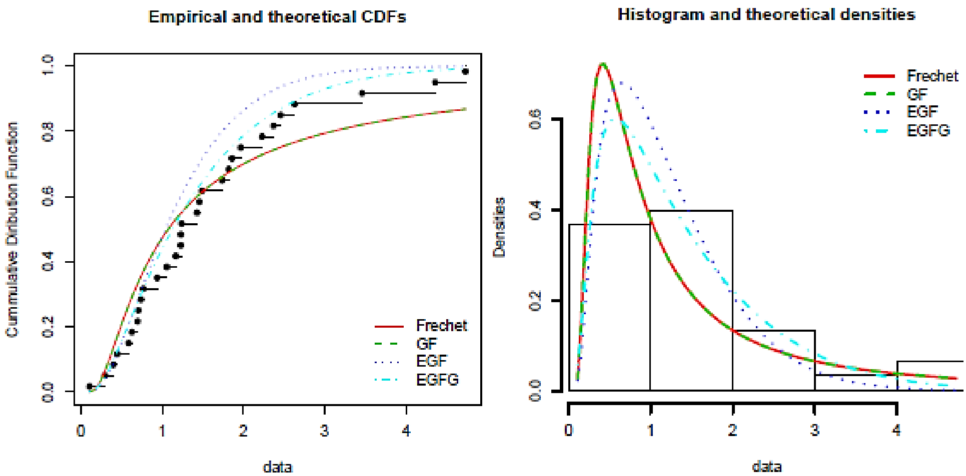


Figure 3: Distributions and Densities of Models for the Time between Failures for Repairable Item

6. CONCLUSION

We introduce a five parameters distribution, named as EGFG distribution which consists four sub-models; the generalized fréchet-geometric (GFG), fréchet-geometric (FG), inverse rayleigh-geometric (IRG) and inverse exponential-geometric (IEG) distribution. We study the shape of the probability function and the failure rate function. The graph shows that the proposed model is unimodal, positively skewed and reverse J- shaped and have increasing and modified upside-down bathtub shaped failure rate for different value of parameters. We calculate some statistical properties such as quantile function, moments and reliability measures. We also derive the close form of density and distribution function. We use the approach of maximum likelihood to estimate the parameters. We conduct the application study to present the flexibility and usefulness of the proposed distribution.

REFERENCES

1. Barreto-Souza, W., de Morais, A.L. and Cordeiro, G.M. (2011). The weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, 81(5), 645-657.
2. Cordeiro, G.M., Ortega, E.M. and da Cunha, D.C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, 11(1), 1-27.
3. da Silva, R.V., de Andrade, T.A., Maciel, D.B., Campos, R.P. and Cordeiro, G.M. (2013). A new lifetime model: The gamma extended fréchet distribution. *Journal of Statistical Theory and Applications*, 12(1), 39-54.
4. Elbatal, I., Altun, E., Afify, A.Z. and Ozel, G. (2018). The generalized burr xii power series distributions with properties and applications. *Annals of Data Science*, 1-27.
5. Elbatal, I. and Muhammed, H.Z. (2014). Exponentiated generalized inverse weibull distribution. *Applied Mathematical Sciences*, 8(81), 3997-4012.
6. Mead, M.E., Afify, A.Z., Hamedani, G. and Ghosh, I. (2017). The beta exponential fréchet distribution with applications. *Austrian Journal of Statistics*, 46(1), 41-63.
7. Murthy, D.P., Xie, M. and Jiang, R. (2004). *Weibull models*, volume 505. John Wiley & Sons.
8. Oguntunde, P.E., Khaleel, M.A., Ahmed, M.T. and Okagbue, H.I. (2019). The gompertz fréchet distribution: Properties and applications. *Cogent Mathematics & Statistics*, 6(1), Article 1568662.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 199-212

A CRITICAL ANALYSIS OF EDUCATION AND EDUCATION SYSTEM: A CASE STUDY IN GC WOMEN UNIVERSITY SIALKOT, PAKISTAN

Zarish Rashid, Mahwish Rabia[§]**, Naima Mubarak** and **Iqra Saleem**
Government College Women University, Sialkot, Pakistan
[§]Corresponding author Email: mahwish.rabia@gcwus.edu.pk

ABSTRACT

This paper attempts to investigate the opinions of students about education and education system. In present study, many aspects are discussed about education system and education. The data is collected using survey method, 71.7% respondents feel that there is gap between rural and urban education system and 75.4% agree that the moral and ethical values are also responsibility of education system. While 62.1% respondents think that there should be one education system in Pakistan and only 43.2% students give opinion that the Pakistan current education system is on high rate and 49.8% agree that the education system is based on technology.

The fitted regression model demonstrates that education would be on high rate in Pakistan if government institutions taught students with updated and new syllabus according to the criteria of the developed countries. This study also shows that 42.2% think that current education produce genius employees. It can be concluded that the current education system has failed to deliver matching response to the call of our national objectives, hopes and aspirations. Finally, it is suggested that Government and public both should work together to improve education and education system and spread awareness about the importance of education.

1. INTRODUCTION

Education is a path used for national development**.** It is used for giving the sense of responsibility among the people. On the other hand, educated people not only realize their duties but also know how to achieve their national, social and individual rights. The goal of education is to give the information to the people about their national and international position as a global citizen. Education promotes economic development (Zaki, 1989).

The position of political stability is low in Pakistan. On the other hand, the country is in deep grip of increasing political turmoil, expanding terrorism, violence, social instability and economic degeneration. All these problems directly or indirectly are the offshoot of a weak education system. Now days, Pakistan faces many problems such as poverty, insecurity and terrorism (Save the children & UNICEF, 2005).

Education is a systematic process of constructing the pillar of society. The literal meaning of education and supplementary definition particularly on the ground of Pakistan poses a critical condition. There are five education levels in Pakistan (i) Primary Level (ii) Middle Level (iii) High Level (iv) Intermediate Level (v) University education. Ahmad et al., (2014) acquired that Pakistan is a developing country. It's economic condition weak. Education is the chief principal in achieving their goals, improve their rules and acquire their political, social and economic condition of country.

Ashraf and Ismat (2016) designed a study to make a comprehensive survey of the causes of failure in achieving better education than other countries. The results of various studies and evaluation reports carried out by the local and foreign agencies were examined to suggest a more effective solution for improving the level of education. Education plays a vital role in the growth and progress of countries in an increasingly competitive, interconnected and globalizing world. It helps to improve living standards and enhance the quality responsibly of life, and can provide essential opportunities for all. Pakistan needs to educate people up to a level where they can understand their own responsibilities. They must be able to earn their own living and contribute to promote the welfare of the society and the country.

Rasool (2007) conducted a study which examines that education provides the facility of reducing poverty and enhancing development. An educational system of poor quality may be one of the most important reasons why poor countries do not grow. In Pakistan, the quality of education has a declining trend and there is acute shortage of teachers. Gerou (1991) studied the main reason of student failure. They defined that there is a situation in which one cannot effort his\her cognitive and emotional imagination about their study and this is supposed to be a major cause of student failure.

Atkinson (2000) examined that the students are responsible for their motivation about their studies and higher need of student's personal involvement in educational process is very important. Laontar (1996) mentioned that there is a soft relationship between teacher and student but most cause of student failure is associated with teacher. The main reason of teacher affiliated student failure is the lack of appreciation, emotionlessness and motivation or bad circumstances of family on behalf of the teacher.

Rashid and Mukhtar (2012) examined that the main part of our body is backbone, so the backbone of our national development is education. There was no need of any planning for education, but now a days the implementation and continuous struggle to produce new planer's, developer's and decision maker's for the national development at any cost has become the main purpose of education. Aziz et al., (2014) examined that there was no planning in Pakistan for vacation's because of limited resources, due to which the students waste their vacations. For getting high level system of better education government must take some serious and hard steps. But due to problems like economy and safety this is not very easy.

2. RESEARCH METHODOLOGY

In this research, the population size of 6,000 students is considered taking BS students, of Government Collage Women University, Sialkot. Using Yamane (1967) formula a sample of 375 using equation (1) is determined from the given population and the procedure is given below:

$$n = \frac{\mathrm{N}}{1+\mathrm{Ne}^2} \tag{1}$$

where N is the total population size and e is the margin of error.

$$n = \frac{6000}{1+6000(0.05)^2} = 375.$$

2.1 Sampling Technique

Considering 6 departments of university as strata the sample is selected using stratified random sampling with proportional allocation. The basic formula of stratified random sampling with proportional allocation is described as:

$$n_h = n \ * \frac{N_h}{N} \tag{3}$$

where "h" denotes number of strata (Cochran, 1977). Using equation (3) required sample size is calculated and shown in Table 1.

Table 1
Sample Size Allocation from each Department

Department	Sample	Department	Sample	Department	Sample
Statistics	58	Computer Science	80	Urdu	62
Botany	43	Political Science	59	Zoology	73

2.2 Analysis of Data

The following statistical techniques are used to analyze the data:

1) Graphs and Frequency Distribution
2) Correlation Coefficient
3) Regression Analysis
4) Mann Whitney U-test (Alternative to t-test)

2.3 Significance of the Study

- To examine the thinking of students about the education system of Pakistan.
- To assess the effects of education system on students for achieving their goals.
- To check the relationship between the technology and rate of education.
- To suggest some points to improve the education and education system.

3. RESULTS AND DISCUSSION

3.1 Reliability Analysis

The reliability of data is as follow:

Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.649	.655	18

In general, 0.7 is good but if it is greater than 0.6 then acceptable. As our reliability value is .649 and it is acceptable. A high value of alpha indicate relationship between items. Adding more relevant items can increase Alpha.

3.2 Histogram of Demographic Variables

Figure 1 shows that 15.5% are statistics department respondents, 11.5% from botany department, 21.3% from CS department, 15.7% from Pol-Sci, 16.5% from Urdu department and 19.5% respondent are from Zoology department.

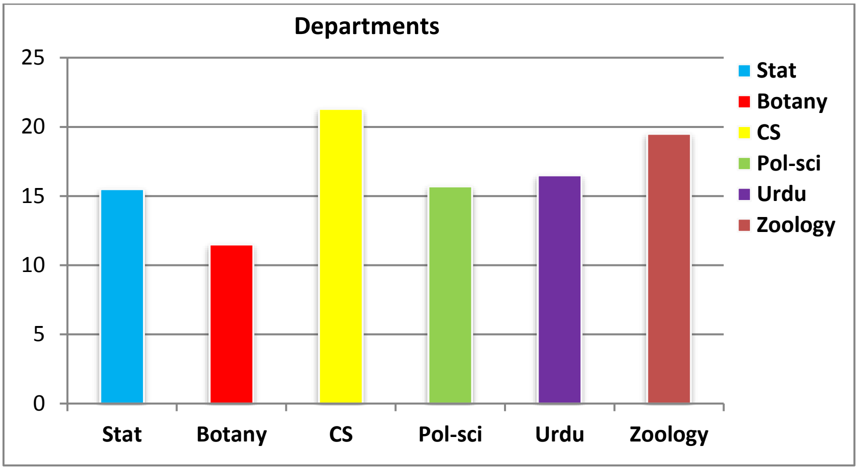


Figure 1: Departments

Figure 2 express that 6.7% students have CGPA between 2.00-2.50, 21.3% students have CGPA between 2.51-3.00, 37.9% respondents have CGPA between 3.01-3.50 and 34.1% respondents have CGPA between 3.51-4.00.

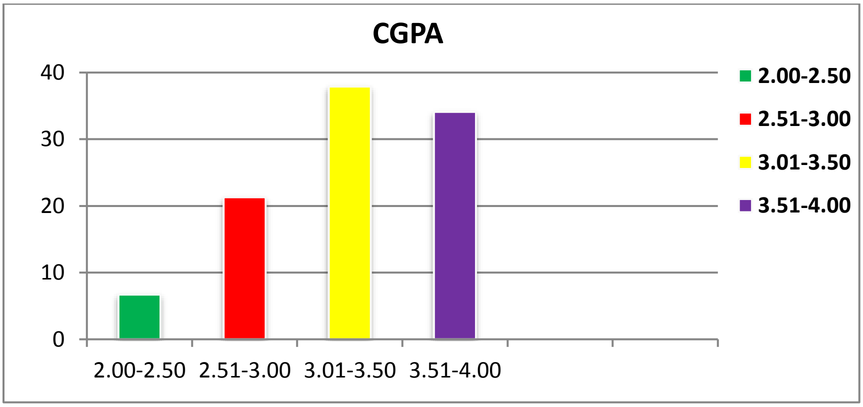


Figure 2: CGPA

Figure 3 demonstrates that 74.1% respondents are of 18-20 age groups, 25.6% of the respondents are of the age of 21-23 years and in age group of 23 or above there are 0.3% respondents.

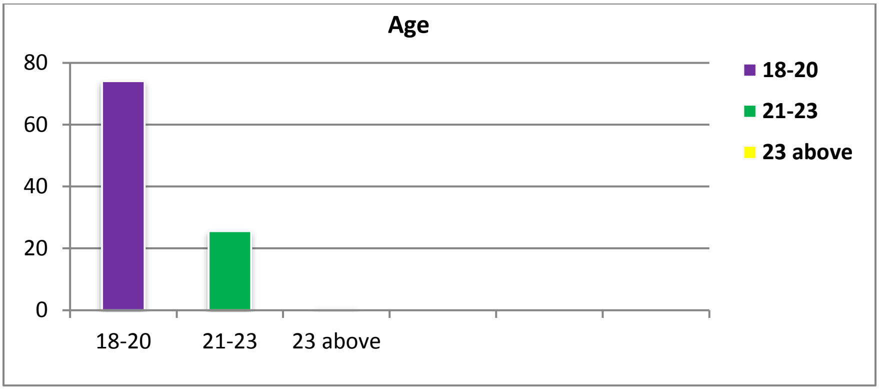


Figure 3: Age

3.3 Descriptive statistic and Discussion

The descriptive statistic like frequencies and percentages are as follows:

Table 1
Present Fee Structure

Variables	Categories	Frequencies (%)
Do you think that the present fee structure is good?	SA (Strongly Agree)	43(11.5)
	A (Agree)	69(18.4)
	N (Neutral)	71(18.9)
	D (Disagree)	91(24.3)
	SD (Strongly disagree)	101(26.9)

Table 1 illustrates that 11.5% respondents are strongly agree, 18.4% agree, 18.9% neutral, 24.3% are disagree and 26.9% strongly disagree that fee structure of present system is not good. The fee is an important factor in education and play a vital role to increase the literacy rate because poor people cannot afford high fees. It is examined that mostly people think that fee structure of present system is not good.

Table 2
Education System is on High Rate

Variables	Categories	Frequencies (%)
Are you satisfied that the Education System is on High Rate in Pakistan?	SA (Strongly Agree)	38(10.1)
	A (Agree)	64(17.1)
	N (Neutral)	86(22.9)
	D (Disagree)	89(23.7)
	SD (Strongly disagree)	98(26.1)

Table 2 show that 10.1% respondents strongly agree, 17.1% agree, 22.9% neutral , 23.7% are disagree and 26.1% are strongly disagree that the present education system is on high rate in Pakistan. A country can be socially and economically strong, if education system would be on high rate. But our country is not developed in this regard because of poor education system.

Table 3
Satisfied with the Syllabus

Variables	Categories	Frequencies (%)
Are you satisfied with the syllabus being taught in government institutions in our country?	SA (Strongly Agree)	51(13.6)
	A (Agree)	72(19.2)
	N (Neutral)	71(18.9)
	D (Disagree)	146(38.9)
	SD (Strongly disagree)	35(9.3)

Table 3 expressed that 13.6% respondents strongly agree, 19.2% agree, 18.9% neutral, 38.9% disagree and 9.3% strongly disagree that the syllabus is satisfied and good in government institutions. The curriculum is the base of education, it enhances the knowledge of students but in current education system syllabus is satisfied but not highly remarkable.

Table 4
One Education System

Variables	Categories	Frequencies (%)
Do you think that there should be one education system in country?	SA (Strongly Agree)	119(31.7)
	A (Agree)	114(30.4)
	N (Neutral)	42(11.2)
	D (Disagree)	61(16.3)
	SD (Strongly disagree)	39(10.4)

The Table 4 represent that 31.7% people strongly agree, 30.4% agree, 11.2% neutral, 16.3% students disagree and 10.4% strongly disagree that there should be one education system. There must be one education system in whole country and it ensures the unity, satisfaction and equality in public.

Table 5
Education System Easily Accessible for Everyone

Variables	Categories	Frequencies (%)
Do you agree that the education is easily accessible for everyone in Pakistan?	SA (Strongly Agree)	48(12.8)
	A (Agree)	83(22.1)
	N (Neutral)	57(15.2)
	D (Disagree)	104(27.7)
	SD (Strongly disagree)	83(22.2)

Table 5 shows that 12.8% respondents are strongly agree, 22.1% agree, 15.2% neutral, 27.7% disagree and 22.2% strongly disagree that in the present system education is not easily accessible for everyone in Pakistan. The results revealed that education is not easily approachable by students due to inequality chances, distance between home and institutions, transport, traffic or other issues.

Table 6
Government is taking any Steps to Improve Education System

Variables	Categories	Frequencies (%)
Do you agree that government is taking any steps to improve education system?	SA (Strongly Agree)	96(25.6)
	A (Agree)	40(10.7)
	N (Neutral)	56(14.9)
	D (Disagree)	147(39.2)
	SD (Strongly disagree)	36(9.6)

Table 6 demonstrate that 25.6% respondents are strongly agree, 10.7% agree, 14.9% neutral, 39.2% disagree and 9.6% strongly disagree that the government is taking steps to improve education system. The government is trying to enhance the education system by different awareness programs, policy making and giving to much importance and emphasize on it but still there need much improvement specially in government schools and colleges.

Table 7
There is Gap between Rural and Urban Education System

Variables	Categories	Frequencies (%)
Do you agree with the statement that there is gap between rural and urban education system in Pakistan?	SA (Strongly Agree)	131(34.9)
	A (Agree)	138 (36.8)
	N (Neutral)	41(10.9)
	D (Disagree)	31(8.3)
	SD (Strongly disagree)	34(9.1)

Table 7 shows that 34.9% people strongly agree, 36.8% agree, 10.9% neutral, 8.3% students disagree and 9.1% strongly disagree that there is gap in rural and urban education system. Currently, in education system there is a gap present between rural and urban institutions in the shape of number of buildings, funds, quality of education and importance etc.

Table 8
Education System is based on Technology

Variables	Categories	Frequencies (%)
Do you agree that the education system is based on technology?	SA (Strongly Agree)	68(18.1)
	A (Agree)	45(12.0)
	N (Neutral)	74(19.7)
	D (Disagree)	69(18.4)
	SD (Strongly disagree)	119(31.7)

Table 8 represented that 18.1% respondents strongly agree, 12.0% agree, 19.7% neutral, 18.4% disagree and 31.7% strongly disagree that the present system is based on technology. The use the audio and video aids enhance the quality of education, in current education system of Pakistan institutions are trying to use technology like mobile phones, projectors and educational software's but still majority of institutions need improvement.

Table 9
Students prefer to use Advanced Technology more than Books

Variables	Categories	Frequencies (%)
Do you agree that the students prefer to use advanced technology more than books in studies?	SA (Strongly Agree)	110(29.3)
	A (Agree)	134(35.7)
	N (Neutral)	42(11.2)
	D (Disagree)	56(14.9)
	SD (Strongly disagree)	33(8.8)

Table 9 shows that 29.3% strongly agree, 35.7% agree, 11.2% neutral, 14.9% disagree and 8.8% strongly disagree that students prefer to use advanced technology more than books in studies. So, from this we can conclude that students not like use books and it's not good. Because, books are in balanced and organized, carefully selected words combination as compared to internet. Therefore, students must readout books to increase their knowledge in better way.

Table 10
Grades Justify the Student's Knowledge

Variables	Categories	Frequencies (%)
Do you think that the grades justify the student's knowledge?	SA (Strongly Agree)	49(13.1)
	A (Agree)	65(17.3)
	N (Neutral)	57(15.2)
	D (Disagree)	91(24.3)
	SD (Strongly disagree)	113(30.1)

Table 10 show that 13.1% strongly agree, 17.3% agree, 15.2% neutral, 24.3% disagree and 30.1% strongly disagree that the grades cannot justify the student’s knowledge. It means that high grades not ensure the capable and outstanding student performance and understanding in real life problems but in current education system students and teachers both are focusing on grades.

Table 11
Current System only Judge the Memory of Student

Variables	Categories	Frequencies (%)
Do you think that the current system only judge the memory of student?	SA (Strongly Agree)	91(24.3)
	A (Agree)	124 (33.1)
	N (Neutral)	65(17.3)
	D (Disagree)	48(12.8)
	SD (Strongly disagree)	47(12.5)

Table 11 expressed that 24.3% strongly agree, 33.1% agree, 17.3% neutral, 12.8% students disagree and 12.3% strongly disagree that system only judge the memory of students. The results show that intelligence and understanding of problem like abilities of a student are not measured in current education system. It is the biggest reason that there is no high development in Pakistan because of poor education structure.

Table 12
Moral and Ethical values are also Responsibility of Education System

Variables	Categories	Frequencies (%)
Do you think that the moral and ethical values are also responsibility of education system?	SA (Strongly Agree)	131(34.9)
	A (Agree)	152(40.5)
	N (Neutral)	42(11.2)
	D (Disagree)	33(8.8)
	SD (Strongly disagree)	17(4.5)

Table 12 demonstrate that 34.9% strongly agree, 40.5% agree, 11.2% neutral, 8.8% students disagree and 4.5% strongly disagree that present system is responsible to teach moral and ethical values. It means that the education system is responsible for moral and ethical values and must teach students how to talk, behave and deal with others.

Table 13
Way of Conveying Lectures of Teacher

Variables	Categories	Frequencies (%)
Do you satisfied with the way of conveying lectures of your teacher?	SA (Strongly Agree)	72(19.2)
	A (Agree)	153(40.8)
	N (Neutral)	60(16.0)
	D (Disagree)	42(11.2)
	SD (Strongly disagree)	48(12.8)

Table 13 shows that 19.2% respondents strongly agree, 40.8% agree, 16.0% neutral, 11.2% students disagree and 12.8% strongly disagree that the teachers conveying lecture way is good. It express that teachers deliver lectures in good way and students understand the lessons.

Table 14
The Current System can Produced the Genius Employees

Variables	Categories	Frequencies (%)
Do you think that the current system can produced the genius employees?	SA (Strongly Agree)	40(10.7)
	A (Agree)	58(15.5)
	N (Neutral)	92(24.5)
	D (Disagree)	118(31.5)
	SD (Strongly disagree)	67(17.9)

Table 14 shows that 10.7% students strongly agree, 15.5% agree, 24.5% neutral, 31.5% students disagree and 17.9% strongly disagree that present system produced genius employees. It is clearly shown from the results that majority of the respondents agrees that current education system do not produce genius employees.

Table 15
Present System is Achieving its Goals

Variables	Categories	Frequencies (%)
Do you think that the present system is achieving its goals?	SA (Strongly Agree)	48(12.8)
	A (Agree)	100(26.7)
	N (Neutral)	92(24.5)
	D (Disagree)	84(22.4)
	SD (Strongly disagree)	51(13.6)

Table 15 show that 12.8% respondents strongly agree, 26.7% agree, 24.5% neutral , 22.4% are disagree and 13.6% are strongly disagree that the present education system is achieving its goals. The system not get its targets like increase literacy rate, free and good quality education for all, remove corruption and inequality etc.

3.4 Inferential Analysis

The process of drawing inferences about population on the basis of sample data is known as inferential statistics.

3.4.1 Correlation Coefficient

The main purpose of the correlation analysis is to measure the strength or degree of linear relationship between the two variables. Though, these two variables are supposed to be random.There is exact linear relationship between the variables; almost all the correlation coefficients have expected positive signs and the correlation between the variables is greatly strong.

Table 16
Correlation between Variables

Variables	Correlation	P-value	Conclusion
Rate of education system and use of technology	0.136*	0.000	Significant
Government should take steps to improve education system and the current education system produce genius employees	0.128*	0.013	Significant

*P-value Significant at 5% level of significance (α =0.05)

Table 16, shows the values of correlation coefficient. When ($r = 0.136$ and p-value = $0.000 < \alpha$), it demonstrates that there is positive correlation between use of technology and rate of education. It determines that to get high rate of education, the use of technology must increase. There should be use of advanced technology to implement educational theories, to share new ideas by social sites. In addition, 3D technology can create environment for practical results in real world and use of projectors can help students to easily understand the problems related to subjects.

The value of the correlation coefficient ($r = 0.128$, $p = 0.013 < \alpha$) shows that there is positive correlation between variables. It demonstrates that if government takes steps to improve educationsystem then in future there will be genius employees in the country, because the intelligent employees are the need of Pakistan and only highly educated and honest employees can solve the economic problems of country.

For the Mann-Whitney test, the hypotheses are

H0: All respondents point of view is same.
H1: All respondents point of view is not same.

Table 17
Mann Whitney Hypothesis Test

Variable	MANN-Whitney Value	P-Value	Conclusion
1. Are you satisfied that the education system is on high rate in Pakistan*Do you think there should be one education system in the country.	64502.500	0.000	Significant
2. Do you agree that education system is based on technology*Do you agree that the student prefer to use advanced technology more than books in studies.	62391.500	0.006	Significant

Table 17 check the significance of variables, between the high rate of education and one education system, Mann Whitney test is applied. The main concern is to check the point of view of students about these variables. From the provided evidence as p-value is $0.000 < \alpha$, result is significant and null hypothesis is rejected and we can conclude that the point of view of respondents about high rate in education and one education system is not same.

From the provided evidence of second hypothesis as p-value is $0.006 < \alpha$, result is significant and null hypothesis is rejected and we can conclude that the point of view of students that education system is based on technology and student's preference to use technology more than books is different.

3.4.2 Model Analysis

The simple linear regression model is defined as:

$$y = \beta_0 + \beta_1 x + e$$

where β_0 is intercept and β_1 slope of the model, Y =education System (High rate, One education system)

X = Teaching Methods (Syllabus: Satisfied, Updated)

Table 18
Simple Linear Regression Model

Attributes	Regular coefficient β	S.E	t-value	Significance (p-value)	R^2
(Constant) Are you satisfied with the syllabus being taught in government institutions in our country?	1.814	.137	13.221	.000	.123
(Dependent).Are you satisfied that the education system is on high rate in Pakistan?	.351	.048	7.247	.000	
R-square=.123, S.E=1.352			F-value = 52.524, p-value = 0.000		
$\hat{y} = 1.841 + .351x$					

In Table 18 the value of $\mathbf{R^2}$ = 0.423, it shows that there is 42.3% variation is explained by the independent variable. The simple linear regression model is based on two variables where dependent variable is education system and teaching methods in government institutions is independent variable.

In addition, both the regression parameters ($\boldsymbol{\beta_0}$= 1.841, $\boldsymbol{\beta_1}$ = 0.351) are significant that is p-value is less than significance value (α = 0.05). The fitted regression model demonstrates that education would be on high rate in Pakistan if government institutions taught students with updated syllabus according to the developed countries.

4. CONCLUSION

The education and education system failures are the main issues in Pakistan, keeping this in view, the main concern of this study is to know the point of view of people about it. The results demonstrate that, the mostly people (50%) think that the present education and education system is not on the high rate. More than half of people feel that there is a gap between rural and urban education system in Pakistan like number of schools, universities, standard of education.

It is also the responsibility of education system to teach moral and ethical values to students. Mostly respondents (50%) think that the education system and education not easily accessible and the fee structure is not good and therefore the poor people are not well educated and there is no free and compulsory education for every citizen available on equal bases. Furthermore, the use of the audio and video aids enhance the quality of education and 50% respondents agrees that in current system there is lack of using advanced technology.

The results of the correlation coefficient r is 0.128 expressed that if government takes steps to improve education system then in future there will be genius employees in the country, because the intelligent employees are the need of Pakistan and only highly educated and honest employees can solve the economic problems of country. The results of simple linear regression model demonstrates that education would be on high rate in Pakistan if government institutions taught students with updated syllabus according to the developed countries because the syllabus is not properly updated and not very highly remarkable. It is concluded that there should be more awareness provided to public about the importance of education and how to apply this education in practical life.

Now, this is the time to realize that there is no progress possible without education. There are more steps required to improve polices and structure of education. The government should pay attention to education and increase funds for education system. It's also our responsibility to work for betterment of our education system, then we can stand our self with western and developed countries.

5. SUGGESTIONS

There are multiple ways present to improve the education system in Pakistan. Education system can be improved through the collective efforts by government, teachers and students. Similarly, for the development of country it is necessary that Government should:

1) Provide the financial and encouragement power to education system
2) Ensure the proper check and balance in the system affairs like every year report etc.
3) Provide more opportunities to use advanced technology and methods
4) Ensure that the system is properly up grated according to the demands of developing world
5) Increase the number of educational institutions specially college and universities in rural areas
6) Student guidance and educational awareness organizations should be build up to promote the education rate.

REFERENCES

1. Zaki, W.M. (1989). *Education of Plans and Projects*, National Book Foundation, Islamabad.
2. Save the Children & UNICEF (2005). Disciplining the child practices and impacts. School & Literacy Department, NWFP.
3. Ahmad, I., Rehman, K., Ali, A., Khan, I. and Khan, A. (2014). Critical Analysis of the Education in Pakistan: Possible Solutions. *International Journal of Evaluation and Research in Education*, 3(2), 79-84.
4. Ashraf, A.M. and Ismat, H.I. (2016). Education Development of Pakistan: A study of Current Situation of Education and Literacy in Pakistan. *US-China Education Review B*, 6(11), 647-654.
5. Rasool, G.M. (2007). Education in Pakistan: The key issues, Problem’s and The New Challenges. *Journal of Management and Social Sciences*, 31(1), 47-55.
6. Gerou, T.D. (1991). *Student failure - Illiteracy - Learning Difficulties: Dyslexia. Athens: Vivliogonia* (in Greek).
7. Atkinson, E. (2000). An investigation into the relationship between teacher motivation and pupil motivation. *Education Psychology*, 20, 45-57.

8. Leontari, A. and Gialamas, B. (1996). Exam stress, self-concept and school perform. *Education Psychology*, 2, 20-39.
9. Rashid, K. and Mukhtar, S. (2012). Education in Pakistan: Problems and their solutions. International *Journal of Academic Research in Business and Social Sciences*, 2(11), 332-343.
10. Aziz, M., Bloom, D.E., Humair, S., Jimenez, E., Rosenberg, L. and Sathar, Z. (2014). Education system reform in Pakistan: why, when, and how? *IZA Policy Papers*, Working Paper No. 76.
11. Yamane, T. (1967). *Statistics: an introductory analysis Harper and row*. New York, Evanston and London and John Weather Hill. Inc., Tokyo.
12. Cochran, W.G. (1977). *Sampling Techniques*. John Wiley and Sons, New York.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 213-224

IRREGULAR MENSTRUAL SIDE EFFECTS ON THE WOMEN'S HEALTH: A CRITICAL ANALYSIS IN GC WOMEN UNIVERSITY SIALKOT, PAKISTAN

Amina Nadeem, Narjis Batool, Mahwish Rabia§**,**
Humma Nawaz and **Saleh Zahid**
Government College Women University, Sialkot, Pakistan
§Corresponding author Email: mahwish.rabia@gcwus.edu.pk

ABSTRACT

This research paper attempts to investigate the irregular menstrual cycle and its effect on the women's health. Mostly women experience different problems during this period like pain, heavy bleeding, and mood disruption. The research is conducted through survey method and data is collected using questionnaire. The result revealed that 38% people gain weight due to anomaly of periods and 72% respondents feel pelvic pain during periods. Moreover majority of respondents give opinion that they face unbalanced menstrual because of stress and eating disorders. The value of chi-square ($\chi^2 = 18.721$) demonstrates that hormonal change effect irregularity of periods. The value of odd ratio depicts that probability of pelvic pain increases when the periods are irregular. It can be conclude that the reasons of irregularity of menstrual cycle are the obesity, stress, eating disorder, excessive exercise, thyroid, hormonal influence, dieting, contraceptive pills etc. The result of such disease is highly dangerous for women's that includes heart disease, chronic disease, infertility, facial hairs. The prolonged imbalance of menstrual cycle leads to the extremism of depressed condition. The Maintenance of healthy weight, exercise, healthy diet, proper medication can cure the problems.

KEYWORDS

Menstrual cycle, women experience, unbalanced menstrual, demonstrates.

1. INTRODUCTION

Periods effects on health are a common issue of women's life now a day. Usually, women have to face problems of period's effect on health which mostly occur at the beginning period or at the time of menopause. It is because of various reason such as stress, low and high body weight, chronic disease, thyroid issue, obesity, contraceptive pills, age, extreme exercise or dieting, Polycystic Ovary Syndrome (PCOS), some other illness etc. In many cases irregular periods are also related to an ovulation or subtle hormone imbalances. Irregular menstrual is not ordinary disease. It is really a serious disease which needs our attentions. It can even create serious problems and can also be a difficulty in getting pregnant. So, we should not take it easy. We should consult a doctor and should have our proper medical checkup. We should also be screened for thyroid disorders and other disorders that many link to irregular menstrual. Most common causes are as follows, stress, obesity, dieting, extreme exercise, contraceptive pills. A health care provider diagnoses period's irregularities using a combination of the following healthy diet, yoga, exercise, medication etc.

Williams and Creighton, [1] explained that menstrual disorders are common in adolescent girls. Periods can be irregular, heavy and painful, especially in the first few years. Serious physician is rare; however, irregular menstrual cycle can have significant effect on daily activities and school absence. The study examines many treatment options which are safe to use in youngsters, although the evidence for their use is extrapolated from adult data. This study presents a clinical review of the current practice, including management of girls with other medical problems and learning difficulties. Kimberly Fitzgerald [2] said that physical and emotional symptoms relating to the menstrual cycle can have different effects for different women's. For women's who experience major symptoms, in other words pain, heavy bleeding and mood disruptions. Health specialist may need to make special medication when working with women who experience irregular periods. Menstrual cycle function may be effected by many factors such as hormones, exercise, body weight, perceived stress, life events and work environment. Menstrual cycle function are often studied because of the impact on fertility. Yeunhee et al., [3] explained that irregular menstruation is an important indicator of current and potential health problem. A women health is greatly influenced by her socioeconomic status. The purpose of this study to examine the prevalence of irregular periods by socioeconomic status among South Korean women. In this study the women's of South Korea are target population from aged 19-54 years, using raw data from the Korea National Health and Nutrition Examination Survey V (2010-1012), a nationally representative survey. Compare to women who graduated from university, the adjusted odd ratios (95% confidence interval) for those who graduated from elementary school or lower, middle school, and high school. Compared to women with a medium-high income level, the adjusted odd ratio (95% confidence interval) for women with the highest household income level. Irregular menstruation was prevalent among adult women and appeared to be associated with socioeconomic status, especially in terms of education and household income. Strine et al., [4] examine the associations of menstrual-related problems with mental health and health behaviors in a U.S. population-based study. In this study the women of U.S. are target population from aged 18-55 years who participated in the 2002 National Health Interview Survey, an ongoing, computer-assisted personal interview of the non-institutionalized U.S. population. In this study 19% of women aged 18-55 years reported experiencing menstrual-related problem (e.g., heavy bleeding, bothersome cramping, or premenstrual syndrome. These women's were significantly more likely than those without menstrual-related problems to report frequent anxiety and depression, insomnia, excessive sleepiness, and pain over the past 12 months. Women's with menstrual-related problems were also significantly more likely to report feeling sad, nervous, restless, hopeless, or worthless and that everything was an effort all or most of the time during the past 30 days. Bae et al., [5] Said that a regular menstrual cycle is an important indicator of a healthy reproductive system. Previous studies reported obesity, stress, and smoking as the factors that are associated with irregular menstruation and early menopause. However, the integrable effect of these changeable risk factors have not been fully understood. This study aimed to investigate the changeable risk factors of menstrual cycle irregularity and premature menopause, as well as their individual and combined effects among adult women in Korea. Yamamoto et al., [6] examine the relationship between menses-associated health problems of women, in other words premenstrual symptoms, period's pain and irregular menstrual cycles, and psychosocial stress. A cross sectional-study was conducted among Japanese college students, measuring psychosocial stress levels by means of IMPS (The Inventory to Measure Psychosocial Stress). A sample of 264 questionnaires was used in which 221 response were collected. The proportions of students who reported premenstrual symptoms, menstrual pain, and the experience of menstrual cycles were 79%, 79%, and 63% respectively.

2. SIGNIFICANCE OF THE STUDY

- To examine factors behind irregularity of periods.
- To observe the side effects or diseases due to irregularity of periods.
- To examine the relationship between ages and irregularity of periods.
- To find out preference among females about treatment.

3. RESEARCH METHODOLOGY

3.1 Purpose of Study

To obtain a deeper understanding of women lived experience, we choose survey method, as it is a data collection tool used to gather information about individuals. We collect the data information from the girls 15-30 years of age. We collect information about the irregular menstrual side effect on women health from the girls of science departments of GC women university of Sialkot.

3.2 Population

All the student of science departments of G.C. Women University Sialkot.

3.2.1 Sampling technique

The data information is collected by using the simple random sampling.

3.2.2 Statistical tools

To observe the effect of irregular menstrual on women's health for this we use the reliability, graphical representation chi-square and odd ratio to analyze the data.

3.3 Determination of sample size

In this research, the population of size is considered taking BS (Hons.) students of science departments. The representative part of the population is called sample. Using the Yamane [7] formula given in equation (1) where "N" is total population size 1663 and "e" is the margin of error, a sample of 323 is determined from the given population. The procedure is given below:

$$n = \frac{N}{1+Ne^2} \tag{1}$$

$$n = \frac{1663}{1 + 1663(0.05)^2} = 323$$

3.4 Statistical Methods

Reliability Analysis
Graphical representation
Chi-Square Test
Odd Ratio.

4. RESULTS AND INTERPRETATION

4.1. Reliability Analysis

"Reliability" is how well a test measures what it should. Table 1 represent the reliability analysis. A value of Cronbach's Alpha will which greater than 0.7 is acceptable. In the present study the value of Cronbach's Alpha is 0.634 which is approximately near the true value.

Table 1
Reliability Analysis

Cronbach's Alpha	Number of Items
0.634	20

4.2 Graphical Representation

Figure 1-3 shows the graphical representation of different variables of menstrual cycle.

Periods Begins Age

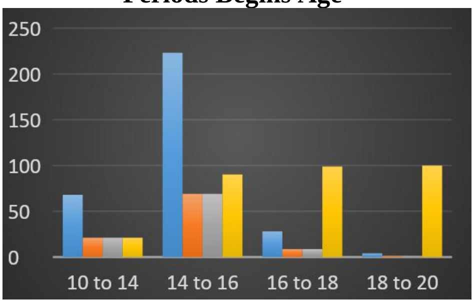


Figure 1

Weights due to Irregularity of Periods

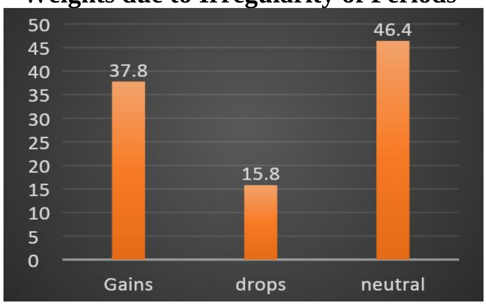


Figure 2

Pelvic Pain during Periods

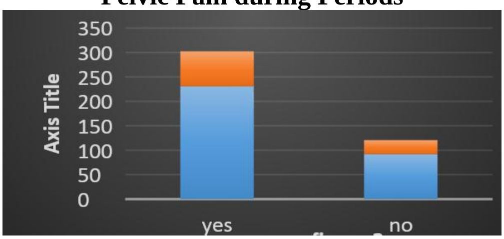


Figure 3

The Figure 1 reveals that 21.1% girls begin periods in age of 10-14 years and 69% girls are begin periods girls 14-16 years and the 8.7% girls are begin periods 16-18 years and only 1.2% girls are begin periods girls 18- 20 years. According the results it is illustrate that girls having periods in 14-16 years of age. Moreover, Figure 2 illustrate that 37.8% girl's weight gain due to irregularity of periods, 15.8% drop the weight and 46.4% girls have neutral weight. According the results it is examined that girls have neutral weight due to irregularity of periods. In addition Figure 3 shows that the 71.5% girls have pelvic pain during periods and the other 28.5% girls have no pelvic pain during periods. The results shown that girls face problem of pelvic pain during periods because of irregular periods.

Table 2
Periods Irregular

Variable	Categories	Number	Percent
Periods Irregular	Yes	281	87.0
	No	42	13.0

Table 2 represent that 87% girls have irregular periods and the other 13% girls have regular periods. The results examined that mostly girls face problem of irregularity of periods.

Table 3
Causes of Irregular Periods

Variable	Categories	Number	Percent
Causes of Irregular Periods	Thyroid	17	5.3
	Stress	239	74.0
	Chronic Disease	18	5.6
	Obesity	49	15.2

Table 3 demonstrates that the cause of irregularity of periods are 5.3% thyroid, 74% stress, 5.6% chronic disease and 15.2% obesity. It is investing that majority of girls face problem of irregularity of menstrual cycle because of stress.

Table 4
Mental Status during Periods

Variable	Categories	Number	Percent
Mental Status During Periods	Relaxing	15	4.6
	Conscious	150	46.4
	Normal	105	32.5
	Other	53	16.4

Table 4 shows the mental status during periods. It clearly depicts that 4.6% respondents feel relaxed, 46.6% conscious, 32.5% normal and 16.4% have other mental status during periods. The results examined that majority of girls have conscious mental status during periods because of irregularity of periods.

Table 5
Polycystic Ovary Syndrome (PCOS) Disease Highly Results in

Variable	Categories	Number	Percent
Polycystic Ovary Syndrome (PCOS) Disease Highly Results in	Infertility	174	55.1
	Chronic disease	24	7.4
	Heart disease	13	4.0
	Facial hair	108	33.4

Table 5 demonstrates that results of polycystic ovary syndrome (PCOS) disease are 55.1% in infertility, 7.4% in chronic disease, 4% in heart disease and 33.4% in facial hairs. According to the results it is illuminated that major part of girls face the problem of infertility because of polycystic ovary syndrome (PCOS) disease.

Table 6
Eating Disorder Harm for Periods

Variable	Categories	Number	Percent
Eating Disorder Harm for Periods	Fats and oil	81	25.1
	Junk foods	173	53.6
	Over eating	9	2.8
	Other	60	18.6

Table 6 reveals that eating disorder like fats and oil, junk food, overeating are harmful for periods. According to results, 25.1% respondent's opinion is fats and oil are harm for the periods, 53.6% belief is junk foods harms for the periods, 2.8% opinion is over eating while 18.6% opinion is other things harms for periods. It is examined that majority of people face irregular periods problem because of junk food.

Table 7
Reasons behind Irregularity of Menstrual Cycle

Variable	Categories	Number	Percent
Reasons Behind Irregularity of Menstrual Cycle	Extreme exercise	15	4.6
	Dieting	94	29.1
	Contraceptive pills	106	32.8
	Others	108	33.4

Table 7 illustrate that the reasons behind the irregularity of menstrual cycle are extreme exercise, dieting, contraceptive pills and others. The results examined that 4.6% irregularity are due to extreme exercises, 29.1% due to dieting, 32.8% due to contraceptive pills and 33.4% due to other things. According to the results it is check out that girls face irregular menstrual problem because of dieting and contraceptive pills.

Table 8
Healthy Menstrual Cycle Range

Variable	Categories	Number	Percent
Healthy Menstrual Cycle Range	21 to 35 days	197	61.0
	19 to 40 days	38	11.8
	20 to 38 days	39	12.1
	25 to 45 days	49	15.2

Table 8 shows that the healthy menstrual cycle range is 61% 21 to 35 days, 11.8% 19 to 40 days, 12.1% 20 to 38 days and 15.2% 25 to 45 days. According the results a healthy menstrual cycle range is between 21 to 35 days. According the survey a healthy menstrual range is under 21 to 35 days and if your periods days out of range that means your periods to be irregular.

Table 9
Flow Experience during Periods is Usually

Variable	Categories	Number	Percent
Flow Experience During Periods is Usually	Heavy	207	64.1
	Moderate	45	13.9
	Light	33	10.2
	Can’t remember	38	11.8

Table 9 demonstrate that your flow experience during periods is usually 64.1% heavy, 13.9% moderate, 10.2% light and 11.8% girls can’t remember the flow experience during periods. According the results shows that girl’s experience of flow during irregular periods is usually heavy.

Table 10
Hormonal Change Effect Irregularity of Periods

Variable	Categories	Number	Percent
Hormonal Change Effect Irregularity of Periods	Yes	295	91.3
	No	28	8.7

Table 10 illustrate that 91.3% hormonal change effects the irregularity of periods and 8.7% hormonal change could not effects the irregularity of periods. According the result it is examined that girls face the problem of irregularity of periods because of hormonal change.

Table 11
Bleeding Days Usually have each Period

Variable	Categories	Number	Percent
Bleeding Days Usually have each Period	2 to 3 days	48	14.9
	3 to 4 days	77	23.8
	4 to 5 days	76	23.5
	5 to 6 days	122	37.8

Table 11 reveals that the bleeding days of each period's usually 14.9% 2 to 3 days, 23.8% 3 to 4 days, 23.5% 4 to 5 days and 37.8% 5 to 6 days. According the results is depicts that majority of girls have 5 to 6 bleeding days usually in each periods when periods are irregular.

Table 12
Remedies can Return Menstrual Cycle to Normal

Variable	Categories	Number	Percent
Remedies can Return Menstrual Cycle to Normal	Exercising	15	4.6
	Medication	36	11.1
	Healthy diet	246	76.2
	Other	26	8.0

Table 12 reveals that remedies can return menstrual cycle to normal 4.6% exercising, 11.1% medication, 76.2% healthy diet and 8% other. According the results it is explained that healthy diet can return menstrual cycle to normal.

Table 13
Kind of Pain Killer you Prefer during Periods

Variable	Categories	Number	Percent
Kind of Pain Killer you Prefer During Periods	As prescribed by doctor	70	21.7
	Without prescription	183	56.7
	None	70	21.7
	Other	26	8.0

The Table 13 illustrate that kind of pain killer you prefer during periods 21.7% as prescribed by doctor, 56.7% without prescription and 21.7% girls could not use any kind of pain killer during periods.

Table 14
Use of Painkillers during Periods

Variable	Categories	Number	Percent
Use of Painkillers During Periods	Yes	235	72.8
	No	88	27.2

Table 14 shows that 72.8% girls use pain killer during periods and 27.2% girls could not use pain killer during periods. According the results it is examined that mostly girls have irregular periods because of use of painkillers during periods.

Table 15
After Medicine Effect during Periods

Variable	Categories	Number	Percent
After Medicine Effect During Periods	Yes	212	65.6
	No	111	34.4

The Table 15 shows that 65.6% girl's effect have after medicine during periods and 34.4% girls have no effect after medicine during periods**.** According the results it is illustrate that majority of girls face problems of irregularity of menstrual because of use of medicine during periods.

Table 16
Which Medicine you use During Periods

Variable	Categories	Number	Percent
Which Medicine you use During Periods	Panadol	127	39.3
	Ponstan	53	16.4
	Brufen	30	9.3
	Other	113	35.0

Table 16. Reveals that 39.3% girls Panadol use during periods, 16.4% Ponstan, 9.3% Brufen and 35% use other medicines during periods. According the results it is depicts that mostly girls use Panadol during periods.

Table 17
Medicine Lead to Irregularity of Periods

Variable	Categories	Number	Percent
Medicine lead to Irregularity of Periods	Yes	236	73.1
	No	87	26.9

Table 17 reveals that 73.1% medicines lead to irregularity of periods and 26.9% medicines could not lead to irregularity of periods. According the results it is explained that the use of medicine lead to irregularity of menstrual.

Table 18
You Suggest others for Using Medicines

Variable	Categories	Number	Percent
You Suggest others for using Medicines	Yes	50	15.5
	No	273	84.5

Table 18 demonstrate that the 15.5% girls suggest the others for using medicines during periods and 84.5% girls are not suggest to others for using medicines during periods. According the results it is examined that girls are not suggest others for using medicine because of use of medicine during periods effect the menstrual to be irregular.

4.2.1 Chi-Square Test

H_0= There is no association between variables
H_1= There is association between variables

Table 19
Relationship between Variables through Chi-Square

Attributes	Chi-Square	p-value	Conclusion
Periods are irregular and hormonal change effect irregularity of periods	18.721	0.000	Significant
Bleeding days and flow experience during periods	19.049	0.025	significant
Pelvic pain and use of painkillers during periods	6.301	0.012	Significant
Use of painkillers and effect you have after medicine	18.266	0.000	Significant
You think that medicines lead to irregularity of periods and you suggest others for using medicines	18.885	0.000	Significant

To check the relationship between different attributes chi-square is calculated. In Table 19 the value of chi-square $(\chi^2 = 18.721, p - value = 0.000)$ shows the significant relationship between the periods are regular and hormonal change effect the irregularity of periods. The value of chi-square examined that hormonal change effect periods and periods to be irregular. The value of chi-square $(\chi^2 = 19.049, p - value = 0.025)$ shows the significant relationship between the bleeding days and flow experience usually have each periods. The value of chi-square $(\chi^2 = 6.301, p - value = 0.012)$ shows the significant relationship between the pelvic pain during periods and use of painkiller during periods. The value of chi-square $(\chi^2 = 18.266, p - value = 0.000)$ shows the significant relationship between the use of medicine and effect you have after using medicine. The value of chi-square illustrate that effect you have after medicine. The value of chi-square $(\chi^2 = 18.885, p - value = 0.000)$ shows the significant relationship between medicines lead to irregularity of periods and suggest others for using medicines. The value of chi-square demonstrate that could not suggest others for using medicines.

4.1.2 Odd Ratio

Table 20
Odd Ratio between Periods Irregular and Pelvic Pain during Periods

		Pelvic Pain During Periods		Total
		Yes	**No**	
Periods Irregular	**Yes**	**203**	78	281
	No	28	14	42
	Total	**231**	**92**	**323**
Odd Ratio = 1.301				

Table 20 illustrates the value of odd ratio=1.301 for pelvic pain and irregular periods. The results examined that the risk factor of pelvic pain during irregular periods is more likely to occur than regular periods as the value of odd ratio is greater than 1. It is also examined that the value 203 in table 20 shows that girls with irregular periods are suffering more from pelvic pain.

Table 21
Odd Ratio between Periods Irregular and use of Pain Killer during Periods

		Use of painkiller during periods		Total
		Yes	**No**	
Periods Irregular	**Yes**	**206**	75	281
	No	29	13	42
	Total	235	88	323
Odd Ratio = 1.231				

Table 21 reveals that the value of odd ratio is 1.231 for use of painkiller and irregular periods. It is clearly shown by the result that use of painkillers during periods increase the risk factor for irregularity of periods. According the contingency table the value of 206 identifies that girls who use painkillers during periods have more chance of irregularity of periods.

Table 22
Odd Ratio between Periods Irregular and Medicine Lead to Irregularity of Periods

		Medicine Lead to Irregularity of Periods		Total
		Yes	**No**	
Periods Irregular	**Yes**	**206**	75	281
	No	30	12	42
	Total	235	236	87
Odd Ratio = 1.099				

Table 22 demonstrates that the value of odd ratio (OR=1.231) is examined for use of excessive medicine which leads to irregularity of periods and respondents having irregular menstrual cycle. The results explained that using different medicines during periods lead to irregularity of periods. Since the value of odd ratio is greater than 1 which clearly depicts that use of medicines increase the risk factor of irregular periods. The value of frequency 206 in contingency table, examined that that girls are suffering more from irregularity of periods by using medicines.

5. CONCLUSION

The purpose of the study is to examine the side-effects of irregularity of menstrual cycle in women. The survey method is used to collect data from the science students of GCWUS by using simple random sampling. It is examined that the irregularity of menstrual cycle has a major impact on women’s health, mostly women faces different problems such as heavy bleeding, pelvic pain and any other problems, where by 38% women’s gain their weight due to anomaly of menstrual cycle and 72% feel pelvic pain & 46.4% of the women’s feel conscious during these days, most of the women feel that they have menstrual disorder due to stress and they have facial hair. 55.1% infertility rate shows that in PCOS disease most common disease is infertility, that’s most commonly diagnose in irregularity of menstrual cycle and is most sensitive issue for women.

Furthermore, 74% women have stress due to disorder of periods and 54% of women have menstrual disorder due to Junk food. The value of chi-square (χ^2=18.721) demonstrates that hormonal change affect irregularity of periods. The value of odd ratio says that the probability of pelvic pain increases when periods are irregular. It can be conclude that the reasons of irregularity of menstrual cycle are the obesity, thyroid, eating disorder, excessive exercise, stress, hormonal influence, dieting, contraceptive pills etc. The results of these types of diseases very dangerous for women, they face infertility, chronic diseases, depression, facial hair and obesity. More than half of the students had their academic performance affected due to stress & depression. The university administration and counseling services should be aware about the results of the study and should provide further facilities and awareness to female students.

SUGGESTIONS

In the light of this study we can say that nowadays irregularity of menstrual cycle is very common issue. All the results of this study show that most of the women's faces different problem due to menstrual disorder. To reduce the ratio of this problem , take healthy food and stop eating junk food, maintain your diet plan and do exercise & yoga daily. If all women's follows this pattern then they will sort out this issue.

REFERENCES

1. Williams, C.E. and Creighton, S.M. (2012). Menstrual disorders in adolescents: review of current practice. *Hormone research in Paediatrics*, 78(3), 135-143.
2. Fitzgerald, K. (2013). Menstrual cycle and workplace issues: review of the literature. *Psique*, 9, 31-55.
3. Kwak, Y., Kim, Y. and Baek, A.K. (2019). Prevalence of irregular menstruation according to socioeconomic status: A population-based nationwide cross-sectional study, *PLoS ONE*, 14(3), e0214071.
4. Strine, T.W., Chapman, D.P. and Ahluwalia, B.I. (2005). Menstrual-related problems and psychological distress among women in the United States. *Journal of Women's Health*, 14(4), 316-323.
5. Bae, J., Park, S. and Kwon, J.W. (2018). Factors associated with menstrual cycle irregularity and menopause. *BMC Women's Health*, 18(1), 36.
6. Yamamoto, K., Okazaki, A., Sakamoto, Y. and Funatsu, M. (2009). The relationship between premenstrual symptoms, menstrual pain, irregular menstrual cycle, and psychosocial stress among Japanese college students. *Journal of Physiological Anthropology*, 28(3), 129-136.
7. Yamane, T. (1967). *Statistics: Introductory Analysis*. Harper and Row, New York.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 225-237

BAYESIAN ANALYSIS OF EXPONENTIATED KUMARASWAMY DISTRIBUTION USING INFORMATIVE AND NON- INFORMATIVE PRIORS

Razia Bibi, Nida Nureen, Naima Mubarak and **Wajiha Nasir**
Department of Statistics, GC Women University, Sialkot, Pakistan
Email: raziacheema56@gmail.com
naima.mubarak@gcwus.edu.pk
wajiha.nasir19@gmail.com

ABSTRACT

In this paper, we estimate a scale parameter of the Exponentiated Kumaraswamy distribution by considering the Bayesian approach under two informative and non-informative priors using four different loss functions. We derive the corresponding posterior distributions for the scale parameter of the Exponentiated Kumaraswamy distribution assuming that the other three parameters are known. The Bayes estimators and associated posterior risks have also been derived using the four different loss functions. The performance of the Bayes estimators are evaluated and compared using a simulation study.

KEY WORDS

Exponentiated Kumaraswamy Distribution, Uniform prior, Jeffrey's prior, Gamma prior, Exponential prior, Loss functions, Efficiency.

1. INTRODUCTION

Kumaraswamy introduced Exponentiated Kumaraswamy distribution for double bounded random processes which has considerable attention in hydrology and related areas. This distribution is applicable to many natural phenomena whose outcomes have lower and upper bounds, such as the heights of individuals, scores obtained on a test, atmospheric temperatures, hydrological data, etc. The probability distribution function and the cumulative distribution function of Kumaraswamy distribution is given as

$$f\left(x;\lambda,\eta,\theta\right)=\lambda\theta\eta x^{\lambda-1}\left(1-x^{\lambda}\right)^{\eta-1}\left[1-\left(1-x^{\lambda}\right)^{\eta}\right]^{\theta-1}\quad \lambda,\eta,\theta>0, x>0 \tag{1}$$

where λ, η and beta are shape parameters and θ is scale parameter.

Marcelino et al. (2011) estimated the model parameters of Kumaraswamy generalized gamma distribution by method of maximum likelihood and a Bayesian procedure had adopted. They used the simulation method of Markov Chain Monte Carlo (MCMC). They had no prior information from historical data or from previous experiment, they assigned conjugate but weakly informative prior distributions to the parameters. Sultan and Ahmad (2015) studied the importance of Bayesian approximation techniques. They observed that the large sample distribution could be improved when prior is taken into account. In all cases normal approximation, Lindley's approximation, T-K approximation, Bayesian estimates under informative priors would be better than those under non-informative priors especially the Inverse levy distribution proves to be efficient with minimum posterior standard deviation and mean square error in case of Lindley's approximation. In case of non-informative priors modified Jeffrey's prior proved to be efficient. They observed that under informative as well as non- informative priors, the normal approximation behaved well than T-K approximation, although the posterior variances in case of T-K approximation were very close to that of normal approximation. Further they concluded that the posterior standard deviation based on different priors tends to decrease with the increase in sample size. Samia A. Adham and Anfal A. ALgfary (2016) used Bayesian estimation of the vector of parameters of the finite mixture of two-component exponentiated Kumaraswamy distribution. The MEKum distribution is studied under the squared error loss function and conjugate priors. R. M. Mandouh (2016) computed approximate Bayes estimates under type II censored samples using the Gibbs sampling procedure which generates samples from the posterior distributions. The approximate Bayes estimators are obtained under the assumptions of non- informative priors. The computation achieved by using the WinBUGS software. A. M. Abd AL-Fattah (2017) estimated parameters of inverted Kumaraswamy distribution by maximum likelihood estimation. The Bayesian approach is considered, under squared error (SE) loss function to estimate the parameters of the inverted Kumaraswamy distribution based on Type II censored samples, using the non-informative prior. The relative absolute biases (RABs), relative mean square errors (RMSEs), variances and estimated risks (ERs) of ML and Bayes estimates of the shape parameters also estimated. And observed that the RABs, variances and RMSEs of the MLEs of the shape parameters decrease when the sample size n increases. Ghosh and Nadarajah (2017) discussed Bayesian estimation of Kumaraswamy distributions based on three different types of censored samples. They obtained Bayes estimates of the model parameters using two different types of loss functions. LINEX and Quadratic Loss Functions were used. The QLF had been found to be the most efficient loss function under each censoring scheme. Similarly, the estimates under single type (II) censoring scheme were found to be associated with the minimum amount of risks. Abbas Pak et al. (2018) estimates the parameters of Kumaraswamy distribution. They had provided likelihood function based on type II hybrid censoring sample and obtained maximum likelihood estimate of the parameters. They had computed the estimate of the unknown parameters by using Tierney and Kadane's approximation under the assumption of both non-informative and informative gamma priors. They had further constructed approximate confidence interval and bootstrap confidence interval of the parameters. The performances of the different methods had been compared gooby Monte Carlo simulations. Based on the results of the simulation study, it had seen clearly that, the Bayesian procedure based on non-informative prior and the ML procedure give similar estimation results. However, using the ML method, approximate confidence interval of the parameters was obtained. On the other hand, they suggested that Bayes estimates of the parameters based on informative priors give better performances than the MLEs in terms of minimum MSEs. As a result, when information about the hyperparameters was available, they suggested Bayesian approach to estimate the parameters of Kumaraswamy distribution. They suggested that the performances of all estimators were improved when the effective sample size had increased. Naser A. Abou-Elheggag et al. (2018) used maximum likelihood method, Bayesian method relative to symmetric and asymmetric loss functions and parametric bootstrap percentile method. They used distribution of the MLE of R to construct exact confidence interval of R they observed that the MSE's of the Bayes estimates of R were smaller than MSE's of the ML estimates. For all methods when n and m increase, the MSE's would be reduce. M. K. Rastogi and P. E. Oguntunde (2018) estimated Kumaraswamy inverse exponential distribution by maximum likelihood estimation method. The symmetric and asymmetric loss functions were used for the Bayesian estimations. Random samples were generated from the posterior distributions using the Metropolis Hastings algorithm procedure and the Bayes estimators were obtained. Comparison was made between the Bayes estimators and the maximum likelihood estimators using Monte Carlo simulations. And concluded that Bayesian estimation is more accurate than the maximum likelihood estimation.

2. POSTERIOR DISTRIBUTION

The posterior distributions are derived by using two non-informative and informative priors. Bayes estimates and posterior risk are calculated by using simulation techniques at different sample size.

2.1 Posterior Distribution Under Non-Informative Prior

2.1.1 Uniform Prior

Uniform prior for parameter θ is:

$$p(\theta) \propto 1, \quad 0 < \theta < \infty \tag{2}$$

The posterior distribution, using uniform prior is:

$$p\left(\theta|x\right) = \frac{\left(-\log \sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right)^{n+1} \theta^{n+1-1} e^{-\theta\left[-\log \sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right]}}{\Gamma\left(n+1\right)} \tag{3}$$

and so $p\left(\theta|x\right) \sim G\left(\alpha,\beta\right)$, where $\alpha = n+1$ and $\beta = -log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}$.

2.1.2 Jeffrey's Prior

Jeffery's prior for parameter θ is:

$$p(\theta) \propto \theta^{-1}, \quad 0 < \theta < \infty \tag{4}$$

The posterior distribution, using Jeffrey's prior is:

$$p\left(\theta|x\right) = \frac{\left(-\log \sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right)^{n} \theta^{n-1} e^{-\theta\left[-\log \sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right]}}{\Gamma\left(n\right)} \tag{5}$$

and so $p\left(\theta|x\right) \sim G\left(\alpha,\beta\right)$, where $\alpha = n+1$ and $\beta = [-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}$

2.2 Posterior Distribution Under Informative Prior

2.2.1 Exponential Prior

Exponential prior for parameter θ is:

$$p(\theta) \propto \frac{1}{b} e^{-\frac{\theta}{b}}, \ 0 < \theta < \infty \tag{6}$$

The posterior distribution, using exponential prior is:

$$p\left(\theta \middle| x\right) = \frac{\left(\frac{1}{b} - \log \sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right)^{n+1} \theta^{n+1-1} e^{-\theta\left[\frac{1}{b}-\log\sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right]}}{\Gamma\left(n+1\right)} \tag{7}$$

and so $p\left(\theta \middle| x\right) \sim G\left(\alpha,\beta\right)$, where $\alpha = n+1$ and $\beta = \left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]$

2.2.2 Gamma Prior

Gamma prior for parameter θ is:

$$p(\theta) = \frac{b^{\alpha}}{\Gamma\left(\alpha\right)} \theta^{a-1} e^{-\theta b}, \ 0 < \theta < \infty \ (8)$$

The posterior distribution, using gamma prior is:

$$p\left(\theta \middle| x\right) = \frac{\left(b - \log \sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right)^{a+n} \theta^{a+n-1} e^{-\theta\left[b-\log\sum_{i=1}^{n}\left(1-\left(1-x_i^{\lambda}\right)^{\eta}\right)\right]}}{\Gamma\left(a+n\right)} \tag{9}$$

and so $p\left(\theta \middle| x\right) \sim G\left(\alpha,\beta\right)$, where $\alpha = a+n$ and $\beta = \left[b - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]$.

3. BAYES ESTIMATOR AND POSTERIOR RISK UNDER UNIFORM PRIOR

The Bayes estimator using SELF is

$$\hat{\theta}_{SELF} = \frac{n+1}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using SELF is

$$\hat{\theta}_{SELF} = \frac{n+1}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Bayes estimator using PLF is

$$\hat{\theta}_{PLF} = \frac{\sqrt{\left(n+1\right)\left(n+2\right)}}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using PLF is

$$\hat{\theta}_{PLF} = \left[\frac{\sqrt{\left(n+1\right)\left(n+2\right)}-n-1}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}\right]$$

The Bayes estimator using SPLF is

$$\hat{\theta}_{SPLF} = \frac{n\left(n+1\right)}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Posterior risk using SPLF is

$$\hat{\theta}_{SPLF} = 2*\sqrt{\frac{n+1}{n}}-1$$

The Bayes estimator using DELF is

$$\hat{\theta}_{DELF} = \frac{n+2}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using DELF is

$$\hat{\theta}_{DELF} = \frac{1}{n+2}.$$

4. BAYES ESTIMATOR AND POSTERIOR RISK UNDER JEFFREYS PRIOR

The Bayes estimator using SELF is

$$\hat{\theta}_{SELF} = \frac{n}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using SELF is

$$\hat{\theta}_{SELF} = \frac{n}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Bayes estimator using PLF is

$$\hat{\theta}_{PLF} = \frac{\sqrt{n\left(n+1\right)}}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using PLF is

$$\hat{\theta}_{PLF} = 2*\left[\frac{\sqrt{n\left(n+1\right)}-n}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}\right]$$

The Bayes estimator using SPLF is

$$\hat{\theta}_{SPLF} = \frac{n\left(n-1\right)}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Posterior risk using SPLF is

$$\hat{\theta}_{SPLF} = 2*\sqrt{\frac{n}{n-1}}-1$$

The Bayes estimator using DELF is

$$\hat{\theta}_{DELF} = \frac{n+1}{\left[-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using DELF is

$$\hat{\theta}_{DELF} = \frac{1}{n+1}.$$

5. BAYES ESTIMATOR AND POSTERIOR RISK UNDER EXPONENTIAL PRIOR

The Bayes estimator using SELF is

$$\hat{\theta}_{SELF} = \frac{n+1}{\left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using SELF is

$$\hat{\theta}_{SELF} = \frac{n+1}{\left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Bayes estimator using PLF is

$$\hat{\theta}_{PLF} = \frac{\sqrt{(n+1)(n+2)}}{\left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using PLF is

$$\hat{\theta}_{PLF} = \frac{\sqrt{(n+1)(n+2)} - n - 1}{\left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Bayes estimator using SPLF is

$$\hat{\theta}_{SPLF} = \frac{n(n+1)}{\left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Posterior risk using SPLF is

$$\hat{\theta}_{SPLF} = 2 * \sqrt{\frac{n+1}{n}} - 1$$

The Bayes estimator using DELF is

$$\hat{\theta}_{DELF} = \frac{n+2}{\left[\frac{1}{b} - log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using DELF is

$$\hat{\theta}_{DELF} = \left(\frac{1}{n+2}\right).$$

6. BAYES ESTIMATOR AND POSTERIOR RISK UNDER GAMMA PRIOR

The Bayes estimator using SELF is

$$\hat{\theta}_{SELF}=\frac{a+n}{\left[b-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using SELF is

$$\hat{\theta}_{SELF}=\frac{a+n}{\left[b-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Bayes estimator using PLF is

$$\hat{\theta}_{PLF}=\frac{\sqrt{\left(a+n\right)\left(a+n+1\right)}}{\left[b-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using PLF is

$$\hat{\theta}_{PLF}=\frac{\sqrt{\left(\alpha+n\right)\left(\alpha+n+1\right)}}{\left[b-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Bayes estimator using SPLF is

$$\hat{\theta}_{SPLF}=\frac{a+n\left(a+n-1\right)}{\left[b-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]^2}$$

The Posterior risk using SPLF is

$$\hat{\theta}_{SPLF}=2*\sqrt{\frac{a+n}{a+n-1}}-1$$

The Bayes estimator using DELF is

$$\hat{\theta}_{DELF}=\frac{a+n+1}{\left[b-log\sum_{i=1}^{n}\left(1-\left(1-x_i\right)^{\lambda}\right)^{\eta}\right]}$$

The Posterior risk using DELF is

$$\hat{\theta}_{DELF}=\left(\frac{1}{a+n+1}\right).$$

7. SIMULATION STUDY

Simulation is a technique in which we conduct an experiments on computer program which would be impracticable. By simulation we get a model that behaves like a real situation. We run simulation for many numbers of time, data are collected, and estimated probability can be calculated. The purpose of using simulation is to check the performance of prior distribution and loss function. We use Mathematica 10 package to generate random sample of size $n \in 30, 100, 200, 500$ from Exponentiated Kumaraswamy distribution and find out the Bayes estimates and Bayes posterior risks of parameter.

N	Prior	α=1.5, β=2, λ=2			
		SELF	PLF	DELF	SPLF
30	Uniform	2.1366 (0.1524)	2.1680 (0.0682)	2.2085 (0.0312)	2.0997 (0.3651)
100		2.0407 (0.0416)	2.0507 (0.0201)	2.0594 (0.0098)	2.0297 (0.2000)
200		2.0199 (0.0204)	2.0268 (0.0100)	2.0313 (0.0049)	2.0116 (0.1414)
500		2.0085 (0.0080)	2.0090 (0.0040)	2.0120 (0.0019)	2.0054 (0.0894)
30	Jeffrey's	2.0746 (0.1485)	2.0989 (0.0682)	2.1401 (0.0322)	2.0274 (0.3713)
100		2.0235 (0.0413)	2.0307 (0.0201)	2.0455 (0.0099)	2.0129 (0.2010)
200		2.0114 (0.0203)	2.0154 (0.0100)	2.0183 (0.0049)	2.0051 (0.1417)
500		2.0040 (0.0080)	2.0053 (0.0040)	2.0085 (0.0019)	2.0023 (0.0895)
30	Exponential	2.0258 (0.1356)	2.0511 (0.0646)	2.0869 (0.0312)	1.9898 (0.3651)
100		2.0071 (0.0402)	2.0155 (0.0198)	2.0299 (0.0312)	1.9987 (0.2000)
200		2.0025 (0.0200)	2.0076 (0.0099)	2.0121 (0.0098)	2.0008 (0.1414)
500		2.0016 (0.0080)	2.0031 (0.0039)	2.0060 (0.0019)	1.9988 (0.0894)
30	Gamma	2.0484 (0.1334)	2.0838 (0.0629)	2.1149 (0.0299)	2.0162 (0.3571)
100		2.0158 (0.0400)	2.0286 (0.0196)	2.0374 (0.0096)	2.0039 (0.1986)
200		2.0096 (0.0200)	2.0134 (0.0099)	2.0187 (0.0049)	2.0052 (0.1409)
500		2.0033 (0.0080)	2.0048 (0.0039)	2.0082 (0.0019)	2.0016 (0.0893)

N	Prior	α=1.5, β=2, λ=1.5			
		SELF	**PLF**	**DELF**	**SPLF**
30	**Uniform**	1.6058 (0.0861)	1.6249 (0.0511)	1.6522 (0.0312)	1.5729 (0.3651)
100		1.5093 (0.0233)	1.5371 (0.0151)	1.5438 (0.0098)	1.5206 (0.2000)
200		1.5155 (0.0114)	1.5189 (0.0075)	1.5217 (0.0049)	1.5128 (0.1414)
500		1.5057 (0.0045)	1.5058 (0.0030)	1.5083 (0.0019)	1.5050 (0.0894)
30	**Jeffrey's**	1.5552 (0.0834)	1.5755 (0.0512)	1.6032 (0.0322)	1.5353 (0.3713)
100		1.5149 (0.0231)	1.5236 (0.0151)	1.5314 (0.0099)	1.5093 (0.2010)
200		1.5074 (0.0114)	1.5101 (0.0075)	1.5139 (0.0049)	1.5021 (0.1417)
500		1.5033 (0.0045)	1.5044 (0.0030)	1.5052 (0.0019)	1.5007 (0.0895)
30	**Exponential**	1.5369 (0.0786)	1.5640 (0.0492)	1.5895 (0.0312)	1.5174 (0.3651)
100		1.5125 (0.0228)	1.5210 (0.0149)	1.5277 (0.0098)	1.5023 (0.2000)
200		1.5073 (0.0113)	1.5088 (0.0074)	1.5116 (0.0049)	1.5023 (0.1414)
500		1.5015 (0.0045)	1.5035 (0.0029)	1.5049 (0.0019)	1.5014 (0.0894)
30	**Gamma**	1.5712 (0..0786)	1.5941 (0.0481)	1.6189 (0.0299)	1.5458 (0.3571)
100		1.5221 (0.0228)	1.5306 (0.0148)	1.5367 (0.0096)	1.5165 (0.1986)
200		1.5088 (0.0113)	1.5162 (0.0074)	1.5183 (0.0049)	1.5070 (0.1409)
500		1.5035 (0.0045)	1.5064 (0.0029)	1.5069 (0.0019)	1.5030 (0.0893)

N	Prior	α=1, β=2, λ=1			
		SELF	PLF	DELF	SPLF
30	Uniform	1.0698 (0.0382)	1.0867 (0.0342)	1.1069 (0.0312)	1.0496 (0.3651)
100		1.0205 (0.0104)	1.0261 (0.0100)	1.0296 (0.0098)	1.0138 (0.2000)
200		1.0103 (0.0051)	1.0128 (0.0050)	1.0148 (0.0049)	1.0074 (0.1414)
500		1.0044 (0.0020)	1.5058 (0.0030)	1.0066 (0.0019)	1.0027 (0.0894)
30	Jeffrey's	1.0326 (0.0367)	1.0528 (0.0342)	1.0669 (0.0322)	1.0181 (0.3713)
100		1.0072 (0.0102)	1.0149 (0.0100)	1.0193 (0.0099)	1.0044 (0.2010)
200		1.0052 (0.0050)	1.0078 (0.0050)	1.0111 (0.0049)	1.0025 (0.1417)
500		1.0026 (0.0020)	1.0033 (0.0020)	1.0031 (0.0019)	1.0006 (0.0895)
30	Exponential	1.5369 (0.0786)	1.5640 (0.0492)	1.5895 (0.0312)	1.0379 (0.0358)
100		1.5125 (0.0228)	1.5210 (0.0149)	1.5277 (0.0098)	1.0111 (0.0102)
200		1.5073 (0.0113)	1.5088 (0.0074)	1.5116 (0.0049)	1.0076 (0.0500)
500		1.5015 (0.0045)	1.5035 (0.0029)	1.5049 (0.0019)	1.0031 (0.0020)
30	Gamma	1.0689 (0.0364)	1.0898 (0.0329)	1.1006 (0.0299)	1.0532 (0.3517)
100		1.0209 (0.0102)	1.0248 (0.0099)	1.0303 (0.0096)	1.0156 (0.1986)
200		1.0104 (0.0050)	1.0139 (0.0049)	1.0139 (0.0049)	1.0062 (0.1409)
500		1.0040 (0.0020)	1.0051 (0.0019)	1.0063 (0.0019)	1.0032 (0.0893)

N	Prior	α=1, β=1, λ=1			
		SELF	PLF	DELF	SPLF
30	**Uniform**	1.0654 (0.0379)	1.0838 (0.0341)	1.1045 (0.0312)	1.0494 (0.3651)
100		1.0187 (0.0103)	1.0252 (0.0100)	1.0284 (0.0098)	1.0162 (0.2000)
200		1.0111 (0.0051)	1.0126 (0.0050)	1.0159 (0.0049)	1.0064 (0.1414)
500		1.0045 (0.0020)	1.0042 (0.0020)	1.0058 (0.0019)	1.0030 (0.0894)
30	**Jeffrey’s**	1.0378 (0.0372)	1.0512 (0.0341)	1.0664 (0.0322)	1.0167 (0.3713)
100		1.0102 (0.0103)	1.0161 (0.0100)	1.0207 (0.0099)	1.0042 (0.2010)
200		1.0061 (0.0050)	1.0072 (0.0050)	1.0112 (0.0049)	1.0015 (0.1417)
500		1.0014 (0.0020)	1.0035 (0.0020)	1.0048 (0.0019)	1.0005 (0.0895)
30	**Exponential**	1.0423 (0.3623)	1.0541 (0.0332)	1.0748 (0.0312)	1.0245 (0.3651)
100		1.0117 (0.0102)	1.0168 (0.0099)	1.0209 (0.0098)	1.0050 (0.2000)
200		1.0055 (0.0050)	1.0088 (0.0050)	1.0104 (0.0049)	1.0037 (0.1414)
500		1.0025 (0.0020)	1.0033 (0.0019)	1.0042 (0.0019)	1.0009 (0.0894)
30	**Gamma**	1.0676 (0.3639)	1.0870 (0.0328)	1.1024 (0.0299)	1.0553 (0.3517)
100		1.0204 (0.0201)	1.0264 (0.0099)	1.0291 (0.0096)	1.0173 (0.1986)
200		1.0101 (0.0050)	1.0130 (0.0049)	1.0168 (0.0049)	1.0078 (0.1409)
500		1.0042 (0.0020)	1.0057 (0.0019)	1.0065 (0.0019)	1.0035 (0.0893)

8. CONCLUSION

We conduct this study to check the scale parameter behavior of Exponentiated Kumaraswamy distribution. To find out the Bayes and posterior risk we use simulation technique, we use uniform, Jefferys non-informative priors and Exponential, G. we use different loss functions like square error loss function, Degroot loss function, precautionary loss function and simple precautionary loss function. After the simulation and comparing different loss functions, the results shows that Jeffreys prior with SPLF is performing better in non- informative prior while in informative prior Exponential prior with SPLF loss function is performing better. Because its Bayes posterior risk is smaller than other loss functions.

REFERENCES

1. Adham, S.A. and ALgfary, A.A. (2016). Bayesian estimation and prediction for a mixture of exponentiated kumaraswamy distributions. *International Journal of Contemporary Mathematical Sciences*, 11(11), 497-508.
2. AL-Fattah, A., EL-Helbawy, A. and AL-Dayian, G. (2017). Inverted kumaraswamy distribution: Properties and estimation. *Pakistan Journal of Statistics*, 33(1), 37-61.
3. Ghosh, I. and Nadarajah, S. (2017). On the bayesian inference of kumaraswamy distributions based on censored samples. *Communications in Statistics-Theory and Methods*, 46(17), 8760-8777.
4. Mandouh, R. (2016). Bayesian inference from the kumaraswamy-weibull distribution with applications to real data. *International Journal of Contemporary Mathematical Sciences*, 11(3), 119-129.
5. Sultan, H. and Ahmad, S. (2015). Bayesian approximation techniques for kumaraswamy distribution. *Mathematical Theory and Modeling*, 5(5), 49-60.
6. Lemonte, A.J., Barreto-Souza, W. and Cordeiro, G.M. (2013). The exponentiated Kumaraswamy distribution and its log-transform. *Brazilian Journal of Probability and Statistics*, 27(1), 31-53.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 239-249

EFFECT OF WATER POLLUTION ON HEALTH: A CASE STUDY IN PAKISTAN

Mahwish Rabia[§]**, Namra Abid, Noor-ul-Ain Shan**
Amina Shabbir and **Ayesha Raana**
Government College Women University, Sialkot, Pakistan
[§]Corresponding author Email: mahwish.rabia@gcwus.edu.pk

ABSTRACT

Water quality issues are a major challenge that humanity is facing in the twenty-first century. The purpose of this study is to identify the effect of water pollution on health. This study is conducted to find out the pollution situation of water and the health problem of the surrounding residents. The results determine that water quality may not be able to sustain the aquatic life and not suitable for using domestic purpose. The study also provides evidence that local communities are suffering from a variety of health problems including skin problem, diarrhea and more other diseases. For this purpose, 314 students of the university are selected to examine and the effect of water pollution on health is studied. The research also examined how much of the earth water is drinkable and how can it effect on our health. The purpose of this study is also to aware the people not to put the garbage in living areas, rivers, lakes and also to aware the people for the limited population. Water quality issues are a major challenge that humanity is facing. Agricultural, chemicals, industries, gases and water wastage sources exert shorter-term effects on regional to local scales. The purpose of this paper is to make a review of the condition of water and their effects on health.

KEYWORDS

Pollution; impact; human health; waste.

1. INTRODUCTION

Water pollution is the main issue affecting human lives in many ways. We all should know the causes, effects and preventive measures of the water pollution to make our lives better. As Earth's population continues to grow, people are putting ever increasing pressure on the planets water resources. We know that pollution is a human problem because it is a relatively recent development in the history of the planet. As industrialization has spread around the globe, so the problem of pollution has spread with it. Oceans, lakes, fields, rivers, and other inland water can naturally clean up a certain amount of pollution by dispersing it harmlessly. This, in turn, could affect the health of all the plants, animals and humans who live in that area. Another way to measure water quality is to examine the quality of the water in rivers. If the river supports no aquatic life at all, the quality is much poorer. A lot of toxic pollution also enters the water from highway runoff. We should take different preventions that can help to tackle the problem through education, laws and economics and they can work together as a term. Making people aware of the problem is the first step to solve this problem. This means that whoever causes pollution should have to pay to clean it up one way or another. We can work together to keep the environment clean so the plants, animals, and people who depend on it remain healthy. We can take community action too by helping out on beach cleans or litter picks to keep our rivers and seas that little bit cleaner. Working together we can make pollution less of a problem and the world a better place.

2. LITERATURE REVIEW

Samrana (2017) a report showed that 91% of the water in Karachi unfits for drinking because of facing with leaky pipes, faulty treatment plants and illegal tapping. The government of Sindh struggled to provide clean and safe water to Karachi's galloping population. Based on physiochemical analysis 21(17.8%) water samples were found unsafe for drinking due to turbidity values beyond the safe limits. Alrumman (2016) reported that over two-thirds of the earth's surface is covered by water; less than a third is taken up by land. As Earth's population continues to grow, people are putting ever-increasing pressure on the planet's water resources. In a sense our oceans, rivers, and other inland water are being "squeezed" by human activities- not so they take up less room, but so their quality is reduced. Halder JN (2015) analyzed that drinking water condition is worst crossing WHO guideline in Karachi city, Sindh province. Previous results of survey 2014 showed that less than 40% of drinking water comes from groundwater is safe other is unsafe or unknown. Owa (2013) analyzed that safe drinking water is necessary for human health all over the world. Being a universal solvent, water is a major source of infection. According to the World Health Organization (WHO), 80% of diseases are water born. Drinking water in various countries does not meet WHO standards. Shafiq (2011) reported that factories are a point source of water pollution, but ordinary people from nonpoint sources pollute quite a lot of water. This is how ordinary water becomes wastewater in the first place. Virtually everyone discharges chemicals of one sort or another down their drains or toilets. A lot of toxic pollution also enters water from highway runoff. Some highway runoff runs away into drains while others can pollute groundwater or accumulate in the land next to a road, making it increasingly toxic as the years ago. Richardson (2009) investigate that another PCRWR team worked on its groundwater samples were collected at 1,184 sites throughout the country taken from wells down into the earth up to 100 feet. The team then used statistical methods to construct a hazard map and to estimate the size of the population exposed to the threat. Kazi (2009) reported that the discharge of domestic and industrial effluent wastes, leakage from water tanks, marine dumping, radioactive waste and atmospheric deposition are major causes of water pollution. Heavy metals that disposed of and industrial waste accumulated in lakes and rivers are harmful for humans and animals. Toxins in industrial waste are the major cause of immune suppression, reproductive failure and acute poisoning. Infectious diseases, like cholera, typhoid fever. Wania and Brown TN (2008) recorded that with billions of people on the planet, disposing of sewage waste is a major problem. Sewage disposal affects people's immediate environment and leads to water-related illness such as diarrhea that kills 525,000 children under five years. Sewage is a completely natural substance that should be broken down harmlessly in the environment 90% of sewage is water. Chemical fertilizers used by farmers also add nutrients to the soil, which drain into rivers and seas and add to the fertilizing effect of the sewage. Arain (2008) recently identified that more than 50 million people in Pakistan were at risk of arsenic poisoning because most of the Pakistani communities use groundwater for drinking and other households polluted with arsenic. Previous studies of 1990's had revealed that groundwater in some areas of Pakistan also contained high levels of arsenic while the extent of those risks was still unknown. Larry (2006) reported that the ground water contamination is the leading worldwide cause of deaths and diseases. It accounts for the deaths of more than 14000 people daily, and most of them are children under five years old. In recent years the widespread reports of pollutants in groundwater have increased public concern about the quality of groundwater. Children are generally more vulnerable to intestinal pathogens and it has been reported that about 1.1 million children die every year due to diarrheal diseases. Smith (2003) publicized that total arsenic composite food in daily dietary intake by adults is higher than in other countries. Irrigation is needed to be systematic while avoiding contaminated water to minimize the accumulation of arsenic in plants which is consumed directly by human's farm animals and wildlife. Ahmed (2000) recorded that 75% to 80% of water pollution is caused by domestic sewage. Waste from the industries like, sugar, textile, electroplating, pesticides, and pulp and paper are polluting the water. A large amount of domestic sewage is drained into the river and most of the sewage is untreated. Domestic sewage contains toxicants, solid waste, plastic litters and bacterial contaminants and these toxic materials cause water pollution. Different industrial effluent that is disposed into the river without treatment is the major cause of water pollution. There is a greater association between pollution and health problem. Disease-causing microorganisms are known as pathogens and these pathogens are spreading the disease directly among humans. Some pathogens are worldwide, and some are found in a well-defined area. Bhattacharya, Kabir and Ali (1995) analyzed that pollution is caused by industries and is more harmful. Solid and liquid waste is discharged into rivers. Water is also contaminated by human excreta. In contaminated water, many bacteria are also found which is harmful to human health. Government is incapable to supply essential needs to citizens because of an increasing number of populations. Sanitation facilities are more in urban areas than in rural areas. Polythene bag and plastic waste is a major source of pollution. Waste is thrown away by putting it into plastic bags.

Objectives:

- To analyze the effect of water pollution on human health.
- To examine the role of human activities in water pollution.
- To assess the role of industries in water pollution.
- To find out whether population growth rate is the cause of population.
- To suggest measure to control water pollution.

3. RESEARCH METHODOLOGY

The present study focused on the point of view, from G.C. Women university, Sialkot about how water pollution effect on human health. A structured questionnaire designed by the researches was used in data collection.

Target Population: All the students of G.C. Women University Sialkot.
Sampled Population: Four department of G.C. Women University. Sialkot.
Sampling techniques: The data is collected using simple random sampling.

$$n = \frac{N}{1+Ne^2} \tag{1}$$

Statistical Tools: To observe the effect of water pollution on human health for this we used Graphical representation and Pearson's correlation analysis to analyze the data.

4. DATA ANALYSIS AND DISCUSSION

Graphical Representation:

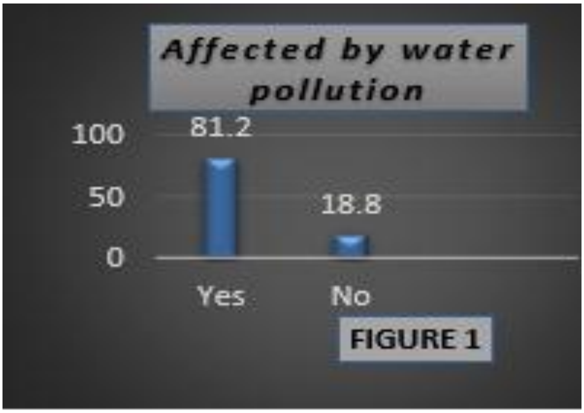


FIGURE 1

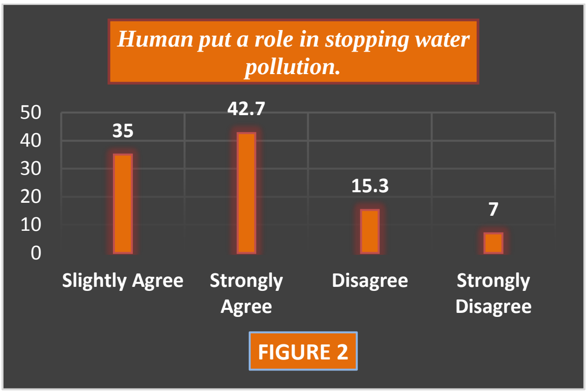


FIGURE 2

Figure 2 shows that 35% of people slightly agree that human put a role in stopping water pollution, 42.7% of people strongly agree that human put a role in stopping water pollution, 15.3% of people disagree that humans put a role in stopping water pollution and 7% of people strongly disagree that human put a role in stopping water pollution. Figure 1 show that 81.2% of people affected by water pollution and 18.2% of people are not affected by water pollution.

Table 1.1
How people Concerned about Water Pollution

Variable	Categories	Number	Percent
People Concerned about Water Pollution	Extremely Concerned	216	68.8
	Very Concerned	64	20.4
	Slightly Concerned	28	8.9
	Not at all Concerned	6	1.9

Table 1.1 shows that 68.8% of people are extremely concerned about the water pollution, 20.4% of people are very concerned, 8.9% of people are slightly concerned about water pollution and only 1.9% of people are not at all concerned about water pollution. Water is a basic resource that guarantees the life of all living beings on the planet. Water pollution has a large set of adverse effects on water bodies (lakes, rivers, ocean, groundwater) caused by human activities.

Table 1.2
Human activities become the Reason of Water Pollution.

Variable	Categories	Number	Percent
Human Activities become the Reason of Water Pollution	Yes	307	97.8
	No	7	2.2

Table 1.2 shows that 97.8% of people respond that human activities become the reason of water pollution and the only 2.2% of people respond that human activities are not the reason for water pollution. Humans are the main cause of water pollution, which is triggered in many ways such as industrial waste and their unplanned leaking, runoff, dumping and disposal. Our water resources face a host of serious threats, all of which are caused primarily by human activities.

Table 1.3
What Percentage do the Industries Pay in Water Pollution

Variable	Categories	Number	Percent
Percentage do the Industries Pay in Water Pollution.	30% to 40%	43	13.7
	40% to 50%	39	12.4
	50% to 60%	63	20.1
	60% to 70%	169	53.8

Table 1.3 shows that 13.7% of people respond that 30 to 40 percent of the industries pay in water pollution, 12.4% people respond that 40 to 50 percent of the industries pay in water pollution, 20.1% of people respond that 50 to 60 percent of the industries pay in water pollution and the 53.8% respond that 60 to 70 percent of the industries pay in water pollution. The industry is a huge source of water pollution, it produces pollutants that are extremely harmful to people and the environment. Many industries use freshwater to move the waste away from the plants; into the rivers, lakes and oceans.

Table 1.4
Increasing Population is the Cause of Water Pollution

Variable	Categories	Number	Percent
You Think Increasing Population is the Cause of Water Pollution	True	255	81.2
	False	59	18.8

Table 1.4 shows that 81.2% of people think that it is true, increasing population is the cause of water pollution and the only 18.8% of people think that it is false. As the population increases, the demand for limited fresh water also increases. Increase in population also accelerates the pollution due to a large amount of waste generated and disposed into the water that ultimately increases the pollution in the water bodies, water from rivers, lakes, underground etc. is used for both industrial and domestic purposes. 80% of the water that is used for domestic purposes comes out as wastewater and the rising number of industries like thermal power plants, engineering industries and textile industries have a large number of impact on water bodies.

Table 1.5
Soil Destruction is the Reason of Water Pollution

Variable	Categories	Number	Percent
Another Reason of Water Pollution is soil Destruction. Do you agree that Statement?	Slightly Agree	139	44.3
	Extremely Agree	142	45.2
	Not Agree	18	5.7
	Agree	15	4.8

Table 1.5 shows that 44.3% people are slightly agreed, 45.2% people are extremely agreed, 5.7% people are not agreed while only 4.8% people are agreed that another reason of water pollution is soil destruction. Indirect sources include pollutants that enter the water bodies via groundwater, soil or through the atmosphere as acid rain.

Table 1.6
Use of Oxygen is a Source of Water Pollution

Variable	Categories	Number	Percent
Use of Different Oxygen is a Source of Water Pollution?	Strongly Disagree	41	13.1
	Strongly Agree	40	12.7
	Disagree	139	44.3
	Agree	94	29.9

Table 1.6 shows that 13.1% people are strongly disagreed, 12.7% people are strongly agreed, 44.3% of people are disagree while the only 29.9% of people are agreed that use of oxygen is a source of water pollution. Organic wastes such as sewage and farm waste impose high oxygen demands on the receiving water leading to oxygen depletion with potentially severe impacts overall eco-system.

Table 1.7
How much of the Earth Water is Drinkable

Variable	Categories	Number	Percent
According to your Opinion how much of the Earth Water is Drinkable?	1%	130	41.4
	5%	95	30.3
	30%	31	9.9
	10%	58	18.5

Table 1.7 shows that 41.4% people respond that 1% of the earth water is drinkable, 30.0% people response that 5% of the earth water is drinkable, 9.9% people respond that 30% of the earth water is drinkable and the 18.5% people respond that 10% of the earth water is drinkable.

Table 1.8
The way you and your household can be more informed about water pollution

Variable	Categories	Number	Percent
The way in which you and your household can be more informed about water pollution.	Internet	63	20.1
	Posters	28	8.9
	School/university	91	29.0
	Community	132	42.0

Table 1.8 shows that 20.1% people respond that through internet we informed about water pollution, 8.9% people respond that through posters we are informed about water pollution, 29.0% people respond that through school and university we are informed about water pollution and the 42.0% people respond that through community we are informed about water pollution. For controlling the water pollution, we should have to inform people through the internet, posters or by using other methods.

Table 1.9
You taken any Initiatives for Protecting yourself from Water Pollution

Variable	Categories	Number	Percent
Have you taken any initiatives for Protecting yourself from water pollution?	Yes	260	82.8
	No	54	17.2

Table 1.9 shows that 82.8% people respond that they take any initiatives for protecting themselves from water pollution and the only 17.2% of people respond that they do not take any initiatives for protecting themselves from water pollution. 82.8% of people protect themselves in a way such that they don’t drink water directly from a pond, stream, river, or lake. Water for domestic use must be tested after a certain period, filter or boil water if needed. We should have to aware the people not to put the garbage near to your house or other places.

Table 2.0
The Severe Diseases are cause by Water Pollution

Variable	Categories	Number	Percent
The Severe Diseases are cause by Water Pollution	Slightly Agree	31	9.9
	Agree	260	82.8
	Extremely Disagree	22	7.0
	Disagree	1	0.3

Table 2.0 shows that 9.9% people are slightly agreed, 82.8% people are agreed. 7.0% people extremely disagree while the only 0.3% people are disagreed that the severe diseases are caused by water pollution. When humans, drinks polluted water then it has many disastrous effects on their health. It causes typhoid, cholera and hepatitis and various other diseases. Humans are harmed by the alteration in the food chain and by contracting illness when drinking or using contaminated water. As water pollution has a great impact on human health.

Table 2.1
Is it necessary to Aware People not to put the Garbage in Water

Variable	Categories	Number	Percent
Is it necessary to aware the people not to put the garbage in water?	Yes	298	94.9
	No	16	5.1

Table 2.1 shows that 94.9% people are respond that it is necessary to aware the people not to put the garbage in the water and the only 5.1% people respond that it is necessary to aware the people not to put the garbage in the water. It is most important to aware the people not to put the garbage in water because it causes pollution in the sea, oceans, river etc. and affects the human health, sea life and the environment.

Table 2.2
What do you do with Garbage Collected near your Society

Variable	Categories	Number	Percent
What do you do with the Garbage Collected near your Society?	Dump it into drainage canal	53	16.9
	Wait for the garbage truck	168	53.5
	Burn it	36	11.5
	Other	57	18.2

Table 2.2 shows that 16.9% people respond that they dump the garbage near to their society into drainage canal, 53.5% people respond that they wait for the garbage truck, 11.5% of people respond that they burn the garbage near to their society and the 18.2% of people respond that they use other methods to dispose garbage. Use cloth bags instead of plastic. Buy food that has less packaging. Burn the garbage, use drainage canal or use other methods for collecting the garbage near to your society.

Table 2.3
As the Household Wastage is the Main Reason of Water Pollution

Variable	Categories	Number	Percent
As the household wastage is the main reason of water pollution? What would be percentage of it?	30-40%	97	30.9
	50-60%	111	35.4
	15-25%	63	20.1
	20-30%	43	13.7

Table 2.3 shows that 30.9% people respond that 30% to 40% household wastage is the main reason of water pollution, 35.4% people respond that 50% to 60% household wastage is the main reason of water pollution, 20.1% people respond that 15% to 25% household wastage is the main reason of water pollution and the only 13.7% people respond that 20-30% household wastage is the main reason of water pollution. Water pollution is caused due to sewage, garbage and liquid waste of households, agricultural lands and factories are discharged into lakes and rivers. The household wastage also affects the water and the environment.

Table 2.4
Correlation Coefficient

Attributes	Correlation Coefficient	p-value	Conclusion
You have taken any initiatives for protecting yourself and what you do with garbage near to your home.	-0.141^{**}	0.00	Significant
Human activities become a reason of water pollution and to aware the people not to put the garbage.	-0.161^{**}	0.00	Significant
Have you ever been affected by water pollution and increasing population is the cause of water pollution?	0.146^{**}	0.01	Significant

** Correlation is significant at 1% level of significance
Source: SPSS

Table 2.4 illustrates that the significant value of correlation coefficient is r=(-0.141) having a p-value less than significance level (α=0.01). It depicts that there is a negative relationship between study variables. Since when people take initiatives to protect themselves from water pollution like burn the garbage, use drainage canal or use other methods to through the garbage then automatically the quantity of polluted water will be decreases. Meanwhile, the value of correlation coefficient -0.161 demonstrates that human activities play a vital role in increasing water pollution such as industrial waste and their unplanned leaking, runoff, dumping and disposal and by increasing the awareness among people not to put the garbage near to society could help to control human activities in order to get clean water. The value of correlation coefficient (r=0.146) shows a positive correlation between the variables that is, respondents ever been affected by water pollution and increasing population is one of the major cause of water pollution. The significant value at 1% level of significance demonstrates that increasing population is the main cause of water pollution.

5. CONCLUSION AND RECOMMENDATIONS

Water pollution is a global issue and world community is facing worst results of polluted water. The purpose of the study is to analyze the effect of water pollution on health. The data is collected from GCWUS. The study provides evidence that local communities are suffering from a variety of health problems that could be a direct or indirect result of the discharge and flow of waste water.

The results of the analysis show that 97.8% respondent believe that human activities become the reason of water pollution and 53.8% illustrates that 60-70% of industries pay their role in water pollution. Moreover, 81.2% demonstrates that the increasing population is the cause of water pollution and 82.8% agree that severe diseases these days are the cause of water pollution. On the other hand, 94.9% gives opinion that it is necessary to aware the people not to put the garbage in the water and 82.8% agreed that some initiatives should be taken to overcome water pollution.

Furthermore, the value of correlation coefficient -0.161 demonstrates that human activities play a vital role in increasing water pollution such as industrial waste and their unplanned leaking, and by increasing the awareness among people not to put the garbage near to society could help to control human activities in order to get clean water.

It can be concluded that since the pressure of increasing population, untreated discharge from industries, wasted garbage, use of chemical fertilizers etc., are causing water pollution. So, to overcome this problem there should be proper waste disposal system and waste should be treated before entering into the river. Government should take initiatives for protecting human's health from water pollution. There should be proper area for the waste garbage and for the industrial wastage. We have to prevent it on time and manage to have clean water. Easy and sustainable surface water treatment should be introduced to the public, especially to the affected areas. Educational and awareness programs should be organized to control the pollution.

Future Study:

In future, few locations that are more representative may be selected to define stations.

1. While carrying out future studies, it should be planned to collect water samples from different places and in different times of the year, preferably at least once during the wet season and other during the dry season.
2. Samples should be tested as early as possible after bringing to the laboratory, preferably by 24 hours. If it is not possible samples should be refrigerated appropriately. The increasing demand for fresh water resources over the next decades will exert an enormous pressure, particularly in arid regions of the world, to protect surface water from pollution.

REFERENCES

1. Ahmed, M.F. and Rahman, M.M. (2000). *Water supply & sanitation: Rural and low income urban communities*. ITN-Bangladesh, Centre for Water Supply and Waste Management, BUET.
2. Alrumman, S.A., EI-Kott, A.F. and Kefsk, M.A. (2016). Water pollution: source and treatment. *American Journal of Environmental Engineering*, 6(3), 88-98.
3. Arian, M.B., Kazi, T.G., Jamali, M.K., Afridi, H.I. and Baig, J.A. (2008). Evaluation of physic-chemical parameters of Manchar Lake water and their comparison with global published values. *Pak. J. Anal. Environ. Chem.*, 9, 101-109.
4. Bhattacharya, D., Kabir, B.N. and Ali, K. (1995). Industrial Growth and Pollution in Bangladesh: A Sectoral Analysis. In symposium on *Environment and Sustainable Development with Special Reference to Bangladesh*", North South University, Dhaka.
5. Brown, T.N. and Wania, F. (2008). Screening chemicals for the potential to persistent organic pollutants: a case study of Arctic contaminants. *Environ. Sci. Technol.*, 42(52), 2-9.
6. Halder, J.N. and Islam, M.N. (2015). Water pollution and its impact on human health. *Journal of Environment and Human*, 2(1), 36-46.
7. Kazi, T.G., Arain, M.B., Jamali, M.K., Jalbani, N., Afridi, H.I., Sarfraz, R.A., Baig, J.A. and Shah, A.Q. (2009). Assessment of water quality of polluted lake using multivariate statistical techniques: case study. *Ecotox Environ Safety*, 72, 301-309.
8. Larry, W. (2006). *World Water Day*. *A Billion people Worldwide Lack Safe Drinking*. Online available at http://environment.about.com/od/environmentalevents/
9. Owa, F.D. (2013). Water pollution: sources effects, controls, and management. *Mediterranean Journal of Social Sciences*, 4(8), 65-8.
10. Richardson, S.D. (2009). Water analysis: emerging contaminated and current issues. *Anal. Chem.*, 81(46), 45-77.
11. Samrana, S., Ali, I., Azizullah, A., Daud, M.K. and Gan, Y. (2017). Arsenic-Based Pollutio n Status in Pakistan. *Ann. Agric. Crop. Sci*., 2(2), 1027.
12. Shafiq, H.B., Ajaz, M. and Rasool, S.A. (2011). Bacterial and toxic pollutants in lakes of rivers. *Pak. J. Bot.*, 43, 1765-1772.
13. Smith, J.V.S., Jankowski, J. and Sammut, J. (2003). Vertical distribution of As (III) and As (V) in a coastal sandy aquifer: factors controlling the concentration and speciation of arsenic in the Stuarts Point groundwater system, northern New South Wales, Australia. *Applied Geochemistry*, 18(9), 1479-1496.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 251-265

IMPACT OF STUDENT-TEACHER RELATIONSHIP ON STUDENT ACADEMIC ACHIEVEMENTS

Soha Tehreem and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: sohatehreem123@gmail.com
zaheerabbas@uog.edu.pk

ABSTRACT

The Student-Teacher relationship is key ingredient for learning process of Students. There are some Formal (Structured Environment and Behavior in Class etc.) and Informal (Outside the class etc.) factors that effects the relationship between teacher and students. The study sample consist of 200 respondents from University of Gujrat by using Simple Random sampling. A well-structured questionnaire was used to collect information regarding the subject. The methods used to analyze the data are Descriptive Statistics, Confirmatory Factor Analysis and Structure Equation Modeling. The results of this study shows that these factors play significant role to predict the relationship so we conclude that there are positive relationship between student-Teacher relationship and achievements and class management is most important factor.

KEYWORDS

Learning process, Student-Teacher relationship, Achievement’s, Students, Structure Equation Modeling, Confirmatory Factor Analysis, Formal and Informal.

1. INTRODUCTION

Getting quality education is a critical foundation in the lives of each person. It is basic that understudies have the gadgets they should be profitable instruments that join inspiration and obligation. For explicit understudies, regardless, inspiration isn’t persistently trademark. It consequently tumbles to others to direct understudies on the way to their very own planning. As instructors contribute an inconceivable extent of imperativeness with their understudies over the course, it is a teacher's commitment to help a tendency for learning. Research has shown that the relationship among instructors and understudies is a basic pointer of scholastic obligation and accomplishment. Believe it or not, the most common weapon teachers have, when attempting to empower a phenomenal learning air, is certain association with their understudies. Understudies who consider their to be as constantly predictable have better accomplishment results (Boynton and Boynton, 2005; Spilt, Komen, and Thijs, 2011; Skinner and Green, 2008; Rimm, Kaufman and Sandilos, 2012; Gehlbach, Brinkworth, and Harris, 2012). Furthermore, the learning condition acknowledge a gigantic action in keeping up understudy premium and obligation. Precisely when understudies feel an estimation of control and security in the assessment hallway, they are legitimately pulled in considering the way that they approach learning with vitality and life. Understudies become dynamic people in their own planning (Skinner and Green, 2008; Maulana, Opdenakker, Stroet, and Bosker, 2013).

Management alludes to every one of the things that an teacher does to compose students, space, time, and materials with the goal that students learning can occur characterizes student focused learning groups as having shared administration, group building, and a harmony between the requirements of the instructors and students (Wong et al., 2012). Students of all learning capacities and social foundations are adapting together in the same classrooms with teachers being considered responsible for every individual students' achievement (Brannon, 2010). Establishing effective classroom management takes time, and differs from teacher to teacher based on their personality and preferred teaching style, as well as being dependent on subject and age group. Due to the fluidity of how classroom management presents itself, there is no set 'how to' on securing complete control of your classroom, however there are guidelines and core components that make up effective classroom management.

Due to societal changes, universities have more behavioral issues that affect the way a teacher manages the classroom (Etheridge, 2010). Today, classroom disciplinary issues are worse than those in the past, which has impacted student achievement (Colavecchio & Miller, 2002; Etheridge, 2010). Classroom management issues have a devastating impact on students' achievement.

Relationships with teachers have been studied from the perspective of the Attachment Theory, in which teachers, are seen as "alternative caregivers" (Howes, 1999) and as extensions of the parent-child relationship (Davis, 2003). Students who feel emotionally secure with their teachers can communicate effectively and are better able to devote their attention to learning (Pianta, 1999). Since teacher-student interactions represent a bidirectional, transactional process between individuals, it is important to recognize and understand what each individual may contribute to these interactions, as well as how other contextual factors may serve as facilitators or impediments. Pianta and Walsh (1996) developed the Contextual Systems Model (CSM), designed to explain the experiences of students, which is a good example of an ecological-contextual model of development.

Over the past decades, researchers have engaged in conducting studies in chemistry classrooms in order to establish an effect of a particular teaching method/strategy on students' academic achievement (Offiah & Egolum, 2006; Okonkwo, 2007). Empirical evidence has shown that studies on effects of using the learning cycle method in teaching (Qarareh, 2012), Demonstration strategy (Giridharan & Raju, 2016), jigsaw method (Abdulkadir, 2016), concept mapping (Qarareh, 2010), experiential and generative learning strategies (Adeyemi & Awolere, 2016), computer simulations (Okwuduba et al., 2018) on academic achievement of students in chemistry and sciences have been successfully carried out.

The concept of motivation is considered as a crucial factor that affects human behavior and performance (Kian et al. 2014; Turan 2015). Especially educational researchers and practitioners express that motivation is one of the most important factors in student achievement and in ensuring continuous achievement (Alkış 2015; Aluçdibi and Ekici 2012; Guay et al. 2010; Pintrich 2003; Pintrich and Schunk 2002). Lin (2012) describes motivation as intrinsic desires which are already present in the individual or which are reflected in the individual while acquiring new information and learning. There are, however, in the literature other definitions of motivation; the latter word was derived from the word “movere” that means moving in Latin (Seiler et al. 2012). In this regard, according to Ertem (2006), motivation is an inner state uncovering individuals’ behavior and directing them to these behaviors; however, according to Baumeister and Vohs (2007), it is a state where the individual displays various attitudes voluntarily in order to achieve a certain goal. Küçüközkan (2015) defined motivation as the sum of the efforts made for mobilizing the individual towards one or more particular goal.

2. SIGNIFICANCE AND OBJECTIVE OF STUDY

The meaning of life is worthless without any goal or aim. Every human being has goals or aims which may be diverse in nature. The desire to achieve higher academic achievement is inherited by every human being since birth for successful survival in society.

It has been observed through the literature review that Class Management, Teachers’ behavior, teaching methodology and effectiveness, motivation and feedback and many other factors plays an important role for high academic achievement of learners.

The purposed study will be very helpful for teachers, students, top level management and policy makers as they will have more additional information about necessary features affecting teaching and learning process. The study will also have significant importance for teachers as they use feedback in classroom. Its impact on academic achievement will help them in teaching learning process. The study may lead to arrangements for continues professional development of teachers in relation to motivation and right type of feedback. The study might be helpful for teachers in relation to effect of classroom environment on learner’s academic achievement. The study will be of great significance as it will helpful to measure effect of Student-Teacher relationship on academic achievement. The study will also lead to creating overall conducive environment for learning and to enhance academic achievement.

Following objectives have been formulated to achieve the above sited goals:

- To explore the basic features of data.
- To fit a measurement model to measure the student academic achievements by using confirmed significant factors.
- To look at instructors and understudies observations with respect to their relationship.
- To investigate connection between instructors understudies relationship and understudies accomplishments.
- To recognize measurements of connection between educators understudies.
- To suggest methodologies for development of instructors understudies relationship.

3. FACTORS IN STUDY

The factors in my study are

1) Formal Factors
   a) Within the Class
      - Provide structure or Class Management
      - Teaching Methodology or Effectiveness
      - Teachers' Attitude
2) Informal Factors
   b) Outside the Class
      - Teachers' Behavior or attitude
      - Motivation and feedback from Teacher

3.1 Class Management

According to Whitaker (2005), the main variable in the classroom is not the student, but the teacher. Great teachers have high expectations for their students, but even higher expectations for themselves. These teachers recognize the importance of connecting with their students, that if they are unable to connect with them emotionally then influencing their minds may be impossible.

"Good teachers put snags in the river of students passing by, and over time, they redirect hundreds of lives… There is an innocence that conspires to hold humanity together…" (Bolman & Deal, 2005, p.124).

3.2 Psychological Effects of Student Teacher Relationship on Students

Students are influenced by perceptions of their teacher's even handedness, competence, caring and support as well as the nature of the teacher-student relationship that results (Stipek, 2005). A student wants to feel connected to people and to feel as though he or she deserves to be loved and respected (Stipek, 2005). According to Stipek many of the students who are not doing well academically, are the same ones who have a poor relationship with their teachers. Typically, the more they fall behind academically, often, the more this relationship is weakened. If they are constantly remains back in class.

3.3 Environmental Influence

Reinforcement theorists argue that motivation is in the environment, not in the person such as the teacher (Stipek, 2005). However, it is the teacher who plays the greatest role in setting the atmosphere (Whitaker, 2005). Whitaker (2005) argues that it is better to create the relationship that will motivate the student to behave well. Climate and culture will enable or restrict classroom instruction and student learning (Stewart, 2008), since students adapt to their environment.

3.4 Role of Student and Teacher

Teacher knowledge and efficacy of student motivation and achievement are crucial component to creating relationships that motivate. Both teachers and students have to value their contribution. A student has to feel worthwhile and appreciated. A teacher needs to recognize that he or she can have a positive effect on their students. Wise man and Hunt (2005) refer to this as “teacher efficacy” and note that the more teacher believes in this, the more they will cause it to happen.

3.5 Teacher’s Expectations

Positive teacher expectations were associated with high academic performance or academic gains, whereas negative teacher expectations resulted in decrease in academic performance. The significance of knowing teacher’s beliefs regarding their roles in student motivation is crucial due to the accepted correlation between this perception and actions.(Tyler and Boelter, 2008) “The quality of teacher, student relationships is the keystone for all other aspects of classroom management” (Marzano & Marzano, 2008).

4. METHODOLOGY

Population of study in hand includes all students of MSc and M.Phil from University of Gujrat in Hafiz Hayat Campus. The total population size is our selected campus is 1000. As we know the population and we can find sample size using the methods where population standard deviation is required. Hence, the sample size *n* can be determined by using Yamne (1967) formula as

$$n=\frac{N}{1+Ne^2}=200$$

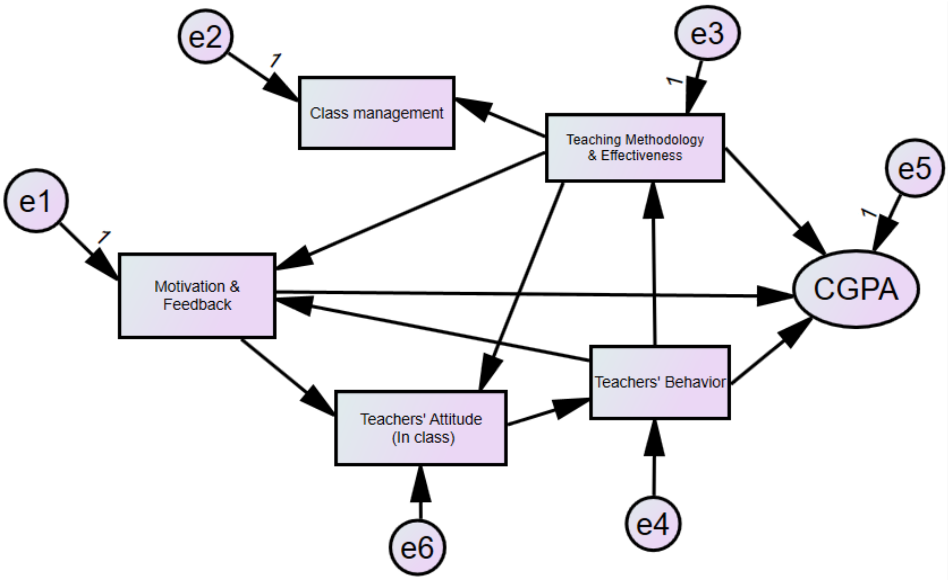
where population size N = 1000 and Margin of error e = 0.05. And Simple Random Sampling was used for sample selection. For analysis Descriptive statistics, confirmatory Factor analysis and Structure Equation Modeling are used.

Descriptive statistics is used to get summary of the data. It provides the features and picture of data. Measures of central tendency, measure of dispersion used in quantitative and percentage is used to get results about qualitative.

Confirmatory Factor analysis (CFA) is used to confirm the measurement theory that is to reject or accept the preconceived measurement theory. In other words it is to use to confirm the hypothesis that a set of variables belongs to a construct known as latent variable. In confirmatory factor analysis (CFA), theory is a systematic set of casual relationships that provide the comprehensive explanation of a phenomenon. In confirmatory factor analysis (CFA), model is a specified set of dependent relationships that can be used to test the theory. In confirmatory factor analysis (CFA), path analysis is used to test structural equations. The path diagram shows the graphical representation of cause and effect relationships of the theory. Usually, statistical software like AMOS, LISREL, EQS, Statistica and SAS are used for confirmatory factor analysis (CFA).

Structural equation modeling (SEM) is a family of statistical models that finds out and explains the relationship among multiple variables. It examines the structure of interrelationships expressed in a series of equations, similar to a series of multiple regression equations. Structural equation modeling (SEM) can be considered of as a unique combination of both types of techniques (interdependence, dependence) because SEM’s base lies in two known multivariate techniques: factor analysis and multiple regression analysis. SEM is the only multivariate technique that allows the simultaneous estimation of multiple equations. Every multi-item constructs in a SEM model can be thought of as variate. Goodness of fit index (GFI), adjusted goodness of fit index (AGFI), root mean square error approximation (RMSEA) provides adequate evidence of model fit.

The path diagram for SEM is



5. RESULTS AND DISCUSSION

Reliability of the questionnaire is an essential factor for statistical results. Table A-1 consists of the test of reliability; the value of Cronbach's alpha based on standardized item is 0.869 which indicates that the data is reliable for further statistical analysis.

Descriptive statistics are considered primarily to explain the basic features of the data. It provides simple summaries about the sample measures. Table A-2 contains the mean and standard deviation of age of the respondents. Results show that the minimum age of the respondent is 20 years and maximum is 28 years. The mean age is 20.60 years with S.D 1.857; it shows that all respondents are within 20 to 28 years of age.

Table A-3 contains the sample information that the students 69.1 percent were from MSc and remaining 30.9 percent were from M.Phil Program. And gander consist of 45.5 percent male and remaining female, the students from Rural area were 52.7 percent with high percentage, Urban with 27.2 percent and Sub-urban with 20.1 percent.

Table A-4 shows the percentage rank of respondents on the statement that Student-Teachers' relation greatly influence your Performance. For first factor overall 47.5 % students are agree with the first factor that class management have positive impact on students' learning , 37.16 %students agree with the factor second that Teaching Methodology and Effectiveness play key role for Students' Learning , 41% students are agreeing that Teachers' behavior has influence on student learning , 44% students are agreeing that motivation and feedback have positive impact on learning and 46.9%

students are agree with the factor Teachers' attitude outside the class have great impact on student learning.

In this paper, we run confirmatory factor analysis on all of the factors which we used in the analysis. In CFA, distinguishing between independent and dependent variables is not necessary. Appendix-B shows the model estimates of confirmatory factor analysis of all factors. Every factor with six items is tested for confirmation, p values of all the confirmed items are significant. So we conclude that our items of the factor are confirmed for the factor. Results for Formal Factors with eighteen items show that Environment parameter estimate value of 0.923, which is high as compared to other variables in the factor; it means that variable is most important for the factor. Evaluation process has a parameter estimate value of 0.140, which is low as compared to other variables in the factor; it means that variable is less important for the factor.

Informal factor with twelve items is tested for confirmation. P-values of all the items are significant so we reject the null hypothesis that all items are not confirm for that factor and conclude that all items of that factor are confirmed for that factor. Results show Encouragement effect on goal has a parameter estimate value of 2.389, which is high as compared to other variables in the factor; it means that variable is most important for the factor. Meeting with teacher outside the class has a parameter estimate value of 0.524 which is low as compared to other variables in the factor; it means that variable is less important for the factor.

Appendix-B shows Goodness of Fit measures. Chi-square (χ^2) is the fundamental Goodness of fit measure used in CFA. Recommended value of ($\chi^2/d.f$) is less than 3and we can see in all factors, the value of ($\chi^2/d.f$) is less than 3. So it supports our estimated model. Goodness of Fit Index (GFI) is absolute goodness of fit measure. The recommended GFI is greater than or equal to ($\geq$) 0.90 and in our all factors we can see that the GFI is $\geq$ 0.90, which supports our estimated model. Another measure to assess the goodness of fit of estimated model is Root Mean Square Error of Approximation (RMSEA). Recommended range of RMSEA is less than .08 and in our factors all RMESA are less than 0.08. So RMSEA is supported to fitted model. All goodness of fit measures meets the recommendation level for this factor. So goodness of fit measures supports our estimated model. After Confirmatory Factor Analysis now we fit the structural equation model on those factors (including items) that are confirmed by Confirmatory FactorAnalysis

In this study we use Structural Equation Modeling to make model for decision making. We use (formal/within class factors) Class management, Teaching methodology, Teaching attitude and (Informal/ outside the class factors) Motivation and feedback and Teachers' behavior making as a constructs. Structure model involves specifying Structural relationships between latent constructs. Table C-1 contains the Parameter estimates of Structure Equation model. P-values of all the relation are significant so we reject the null hypothesis that the coefficients are zero. So we conclude that all relations are significant. Table C-2 shows the Goodness of Fit measures of SEM. p-value of Chi-Square test is significant. So our model is fit. Recommended value of ($\chi^2/d.f$) is less than 3. In this case, the value of ($\chi^2/d.f$) is 2.349 that is less than 3. So it also supports our estimated model. In this case, GFI value is 0.927 the value of AGFI is .0.898, which supports our estimated model. In this model, the value of RMSEA is .061 that is less than .08, which is also supported to our fitted model. All the important Goodness of fit measures indicates that our estimated model is best fitted.

From all above analysis we conclude that all the factors in our study are significant to Testing our hypothesis so we can’t reject H_o and conclude that there is positive relation between Student-Teacher relationship and Student Academic performance.

6. CONCLUSION

In this study we determine the different factors that are related to Student-Teacher relationship and student academic achievements of University of Gujrat by using Structural equation modeling. Study findings reveal that, most significant factor which influenced student academic achievements is class management. The effects mentioned in estimated equations show that we can’t reject our null hypothesis and conclude that there is positive relationship between Student-Teacher relationship and student academic achievements. In other words we can say there is positive impact of student-Teacher relationship on student grads/CGPA.

REFERENCES

1. Ayaz, M., Shah, R.U. and Khan, G.A. (2013). Impact of student-teacher relationship on student’s academic achievement at secondary level in Khyber Pakhtunkhwa. *International Journal of Learning and Development*, 3(1), 181-19.
2. Camp, M.D. (2011). *The power of teacher-student relationships in determining student success* (Doctoral dissertation, University of Missouri--Kansas City).
3. Gablinske, P.B. (2014). A case study of student and teacher relationships and the effect on student learning. Master's Theses, Dissertations, Graduate Research and Major Papers Overview. 106. https://digitalcommons.ric.edu/etd/106
4. Rebrean, L.M. (2003). The Effects of Teacher-Student Relationships on Academic Achievement–a College Survey. *Journal of Early Adolescence*, 23(1), 5-6.
5. Creasey, G., Jarvis, P. and Knapcik, E. (2009). A measure to assess student-instructor relationships. *International Journal for the Scholarship of Teaching and Learning*, 3(2), Article 14.
6. Bergeron, J., Chouinard, R. and Janosz, M. (2011). The Impact of Teacher-Student Relationships and Achievement Motivation on Students' Intentions to Dropout According to Socio-Economic Status. Online Submission, *US-China Education Review* B 2 p273-279.
7. Al Nasseri, Y.S., Renganathan, L., Al Nasseri, F. and Al Balushi, A. (2014). Impact of Students-Teacher Relationship on Student's Learning: A Review of Literature. *International Journal of Nursing Education*, 6(1), 167.
8. Wanders, F.H., Dijkstra, A.B., Maslowski, R. and Van der Veen, I. (2019). The effect of teacher-student and student-student relationships on the societal involvement of students. *Research Papers in Education*, 1-21. DOI: 10.1080/02671522.2019.1568529
9. Varga, M. (2017). *The Effects of Teacher-Student Relationships on the Academic Engagement of Students*. Master thesis, Goucher College.

10. Wood, K., Hojnoski, R.L., Laracy, S.D. and Olson, C.L. (2016). Comparison of observational methods and their relation to ratings of engagement in young children. *Topics in Early Childhood Special Education*, 35, 211-222.
11. Wentzel, K.R. (1997). Student motivation in middle school: The role of perceived pedagogical caring. *Journal of Educational Psychology*, 89(3), 411-419.
12. Wentzel, K.R. (1998). Social Relationships and Motivation in Middle School: The Role of Parents, Teachers, and Peers. *Journal of Educational Psychology*, 90(2), 202-209. doi:10.1037/00220663.90.2.202.
13. Chouinard, R. (2005). *The motivation to learn, a question of perception and context*. Montreal: CTREQ.
14. Crane, J. (1991). The epidemic theory of ghettos and neighbourhood effects on dropping out and teenage childbearing. *American Journal of Sociology*, 96, 1226-1259.
15. Eccles, J., Wigfield, A. and Schiefele, U. (1998). Motivation to succeed. In W. Damon, & N. Eisenberg (Eds.), *Handbook of child psychology* (5th ed., Vol. 3). NS: John Wiley.
16. Creasey, G., Jarvis, P. and Gadke, D. (2009). Student Attachment Stances, Instructor Immediacy, and Student–Instructor Relationships as Predictors of Achievement Expectancies in College Students. *Journal of College Student Development*. 50, 353-372. 10.1353/csd.0.0082.
17. Davis, H. (2003). Conceptualizing the role and influence of student-teacher relationships on children’s social and cognitive development. *Educational Psychologist*, 38, 207-234.
18. Karen, K.; Thisjs M., and Schakel L. (2002). The relationship of emotional intelligence with academic intelligence and the Big-Five. *European Journal of Personality*, 16(2), 103-125.

APPENDIX-A

Table A-1
Test of Reliability

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
0.869	0.871	30

Table A-2
Descriptive Statistics of Quantitative Variable

Descriptive Statistics				
Variables	**Minimum**	**Maximum**	**Mean**	**Std. Deviation**
Age in years	20	28	20.60	1.857

Table A-3
Percentages of Demographic Variables

<table>
<tr><th>Variable</th><th colspan="3">Categories</th></tr>
<tr><td rowspan="2">Gender</td><td colspan="2">Male</td><td>Female</td></tr>
<tr><td colspan="2">45.5%</td><td>54.5%</td></tr>
<tr><td rowspan="2">Program</td><td colspan="2">MSc</td><td>M.Phil</td></tr>
<tr><td colspan="2">69.1%</td><td>30.9%</td></tr>
<tr><td rowspan="2">Region</td><td>Rural</td><td>Urban</td><td>Suburban</td></tr>
<tr><td>52.7%</td><td>27.2%</td><td>20.1%</td></tr>
</table>

Table A-4
Percentages, Mean and Standard Deviation of all Ordinal Scale Variables

Factor/Variables	SD%	DA%	N%	A%	SA%	Mean	S.D
Provide Structure / Conduct Management							
Friendly Behavior	4.0	3.0	4.0	56.5	32.5	4.11	0.96
Friendly Environment	2.0	4.0	6.0	51.0	37.0	4.17	0.83
Student-Teacher Interaction	3.5	2.5	8.0	49.5	36.5	4.13	0.92
Few minutes break	5.5	5.0	22.0	31.0	36.5	3.88	1.12
Encouragement by teacher	5.5	4.5	10.0	47.5	32.5	3.97	1.05
Opportunities satisfaction	5.0	5.0	23.0	49.5	17.5	3.70	0.93
Overall	4.25	4.0	12.16	47.5	32.08	3.99	0.96
Teaching Methodology and effectiveness							
Difficulty in class work	2.5	29.5	24.5	29.0	14.5	3.24	1.10
Complete Class work at Time	9.5	10.5	7.5	46.0	26.5	3.70	1.23
Teachers' ability to make course Martial	9.0	8.5	9.0	50.0	23.5	3.71	1.81
Positive Relation of Strict Environment	5.0	18.0	19.5	36.0	21.5	3.51	1.61
Effect of checking Assignments & Quizzes	4.0	3.5	9.5	49.5	33.5	4.05	0.96
Evaluation Process	5.5	6.0	42.0	30.5	16.0	3.46	1.01
Overall	5.91	12.66	21.91	37.16	22.58	3.61	1.08
Teachers Attitude (In class)							
Equal importance	12.0	14.0	11.0	27.5	35.5	3.61	1.40
Opinions Sharing	9.5	5.5	11.5	54.5	19.0	3.68	1.13
Environment	8.0	6.0	17.5	36.5	32.0	3.79	1.19
Teachers' experience	8.5	5.0	11.0	41.0	34.5	3.88	1.18
Supportive Behavior	3.0	8.0	10.5	46.5	32.0	3.97	1.00
Teachers' Expectations	7.5	8.5	14.5	40.0	29.5	3.76	1.18
Overall	8.08	7.83	12.66	41	30.41	3.78	1.08
Motivation & feedback from Teacher							
Encouragement effect on performance	9.7	31.7	22.6	25.1	10.9	2.95	1.18
Activities	4.6	30.6	29.1	30.9	4.9	3.00	0.99
Encouragement effect on Creativity	2.9	10.6	19.4	57.4	9.7	3.60	0.91
Encouragement effect on Goal	3.1	10.3	14.3	59.7	12.6	3.68	0.93
Feedback from Teacher	2.0	7.4	20.3	49.1	21.1	3.80	0.92
Feedback Effect on Performance	5.1	13.4	26.3	41.7	13.4	3.45	1.05
Overall	4.6	17.3	22.0	44.0	12.1	3.4	1.1
Teachers' Behavior /Attitude (Outside the class)							
Meeting with teachers	6.0	12.6	21.4	50.0	10.0	3.45	1.03
Discussing the course material	4.3	16.0	25.1	46.3	8.0	3.38	0.98
Solution of learning problems	4.6	19.7	25.4	41.7	8.6	3.29	1.02
Students communication	5.7	22.9	25.4	36.9	9.1	3.20	1.07
Meeting with teachers outside the class	2.6	8.9	20.0	54.0	14.6	3.69	0.92
Performance Improvement	2.9	6.9	13.1	52.6	24.6	3.89	0.95
Overall	4.3	14.5	21.8	46.9	12.5	3.5	1.0

APPENDIX-B

(a-Formal Factors)

Table B-1
Model estimate of Confirmatory Factor Analysis of Class Management

Variables	Parameter	Standard	T	Prob.
Class Management → Behavior	0.478	.050	4.185	0.000
Class Management → Friendly Environment	0.525	.074	5.177	0.000
Class Management → Student-Teacher Interaction	0.207	.080	2.445	0.000
Class Management → Few minutes break	0.589	.065	5.537	0.000
Class Management → Encouragement by teacher	0.732	.061	6.064	0.000
Class Management → Opportunities satisfaction	0.643	.065	5.790	0.000

Table B-2 Measure of Goodness of Fit of Class Management

	χ^2	d.f	p-value	$\chi^2/d.f$	GFI	AGFI	RMSEA
Class Management	21.509	9	0.0000	2.389	.957	0.915	0.061
Recommended				≤ 3	≥ 0.90	≥ 0.90	≤ 0.08

Table B-3
Model estimate of Confirmatory Factor Analysis
of Teaching Methodology and Effectiveness

Variables	Parameter	Standard	T	Prob.
Teaching Methodology and effectiveness → Difficulty in class work	0.285	.058	6.169	0.000
Teaching Methodology and effectiveness → Complete class work at time	0.590	.087	10.724	0.000
Teaching Methodology and effectiveness → Ability to make course martial	0.723	.084	6.742	0.000
Teaching Methodology and effectiveness → Influence of strict Environment	0.344	.082	4.506	0.000
Teaching Methodology and effectiveness → Effect of checking Assignment and Quizzes	0.285	.068	6.069	0.000
Teaching Methodology and effectiveness → Evaluation Process	0.140	.072	11.492	0.000

Table B-4
Measure of Goodness of Fit of Teaching Methodology and Effectiveness

	χ^2	d.f	p-value	$\chi^2/d.f$	GFI	AGFI	RMSEA
Teaching Methodology and effectiveness	25.483	9	0.000	2.831	.959	.910	0.016
Recommended				≤ 3	≥ 0.90	≥ 0.90	≤ 0.08

Table B-5
Model estimate of Confirmatory Factor Analysis of Teachers' Attitude (In class)

Variables	Parameter	Standard	T	Prob.
Teachers' Attitude → Equal importance	0.610	0.078	8.982	0.00
Teachers' Attitude → Opinions Sharing	0.923	0.096	9.571	0.00
Teachers' Attitude → Environment	0.935	0.100	9.351	0.00
Teachers' Attitude → Teachers' Experience	0.361	0.094	3.836	0.00
Teachers' Attitude → Supportive Behavior	0.253	0.080	3.178	.001
Teachers' Attitude → Teachers' Expectations	0.860	0.098	8.782	0.00

Table B-6
Measure of Goodness of Fit of Teachers' Attitude (In class)

	χ^2	d.f	p-value	$\chi^2/d.f$	GFI	AGFI	RMSEA
Class Management	9.674	9	0.000	1.074	0.962	0.901	0.03
Recommended				≤ 3	≥ 0.90	≥ 0.90	≤ 0.08

(b-Informal Factors)

Table B-7
Model estimate of Confirmatory Factor Analysis of Motivation & feedback from Teacher

Variables	Parameter	Standard	T	Prob.
Motivation & feedback → Encouragement effect on performance	1.580	.342	4.627	0.000
Motivation & feedback → Activities	.704	.217	3.241	0.001
Motivation & feedback → Encouragement effect on creativity	1.683	.360	4.678	0.000
Motivation & feedback → Encouragement effect on Goal	2.389	.500	4.783	0.000
Motivation & feedback → Feedback from Teacher	1.480	.342	4.627	0.000
Motivation & feedback → Feedback effect on performance	.698	.239	2.927	0.003

Table B-8
Measure of Goodness of Fit of Motivation & Feedback from Teacher

	χ^2	d.f	p-value	$\chi^2/d.f$	GFI	AGFI	RMSEA
Class Management	7.6632	9	0.000	0.851	0.981	0.942	0.06
Recommended				≤ 3	≥ 0.90	≥ 0.90	≤ 0.08

Table B-9
Model estimate of Confirmatory Factor Analysis
of Teachers' Behavior /Attitude (Outside the class)

Variables	Parameter	Standard	T	Prob.
Teachers' Behavior(Outside the class) → Meeting with Teachers	0.916	0.140	4.980	0.00
Teachers' Behavior(Outside the class) →Discussing Course martial	0.608	0.182	3.345	0.00
Teachers' Behavior(Outside the class) →solution of learning problems	0.975	0.226	4.323	0.00
Teachers' Behavior(Outside the class) →Students communication	0.966	0.240	4.030	0.00
Teachers' Behavior(Outside the class) →Meeting with teachers outside the class	0.524	0.172	3.047	0.002
Teachers' Behavior(Outside the class) →Performance Improvement	0.619	0.074	0.250	0.002

Table B-10
Measure of Goodness of Fit of Teachers' Behavior /Attitude (Outside the class)

	χ^2	d.f	p-value	$\chi^2/d.f$	GFI	AGFI	RMSEA
Class Management	25.01	9	0.000	2.77	0.945	0.90	0.01
Recommended				≤ 3	≥ 0.90	≥ 0.90	≤ 0.08

APPENDIX-C

Table C-1
Model estimate of Structure Equation Modeling

Variables	Parameter	Standard	T	Prob.
Teaching Methodology and effectiveness →CGPA	1.223	0.025	17.216	0.00
Teaching Behavior →CGPA	1.201	0.235	14.342	0.00
Teaching Behavior →Teaching Methodology	1.234	0.034	16.822	0.00
Motivation and feedback →CGPA	1.00	0.012	7.925	0.00
Motivation and feedback →Teaching methodology	1.053	0.011	15.066	0.00
Class management →Teaching Methodology	1.025	0.035	25.521	0.00
Teachers' attitude (in class) →Class Management	0.998	0.021	11.251	0.00
Motivation and feedback →Teaching Methodology	0.962	0.037	20.235	0.00
Class Management →CGPA	1.521	0.035	18.978	0.00

Table C-2
Measure of Goodness of Fit of Model

	χ^2	d.f	p-value	$\chi^2/d.f$	GFI	AGFI	RMSEA
Class Management	626.558	267	0.0000	2.349	.927	0.898	0.061
Recommended				≤ 3	≥ 0.90	≥ 0.90	≤ 0.08

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 267-273

COMPARISON OF CLASSIFICATION TECHNIQUES FOR SENTIMENT ANALYSIS OF GLOBAL WARMING

Zaheer Abbas, Erum Shahzadi and **Amna Qadeer**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: zaheerabbas@uog.edu.pk

ABSTRACT

Any judgment of an individual through which the feelings, attitudes and thoughts can be expressed is known as sentiment. Sentiment analysis derives emotional information from text and classifies it either positive, negative or neutral. Global warming is the gradual increase in the average temperature of earth's atmosphere. A big amount of the heat colliding the earth from the sun is being trapped in the atmosphere and not scattered out into space. The earth's climate has changed throughout the history. There's the old saying that you should never discuss politics or religion in polite company. Nowadays, it seems climate change has joined that list. It is the need of an hour to get mass opinion in this regard. Internet based life have gotten more consideration these days. In this study, a large dataset of geotagged tweets containing certain keywords relating to climate change is analyzed. R language is used in this research to implement the classification algorithm on the collected data. I have trained Naive Bayes Classifier; Maximum Entropy Classifier and SVM classifier on several training datasets to optimize for best accuracy. SVM classifier got high accuracy as compared to the rest of classifiers.

KEYWORDS

Sentiment Analysis, NLP, Support Vector Machine, Naive Byes Classifier, Maximum Entropy.

1. INTRODUCTION

Global warming is the latest topic of this century as very tragic climatic events keep on damaging the whole planet. The term global warming and green house effect are analogue with each other. The indication of global warming includes heat wave, sea- level rise, flooding, melting glaciers, earlier spring arrival and the spread of disease. Machine learning and statistical techniques are often used in the study of twitter data for diagnoses of sentiments or opinions . The study of social media data can determine the actual time awareness into world developments by giving a stage for users to explicit their opinions, talk on important happening or broadly take up on different topics of interest.

For this research we concentrate on Twitter that is the freely accessible platform. A lot of researchers have favorably integrated on Twitter data in many research areas containing flu forecasting, political events and community anarchy. Tweets are also called as the micro blog owing to its short text. Micro blogging websites have emerged to enhance origin of different type of facts. Opinions are important to nearly all human task because they are important tastemakers of our attitudes. When we have to take some decisions we desire to know other opinions. Business and organization always need to discover consumer or public opinions related to their productions and supply. Sentiment analysis is generally used to reviews and social media for different type of applications roaming from marketing to customer service.

Basically, sentiment analysis or sentiment classification take place into the extensive sort of classification task. In this task we are supplied with a phrase or list of phrases and our classifier is assumed to narrate what is the sentiment in the back ground of this phrase or phrases. Nonetheless in this research we will concentrate on personal website data like Twitter. The intention of this research is to find out in what way people's view point about global warming has changed over time. Different classification techniques like Naive Bayes Classifier, Maximum Entropy Based classifier and Support Vector Machine(SVM) will be applied to classify different perception related to global warming.

2. LITERATURE REVIEW

Much work has been done in the area of sentiment analysis on twitter by many researchers in recent years. At start it was designed for paired classification that empowers opinions or reviews to two-faced classes such as positive or negative only.

Parik and Movassate [1] executed two models Naive Byes model and Maximum Entropy model to categorize tweets. They raised the fact that Naive Bayes classifier processed enough superior than the maximum entropy model.

Go and L. Huang [2] planned an explanation for sentiment analysis of twitter data by utilizing specific supervision. In this research their training data contained of tweets with emoticons which dealt as boisterous tags. They formed models using Naive Bayes, Maximum Entropy and Support Vector Machine(SVM). Their features slot consisted of unigrams and bigrams. They determined that SVM surpassed other models. They also concluded that unigrams were more adequate as features.

Pak and Paroubek [3] recommended a model to classify the tweets as objective, positive and negative. They designed a twitter corpus by compiling Tweets using twitter API. Working on that body of text they refined a sentiment classifier established on the multinomial naive bayes method. The training set they utilized was little adequate because it included only those tweets containing emoticons.

Barbosa [4] created a two stage automated sentiment analysis approach for categorizing tweets. They categorized tweets as subjective or objective. At second stage they categorized the subjective tweets as positive or negative. Bifet and Frank [4] utilized twitter date administered by firehouse API, which provide all messages from each users which are openly accessible in actual time. They examined multinomial bayes, stochastic gradient descent (SGD) and Hoeffiding tree models. They reached at a decision that SGD model performed well enough as compared to other above mentioned models.

Anurag [5] introduced suggestion for analyzing the sentiments of users applying data mining classifiers. Furthermore it measures the accomplishment of distinct classifier for sentiment analysis of assemblage classifiers. Their results explained that k-nearest neighbor classifier provides great anticipating certainty as compared to other models particularly Random Forest and Naive Bayes models. It was also concluded from the results of test data that data mining classifiers is a good decision for sentiments prediction of twitter data.

3. VARIOUS LEVELS OF SENTIMENT ANALYSIS

In generic, sentiment analysis has been considered at three main levels. Document level, Sentence level and Entity and aspect level. Sentiment analyses has been organized as a Natural language processing (NLP) function through frequent levels of communication. Beginning from being a document level classification job (Pang and Lee [6]) it has been directed at the sentence level (Kim and Hovy [7]) and further freshly at the phrase level (Agarwal [8]). Different social media platforms such as twitter, instagram, facebook and flicker cater researchers with abundant volume of data. Many researches in analysis of social media data report on three main groups: opinion mining corresponding to human behavior, grouping user association, and sentiment and emotion classification. Some researchers also uses social media data to depict a particular use instance, one of which is global warming. Social media has been used to comprehend human attitude over frequent areas of expertise including healthcare (Paul [9]), politics(Wang [10]) and crisis management (Pathak [11]). To discover emotions and opinions in social media researchers have utilized trained model. There has been an arising concern in examining human behavior toward climate change by the way of social media.

The explanation of these levels is as follow:

- **Document Level**

The function at this level is to classify either an entire opinion document explicit a positive or negative sentiment. For instance, provided with different tweets or reviews, the system decided either the review suggested a complete positive or negative opinion related to product or an issue. This project is usually known as document-level sentiment classification. This level of analysis infers that every document gives opinions on a distinct substance e.g. on a single product or issue.

- **Sentence Level**

The function at this level enlists the sentences and decides either each sentence signified a positive, negative or neutral opinion. Neutral generally suggests no opinion. This level of search is approximately connected to subjectivity classification. The document is not anything but the accumulation of the sentences composed. The accuracy of sentence level analysis is enough improved as compared to document level.

- **Entity and Aspect level**

This level of sentiment analysis executes admirable analysis. In place of looking at the language manufactures (documents, paragraphs, clauses, sentences or phrases), aspect level precisely examines the opinion itself. This is established at the concept that an opinion subsists of a sentiment and a target. This level gives enough better results than other two above mentioned levels. The document level and the sentence level analysis do not find what precisely people appreciate or not appreciate.

4. DATA PRE PROCESSING AND CLASSIFIERS

Twitter data related to global warming issue by Kent Cavender-Bares. It is downloaded from figure-eight.com webpage. Various twitter entries were evaluated for belief in the existence of global warming or climate change. The possible answers were "Yes" if the tweet suggests global warming is occurring, "No" if the tweet suggests global warming is not occurring. Data preprocessing is a process to remove the unwanted words from tweets that does not amount to any sentiments. Emotional Icons identifies emotional icons and remove them. URLs-does not signify any sentiment. Replaced it with a word

|URL|. Stop words- words as "a‟, "is", "the" does not indicate any sentiment. User Names and HashTags- @ symbol before the username and # for topic, both replaced with AT_USER. . Repeated Letters- climatee, climaaat, climaatee into the token "climate". Slag Words were also removed. We have trained Naive Bayes Classifier, Support Vector Machines Classifier and Maximum Entropy Classifier on training dataset to optimize for best accuracy.

NAIVE BAYES CLASSIFIER

Naive Bayes Classifier is a basic classifier in machine learning. It is based on Bayes theory assuming all variables are independent. According to Bayes theorem

$$P\left(\frac{A}{B}\right) = \frac{P\left(\frac{B}{A}\right)P(A)}{P(B)}$$

The terms in above expression can be defined as follow

- P(A) = Probability of A
- P(B) = Probability of B
- $P(\frac{A}{B})$ = Probability of A given B
- $P(\frac{B}{A})$ = Probability of B given A.

SUPPORT VECTOR MACHINE CLASSIFIER

Support Vector Machines is an effective machine learning algorithm used for regression and classification purposes. It segregates classes by finding the best hyper plane between different classes. Linear support vector classification is analogous to support vector classification using the linear algorithm. Though linear support vector machines are supplementary elastic in consequences and losses. Hence support vector machine acts proficiently for bigger datasets.

MAXIMUM ENTROPY CLASSIFIER

Maximum Entropy is a machine learning method established for numerical data. Alternate of naive Bayes classifiers. It does not assume statistical independence of the random variables (commonly known as features) that serve as predictors. The classification system is well described by Adwait Ratnaparkhi [12] as:

"….Maximum Entropy models offer a way to combine diverse pieces of contextual evidence in order to estimate the probability of a certain linguistic class occurring with a certain linguistic context….in which task is to estimate the probability of class 'a' occurring with context 'b'….." .

CONFUSION MATRIX

A confusion matrix is a technique for summarizing the performance of a classification algorithm. Classification accuracy alone can be misleading if we have an unequal number of observations in each class or if we have more than two classes in your dataset. Calculating a confusion matrix can give us a better idea of what our classification model is getting right and what types of errors it is making. Confusion matrix obtained by applying above mentioned classifiers are as follow

Confusion matrix for SVM Classifier

Predicted Opinion	**Total Population**	**Actual Opinion**	
		Actual Yes	**Actual No**
	Predicted Yes	142 (TP)	187 (FP)
	Predicted No	48 (FN)	680 (TN)

Sensitivity, recall or true positive rate = $\frac{TP}{TP+FN}$ = 74%

Specificity, selectivity or true negative rate = $\frac{TN}{TN+FP} = 78\%$

Accuracy =$\frac{TP+TN}{TP+TN+FP+FN} = 77\%$.

Confusion matrix for Maxent Classifier

Predicted Opinion	**Total Population**	**Actual Opinion**	
		Actual Yes	**Actual No**
	Predicted Yes	182 (TP)	147 (FP)
	Predicted No	131 (FN)	597 (TN)

Sensitivity, recall or true positive rate = $\frac{TP}{TP+FN}$ = 58%

Specificity, selectivity or true negative rate = $\frac{TN}{TN+FP} = 80\%$

Accuracy =$\frac{TP+TN}{TP+TN+FP+FN} = 73\%$.

Confusion matrix for Naive Bayes Classifier

Predicted Opinion	**Total Population**	**Actual Opinion**	
		Actual Yes	**Actual No**
	Predicted Yes	280 (TP)	205 (FP)
	Predicted No	126 (FN)	657 (TN)

Sensitivity, recall or true positive rate = $\frac{TP}{TP+FN}$ = 69%

Specificity, selectivity or true negative rate = $\frac{TN}{TN+FP} = 76\%$

Accuracy =$\frac{TP+TN}{TP+TN+FP+FN} = 75\%$.

5. RESULTS AND CONCLUSION

We have trained above mentioned classification algorithms in R Studio with 3171 tweets (75% of original data) and tested on test data set of 1056 tweets (25% of original data). The following results are obtained by using different classifiers. Thus it is concluded that support vector machine classifier performs well as compared to rest classifiers for our dataset.

Algorithm i	**Accuracy i**
NaïveiBayes i	75% i
SVM i	77% i
Maxent i	73% i

6. ACKNOWLEDGEMENT

Figure Eight is the essential Human-in-the-Loop machine learning platform. They help the most innovative companies in the world with their most ambitious projects. We would like to thank Figure Eight team to provide us data on global warming. We are very grateful for all the support and guidance that we have received during the research process**.**

REFERENCES

1. Parikh, R. and Movassate, M. (2009). Sentiment analysis of user-generated twitter updates using various classification techniques. *CS224N Final Report*, 118.
2. Go, A., Bhayani, R. and Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report, Stanford*, *1*(12), p.2009.
3. Barbosa, L. and Feng, J., (2010). Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd international conference on computational linguistics: posters* (pp. 36-44). Association for Computational Linguistics.
4. Bifet, A., Holmes, G., Pfahringer, B. and Frank, E. (2010). Fast perceptron decision tree learning from evolving data streams. In *Pacific-Asia conference on knowledge discovery and data mining* (pp. 299-310). Springer, Berlin, Heidelberg.
5. Jain, A.P. and Katkar, V.D. (2015). Sentiments analysis of Twitter data using data mining. In *2015 International Conference on Information Processing (ICIP)* (pp. 807-

810), IEEE.
6. Pang, B. and Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends® in Information Retrieval*, 2(1-2), 1-135.
7. Kim, S.M. and Hovy, E. (2006). Extracting opinions, opinion holders, and topics expressed in online news media text. In *Proceedings of the Workshop on Sentiment and Subjectivity in Text* (pp. 1-8). Association for Computational Linguistics.
8. Agarwal, A., Biadsy, F. and Mckeown, K.R. (2009). Contextual phrase-level polarity analysis using lexical affect scoring and syntactic n-grams. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 24-32). Association for Computational Linguistics.
9. Paul, M.J. and Dredze, M. (2017). Social monitoring for public health. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 9(5), 1-183.
10. Vaziripour, E., Giraud-Carrier, C. and Zappala, D. (2016). Analyzing the political sentiment of tweets in Farsi. In *Tenth International AAAI Conference on Web and Social Media (ICWSM 2016)*, 699-702.
11. Pathak, N., Henry, M.J. and Volkova, S. (2017). Understanding Social Media's Take on Climate Change through Large-Scale Analysis of Targeted Opinions and Emotions. In *2017 AAAI Spring Symposium Series*.
12. Ratnaparkhi, A. (1997). A simple introduction to maximum entropy models for natural language processing. *IRCS Technical Reports Series*, p.81.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 275-282

EATING BEHAVIOR AND DIETARY HABITS AMONG FEMALE STUDENTS OF UNIVERSITY OF GUJRAT

Zaheer Abbas[§] and **Rabia Mazher**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: zaheerabbas@uog.edu.pk

ABSTRACT

Food is an essential part of every one's lives. For many people, food choices are very complex and complicated decision. The study was used to explore the eating behavior and dietary intake among students of university of Gujrat.

This study comprised on 300 female students of university of Gujrat. It was a cross sectional study and data is collected by stratified random sampling by using proportional allocation. Statistical technique like neural network, confirmatory factor analysis and structural equation modeling are applied on the data.

Analysis results revealed that all the factors have significant effect on eating behavior of students. Result of neural network showed that Students eating behavior is greatly impacted by university life style, body mass index and knowledge of nutritional diet. Results of SEM (Structural equation modelling) showed that there is perfect association between eating habit and knowledge of nutritional diet. It was concluded that most of the students prefer to eat unhealthy food because of tough routine of university and lack of nutritional knowledge.

1. INTRODUCTION

Food is basic need of human and essential part of every one's lives. Diet that contain sufficient amount of nutrition is essential for health. Eating behavior is defined as dietary pattern and choice of food that a person consumes (Patel, et al. 2013). In the recent years, health organization focus on people healthy eating behavior, in result, consumption of nutritive food has been increasing day by day (Scaglioni et al. 2018). Nutritional diet have a major role on the health of a person and prevent illness. It contains supplements and vitality to grow up. Healthy diet is considered as most important component that provides energy to youngster so that they can play their role for development of society (Naeeni et al. 2014).

Now days, People are looking for a few sustenance decisions every day and their food choice dependent on a few criteria like individual perception and knowledge of nutritional diet etc. (Oti 2018). For many people, nourishment choices are very complex and complicated decisions because every person have different taste and choice of food (Deliens et al. 2014).

In this era of fast food, it is very difficult for youngster to eat healthy diet, because they like to eat spicy food. At the stage, when they move from school to college then a big change in their eating habit occur because they are not bound to take home made lunch.

Therefore, they try to eat from cafeteria and this could have negative affect on students eating behavior (Shannon et al. 2002).

Lack of nutrition knowledge is the main reason that people take interest in unhealthy food. Due to unhealthy eating habit our society is facing medical problems like anemia, blood pressure and obesity (Spronk et al. 2014). It is necessary to aware people about the food that contain high amount of nutrition. People change their eating behavior and become diet conscious when they gain information from concealing sessions, seminars and other major sources of awareness. The main purpose of these events should be to provide knowledge about nutrition (Abraham et al. 2018).

3. MATERIAL AND METHODS

The sample was selected by stratified random sampling with proportional allocation. Data was collected from 300 female students of university of Gujrat. This was a cross-sectional study. Student was selected from different departments like faculty of sciences, faculty of arts, faculty of social sciences, faculty of computing and information etc. A questionnaire was constructed to collect data from the respondent.

Data analysis techniques

For the data analysis we use different software and packages of the statistics. The software that are used include SPSS was used for neural network and descriptive statistics, statistical AMOS for SEM and STATISTICA for confirmatory factor analysis.

Study dependent variable was eating habit. Predictors of this study were individual perception, nutrition knowledge, social impact, environmental impact, economic status, university life style and psychological behavior.

4. COMMENTS AND CONCLUSION

According to the descriptive analysis 50 percent of our respondent (girls) belong to rural area and other 50 percent respondent (girls) belongs to urban area. The students who belongs to BS program are 83.7%, students that belongs to MSC program are 11.7% and 4.6% of student’s are doing MPhil. Percentage of participant belong to nuclear family system (73.3) are greater than the joint family system (26.3). Student who take breakfast are 62.6% whereas 37.3% don’t take breakfast in daily routine. Student who take breakfast from home are 88% whereas 12 % students take breakfast from university.

The results of confirmatory factor analysis show that different measures that include p value, GFI, AGFI, and RMSEA that are used to demonstrate factors are significant or not. The first factor is eating habit of students. P value of eating habit is 0.000 which is less than 0.05, chi.sq/d.f value is 2.88 which is less than 3, GFI and AGFI values are 0.95 and 0.89 respectively that is close to 0.9. RMSEA value is 0.01 that is less than 0.08. This shows that factor of eating habit is confirmed. All the other factors of this study are confirmed because they meet to them require criteria. So it means our model is best fitted.

Structural equation modelling results shows that the measures of goodness of fit. If P-value of chi square value is less than 0.05, GFI and AGFI values are close to or greater than 0.90 and RMEA is less than 0.08, then our model is best fitted. In this model of eating habit p value of chi square value is 0.00. GFI and AGFI values are 0.855 and 0.830 respectively which are close to 0.90. RMSEA value is 0.054, which indicate that our model is fitted adequately.

Eating habit = 0.085 University (Uni) (1)

Eating habit = 0.469 Psychology (Psy) (2)

Eating habit = 0.780 knowledge (kno) (3)

Eating habit = 0.129 Social (So) (4)

Combining the equation (1), (2), (3) and (4)

Eating habit = 0.085(Uni) + 0.469(Psy) + 0.380(Kno) + 0.129(So)

Figure 1 shows the direct effect of university life style, psychological behavior, nutrition knowledge and social impact on the eating habit.

Soc → Env → Kno → Eco → Ind → Uni → EH = 0.570

Figure 1 show the indirect relation of social impact, environmental impact, nutrition knowledge, economic status, individual perception and university lifestyle on eating habit of students.

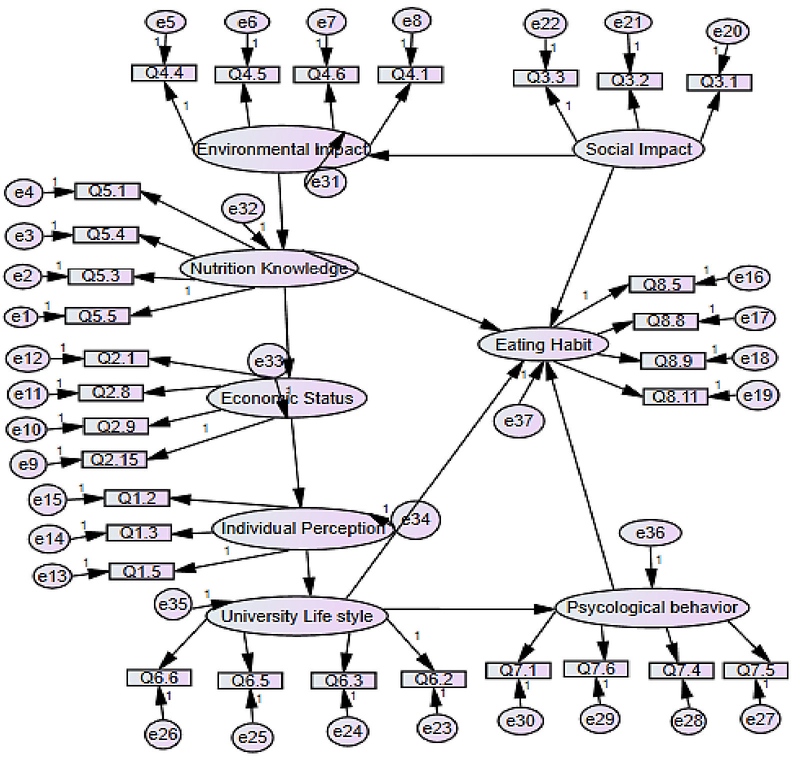


Figure 1: Path Diagram of SEM of Eating Habit

In classification matrix, diagonal cells represent correct classification and off diagonal cells represent incorrect classification. Table 4.1 represent that in training sample 100 respondents out of 112 respondents were correctly classify as they have unhealthy eating habit. The 89.3 percent of respondent are correctly classified whose diet was unhealthy. In case of healthy eating pattern 89 cases out of 96 respondents are correctly classify. This indicate that most of the student prefer to eat unhealthy diet. In testing sample 38 out of 44 cases are correctly classify in unhealthy eating behavior. Overall correct classification for both training and testing sample was 90.9% and 89.1% respectively. Correct percent for testing sample is less than training sample, it indicates that our model is good.

Table 4.1
Classification

		Predicted		
Sample	**Observed**	**Unhealthy**	**Healthy**	**Correct Percentage**
Training	Unhealthy	100	12	89.3
	Healthy	7	89	92.7
	Overall Percentage	51.4	48.6	90.9
Testing	Unhealthy	38	6	86.4
	Healthy	4	44	91.7
	Overall Percentage	45.7	54.3	89.1

Dependent Variable: Eating Habit

In Figure 2, the area under the ROC curve is greater than 0.90 % this shows that our model is good.

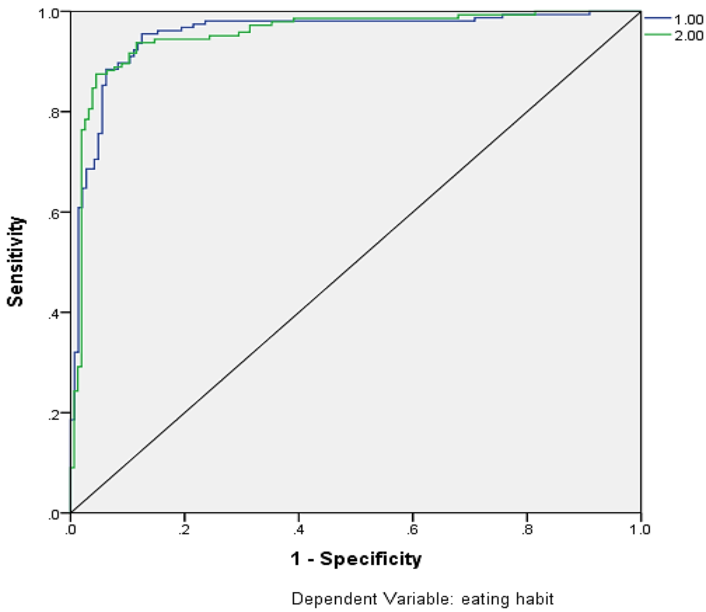


Figure 2: ROC Curve

The Figure 3 represent the importance of independent variable. All of independent variables are important because they have effect on eating habit of students. Nutrition Knowledge was most important factor in selection of food because it fully participates (100%) in classification of eating habit. Psychological behavior was second important variable with 60 % of participation. Residence was least important factor in selection of food with 4.8 % participation in classification.

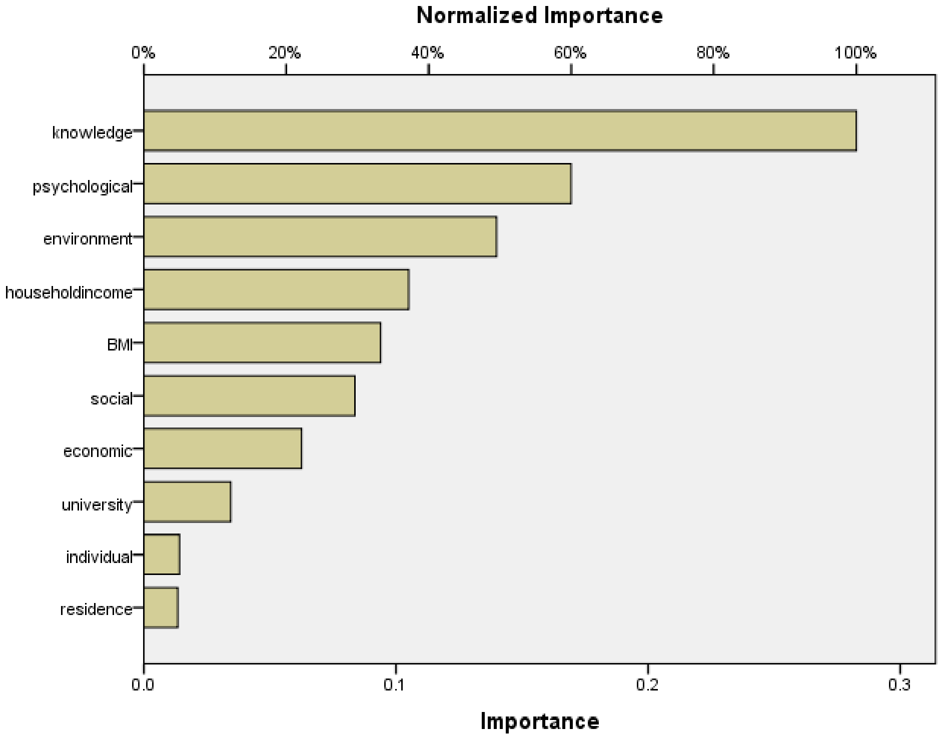


Figure 3: Importance of Independent Variable

World and Health Organization 2015 stated that, the different varieties of food open doors for an individual to change in their eating pattern. Demand of healthy and nutritional diet increased in college student's because it is the stage of rapid growth in their body. Therefore, it is essential to focus on child eating pattern because inadequate nutrition not only diminish physical development as well as the learning capacity of students (Abraham, et al. 2018).

The level of knowledge regarding the nutritional value of food items played an important role in developing eating habits. Our analysis showed the students had poor nutritional knowledge which indicated that they are consuming more fat-like substance. This analysis is concordant with other studies; for example, Ozgen (2016) observed a low level of nutritional knowledge in the agricultural university. Ping et al.

(2016) study also observed that university students in China possessed poor knowledge regarding healthy diets and developed an increased tendency of eating outside. The information acquired through our study suggested that many students' eating habits and dietary intake was notably influenced by academic activities, especially exam pressures. Studies in German universities observed that study pressure reduced the time available for cooking, which lead to unhealthy eating among the students Kabir et al. (2018). As the majority of the students belong to lower middle-class families, and thus have a limited financial capacity to afford food. Ranjbar et al. (2017) study also present similar views that economic status is associated with food habits. Our study revealed that taste preference plays important role to determine eating habit among students. Another study suggest that taste preference plays a central role in food selection Drewnowski et al. (2005). Our analysis suggested that societal factors influence students' eating behavior and food choice. Students' food choice was clearly influenced by their social network (roommates, and friends) in and around the campus. Previous studies support our analysis; for example, Contento et al. (2006) reported that peers influence has a great effect on food selection.

CONCLUSION

This study analyzes the eating behavior of university students by using different independent factors. Descriptive statistics shows that 50 % of respondent belong to urban and 50 % belongs to rural area. 62.6 % respondent take breakfast. While, 37.3% respondent don't take breakfast. This study shows that effect of all the factors on eating habit is significant. The results of CFA show all the items are confirmed under the same pattern. The results of SEM shows that eating behavior of student greatly depend on knowledge of nutritional diet and university life style of an individual. An overall Findings of the neural network shows that nutritional knowledge, body mass index, psychological behavior and university life style are most important in classifying the variable of eating behavior. It was concluded that most of the students prefer to eat unhealthy food because of tough routine of university and lack of nutritional knowledge. In other words, this study illustrates that the people who prefer to the healthy food has more nutrition knowledge, appropriate BMI, Face minimum academic burden and mental problem.

RECOMMENDATION

This study helps out in future to increase nutrition knowledge and decrease burden of academic activities among students to improve their dietary plan. The teachers and parents should bound the youngsters to eat healthy diet. Students should drink at least 8 glasses of water and take exercise regularly. We should have to increase knowledge of nutritional diet among students by arranging seminars and conferences.

The university cafeteria should offer fresh vegetables and fruits and less varieties of fast food to the students.

REFERENCES

1. Abraham, S., Noriega, B.R. and Shin, J.Y. (2018). College students eating habits and knowledge of nutritional requirements. *Journal of Nutrition and Human Health*, 2(1), 13-17.
2. Ansah, C.O. (2006). Food habits and preferences as a factor in the choice of meals by students in the University Cape Coast. *Journal of Nutrition and Health*, 21(3), 151-172.
3. Bielak, M.J., Dolowy, A.K., Santoro, A., Ostan, R., Berendsen, A.M., Jennings, A., Meunier, A., Caumon, E., Lisette, C.P.M. and Pietruszka, B. (2017). Nutrition-Related Knowledge and Attitudes Reflected in Lifestyle and Health Among Elderly People. *Journal of Nutrition Education and Behavior*, 4(3), 184-193.
4. Bin Zaal, A.A., Musaiger, A.O. and Souza, R.D. (2007). Dietary habits associated with obesity among adolescents in Dubai, United Arab Emirates. *European Journal of Clinical Nutrition*, 24(4), 437-444.
5. Contento, I.R., Williams, S.S., Michela, J.L. and Franklin, A.B. (2006). Understanding the food choice process of adolescents in the context of family and friends. *Journal of Adolescent Health*, 38(5), 575–582.
6. Deliens, T., Clarys, P. and Deforche, B. (2014). Determinants of eating behavior in university students: a qualitative study using focus group discussions. *Journal of Nutritional Education and Behavior*, 14, doi https://doi.org/10.1186/1471-2458-14-53.
7. Drewnowski, A. (2005). Energy intake and sensory properties of food. *Am. J. Clin. Nutr.*, 62(5), 1081-1085.
8. Ganasegeran, K., Al-Dubai, S.A., Qureshi, A.M., Al-abed A.M.R. and Aljunid, S.M. (2012). Social and psychological factors affecting eating habits among university students in a Malaysian medical school: a cross-sectional study. *Nutrition Journal*, 11(1), 11-48.
9. Hernandez, J., Bamwesigye, D. and Horak, M. (2016). Eating Behaviors of University Students. *American Journal of Public Health*, 100(2), 216-22.
10. Kabir, A., Miah, S. and Islam, A. (2018). Factors influencing eating behavior and dietary intake among resident students in a public university in Bangladesh: A qualitative study. *The American Journal of Clinical Nutrition*, 63(6), 997-1001.
11. Lee, S.Y., Ha, S.A., Seo, J.S., Sohn, C.M. and Kim, K.W. (2014). Eating habits and eating behaviors by family dinner frequency in the lower-grade elementary school students. *Journal of Nutrition Research*, 8(6), 679-687.
12. Naeeni, M.M., Jafari, S., Fouladgar, M., Heidari, K., Farajzadegan, Z., Fakhri, M., karami, P. and Omidi, R. (2014). Nutritional Knowledge, Practice, and Dietary Habits among school Children and Adolescents. *International Journal of Preventive Medicine*, 5(2), 171-178.
13. Ozgen, L. (2016). Nutritional Knowledge, Attitudes and Practices among University Students in Turkey and the US. *Journal of the American Dietetic Association*, 26(3), 158-166.
14. Oti, J.A. (2018). Perception of public and private junior high school adolescent student in asiakwa in the eastern region of Ghana. *Journal of Food and Nutrition*, 3(1), 103.
15. Park, E.S., Lee, J.H. and Kim, M.H. (2015). Eating Habits and Food Preferences of Elementary School Students in Urban and Suburban Areas of Daejeon. *Journal of Nutrition Education*, 4(3), 190-200.

16. Patel, H., Solanki, H., Gosalia, V., Vora, F. and Singh, M.P. (2013). A Study of Awareness of Nutrition & Anaemia among College Going Students of Mahila College of Bhavnagar. *National Journal of Community Medicine*, 4(2), 300-30
17. Ping, H., Huang, W., Bai, R., Zhang, F., Sharma, M. and Shi, Z. (2016). Knowledge, Attitude, and Behaviors Related to Eating Out among University Students in China. *Int. J. Environ. Res. Public Health*, 13(7). 696.
18. Ranjbar, K., Barzegari, A., Ebrahimi, M. and Azizi, M. (2017).A Study of Nutrition Knowledge, Attitudes and Food Habits of College Students. *World Applied Sciences Journal*, 15(7), 1012-1017.
19. Scaglioni S., De Cosmi, V., Ciappolino, V., Parazzini, F., Brambilla, P. and Agostoni, C. (2018). Factors Influencing Children's Eating Behaviors. Nutrients, 10(6), 706.
20. Shannon, C., story, M., Fulkerson, J.A. and French, S.A. (2002). Factors in the school cafeteria influencing food choices by high school students. *Journal of School Health*, 72(6), 229-34.
21. Spronk, I., Kullen, C., Burdon, C. and Conner, H.O. (2014). Relationship between nutrition knowledge and dietary intake. *The British Journal of Nutrition*, 111(10), 1-14.
22. Vabo, M. and Hansen, H. (2016). The Relationship between Food Preferences and Food Choice: A Theoretical Discussion. *International Journal of Business and Social Science*, 5(7), 145-157.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 283-290

LONG TERM EFFECTS OF CHILD ABUSE ON FEMALE ADULT SURVIVOR

Rumana Khalid and **Zaheer Abbas**[§]
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: zaheerabbas@uog.edu.pk

ABSTRACT

Adult female who were abused in childhood feel anxiety and peevishness due to self-blame, degradation, lack of protection, or feeling unsafe. This research focuses on factors like Trust Issues, Personality, Emotion Process, Physical Health, and societal Issues. The main objective of this research is to assess the long term effects of child abuse on adult survivor at University of Gujrat.

Data is collected through questionnaire from 200 female students. Population of the students is selected through snowball sampling. Statistical techniques like reliability analysis, descriptive statistics, chi-square test of association, confirmatory factor analysis, and structure equation modeling (SEM) are used for the analysis of data.

Results of this research show that the mostly the relatives are perpetrators of females survivors. Results also show that there is significant association between types of abused with, adult survivor The results of CFA shows all the variables are confirmed and their behavior as best in the Questionnaire. The results of SEM show that adult survivors of child abuse relate to each factor and all the relations are significant. But the relation between Societal Issues and Physical Health is very strong.

1. INTRODUCTION

Child abuse is known to be one of the major problems which drive an attention of the nation. Child abuse captured the world attention in 1962, when an article published in the Journal of the American Medical Association described symptoms and deemed of child abuse to be medically diagnosable. In 2002 the issue of child abuse as a public-health problem presented by the World Health Organization (WHO). Childhood abuse is clearly detrimental in the developing children as the experiences of being abuse are long-lasting and definitely have a strong influence in life (Razak et al. 2017).

Ericsson, et al. (2009) examined that adult survivors of childhood abuse have significantly more health problems and more painful symptoms in their life. Experiencing child abused is associated with long-term negative effects that can last into adulthood, and include relational, unexplained fear, emotion processing (Briere et al. 2017). Child abuse has direct or indirect impact on psychological symptoms of adult survivor (Young & Widom 2014).

Child abuse is a societal and public health issue, which can lead to long term effects into adulthood (Norman et al., 2012). Child abuse meet many areas of adult survivor's life which include risk of suicide, personality development, mental health problems, physical health problems, trust problems, and societal issues (McSherry, 2011). Children who experience abuse are exposed to various risk factors that resilience of the child is a major determinant of the extent of damage of the abuse in the long term (Mcdonagh et al. 2005). Female with history of sexual child abuse significantly had a number of sleep difficulties, increased risk of depression (Tackett et al. 2013).

Child abuse is very harmful because the abuse occurs when the children are developing their identity (Jonzon and Linblad 2004). Early childhood trauma can lead to difficulty relating to understanding others and inability to regulate reactions to stressful events. Because survivors face trouble to identifying their own needs they may be unable to set boundaries to protect themselves **(**Johnson and James, 2016).

In this study we will discuss the factors that are associated to determine the long term effects of child abuse on adult survivors. These factors are Trust Issues, Personality, Emotion Process, Physical Health and Societal Issues.

2. METHODOLOGY

Sample of 200 female students is selected using snowball sampling method. I identify those female students who are abused in childhood and willing to participate in study and they will almost certainly know other females who are abused in childhood and can help me locate them. Those female students will know other students, and so on. This study is survey type therefore questionnaire is used for data collection. Questionnaire contained 2 sections. The first section contains information about demographic characteristics like age, family cast, abused time, region, marital status, childhood age and abuse perpetrator. Second section of questionnaire contained questions associated to various factors measured on Likert scale.

3. DATA ANALYSIS TECHNIQUES

Different statistical software was used for the analysis of data which are Statistica, Amos, Statistical packages for social science (SPSS). Statistical techniques used for the analysis of data are Descriptive Statistics, Reliability Analysis, Chi-Square test of Association, Confirmatory Factor Analysis (CFA) and Structural equation modeling (SEM).

4. RESULTS

Descriptive Statistics are discussed summaries about sample data.

Table (A)
Descriptive Statistics of Age, and Childhood Abuse Age of Female's Students

Variables	Minimum	Maximum	Mean	Std. Deviation
Age	19	25	21.0800	1.41194
Abuse Age of Respondents	10	17	14.1150	2.01795

Results of this study show that there are 44 percent respondents who are abused sexually and 56 percent respondents who are abused emotionally. The percentage of respondents who belongs to urban area is 47.0 and the percentage of respondents who belongs to rural area is 53.0. We also find that relatives are more perpetrators of females abused as compare to others.

Table shows the association between family system and type of abuse. First of all this table shows the p-value which is 0.03 for the variable type of abuse and family system and its p-value is less than 0.05 so we reach the conclusion to reject null hypothesis and conclude that there is association between type of abuse experience and family system. With 95% confidence the odds of having respondent to be sexually or emotionally abuse is between 0.9 to 3.6 as large when respondent belong to joint family system.

The table (B) shows the Goodness of Fit measures. Chi-square, GFI, AGFI, CFI and RMSEA are the indices for checking model fit. The value of GFI, AGFI and CFI close to 0.9 shows that model is good fit. When the value of RMSEA is less than 0.08 than model is good fitted. In trust issues factor chi-square/d.f =5.6 which is greater than 3 which means that model is not good it happens due to the sample size even the smallest deviation of the data from the model being tested will yield significant chi-square. Thus, it is very common to always get a significant chi-square. The values of GFI and AGFI are 0.86 and 0.76 respectively which are close to 0.9 indicate that model is good fitted. The value of RMSEA is 0.16 which is greater than 0.08. Overall these indices show that model is fit sufficiently. In Personality factor the values of GFI and AGFI are 0.89 and 0.77 respectively which indicate that model is good fitted. The value of RMSEA is 0.18 which is greater than 0.08 which indicate that model is not good fitted. Overall all these indices show that model is fit adequately. In Emotion Process factor the values of GFI and AGFI are 0.88 and 0.79 respectively which indicate that model is good fitted. The value of RMSEA is 0.20 which is greater than 0.08 which indicate that model is not good fitted. Overall all these indices show that model is fit fairly.

Table (B)
Measure of Goodness of Fit of Model for Confirmatory Factor Analysis

Factor/Variable	$\chi 2$	d.f	p-value	$\chi 2$/ d.f	GFI	AGFI	RMSEA
Trust Issues	113.129	20.00	0.000	5.6	0.86	0.76	0.16
Personality	66.374	9.00	0.000	7.3	0.89	0.77	0.18
Emotion Process	112.585	9.00	0.000	12.5	0.88	0.79	0.20
Physical Health	35.24	9.00	0.000	3.9	0.94	0.91	0.91
Societal Issues	58.87	9.00	0.000	6.5	0.90	0.81	0.17
Adult Survivor	79.58	14.00	0.000	5.7	0.88	0.77	0.16
Recommended				≤ 3	$\geq .90$	$\geq .90$	≤ 0.08

The table (C) shows the goodness of fit measures. Indices Chi-square GFI, AGFI, NFI, CFI and RMSEA are used for the checking of model fit. The p-value of chi-square is 0.000 which is less than 0.05 it means that value is significant it happens because of large sample size. Mostly we get chi-square value is significant so it is not a big problem. The value of test statistic / d.f is 2.08 which is less than 3 it means that model is good fitted. The value of RMSEA is less than 0.08 than model is good fitted. Here the value of GFI is .875 which is close to 0.9 it means that model is good fitted. The value of AGFI is .823 it means that model is good. The value of NFI and CFI is .848 and .914 respectively it means that model is good. The value of RMSEA is .071 which is less than 0.08 which mean model is good fitted. Overall we say that model is fitted sufficiently.

Table (C)
Measure of Goodness of Fit Model

	χ2	**d.f**	**p-value**	**χ2 /d.f**	**GFI**	**AGFI**	**RMSEA**	**NFI**	**CFI**
SEM	265.79	128	.000	2.08	.875	.823	.071	.848	.828
Recommended				≤ 3	≥ .90	≥ .90	≤ 0.08	≥ .90	≥ .90

The direct effects of trust issues and emotion process on adult survivors are given below:

Adult Survivor = 0.248 Trust Issues (TI) (1)
Adult Survivor = 0.246 Emotion Process (EP) (2)

Comparing Equation (1) and (2)

Adult Survivor = 0.248(TI) + 0.246(EP)

Result of SEM shows model estimates that are interpreted as regression coefficients. Estimates output shows that all the relations are significant for the adult survivor of child abuse. The relation between Societal Issues and Physical Health is strong than all other relations as its value is .754. Also emotion process affects the adult survivor. The relation between Societal Issues and Trust Issue is low than other relations.

Table (D)
SEM Results for Adult Survivor

Variables	**Parameter Estimate**	**Prob. Level**
Trust Issues ← Personality	.425	***
Trust Issues ← Societal Issues	.211	.003
Physical Health ← Societal Issues	.754	***
Emotion Process ← Physical Health	.599	***
Emotion Process ← Trust Issues	.369	***
Adult Survivor ← Trust Issues	.248	.002
Adult Survivor ← Emotion Process	.246	.002

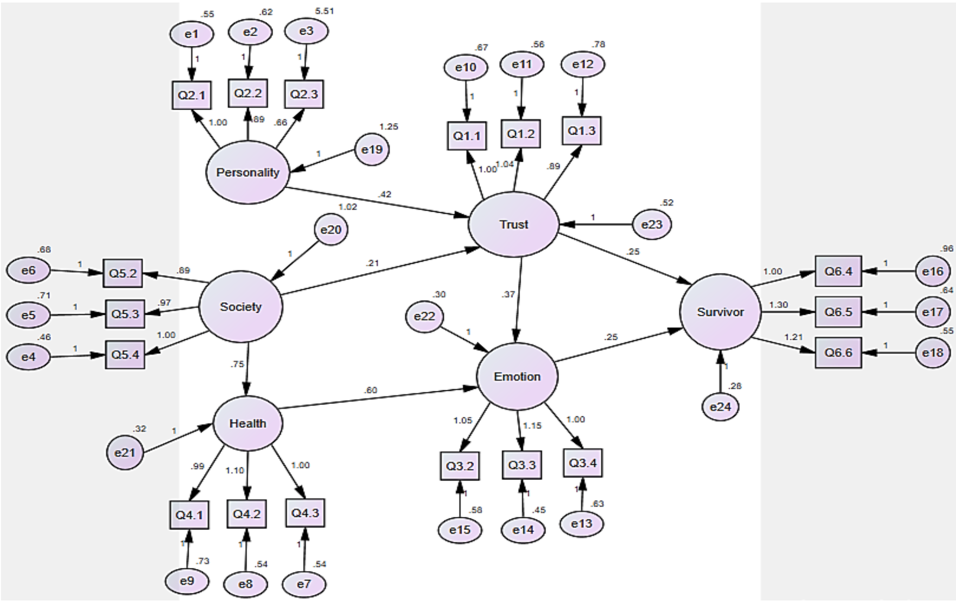


Figure (1): Path Analysis

5. DISCUSSION

A person who was abused as a child is an adult survivor of child abuse. Adult who were abused in childhood may display symptoms of Post-Traumatic Stress Disorder which includes flashbacks, bad dreams, physical reactions to reminders of the event, avoidance, memory loss, social and emotional detachment, hopelessness, numbing, insomnia, anger, difficulty concentrating, self-blame, suicidal thoughts, and feelings of mistrust (Burns, et al. 2010).

This study analyzes the long term effects of child abuse on female adult survivor by using different independent factors Trust Issues, Emotion Process, Personality, Physical Health and Societal Issues. Descriptive statistics shows that that minimum age of respondents of this study is 19 and the maximum age is 25. The minimum and maximum abused childhood age of respondents in the study is 10 and 17 respectively.

Our analysis showed the females who are abused in childhood face more societal issues and have high health problems. In data collection a female share her abused experience she tells that she was sexually abused in 12 year age from her teacher. She reported that after this incident she has sleeping problem and she feel afraid from society. This analysis is concordant with other studies for example, Springer, et al. (2003) observed that experience of childhood abuse in women caused a long term health problems including back pain, severe headaches, eating disorder and sleeping problem. Ahrens et al. (2010) investigate that an adult survivor of child abuse mostly feels alone and shows from their expressions that this society suffocates them. The information acquired through our study observed that there is significant relation between female's adult survivor with trust issues and emotion process. Its means that after abuse females cannot trust on their friends r relatives they feel difficulty to make relationships and they mostly feel stress or anxiety. Ratican (1992) find that survivors of child abuse face difficulty in establishing interpersonal relationships with friends. Long et al. (2006) observed that Stress and anxiety are often long-term effects of child abuse. These anxiety-related symptoms are frequently associated with PTSD responses. Our analysis suggests that personality factor influence on female adult survivors they feel inferior to other peoples. Corsentino et al. (2014) find that Child abuse is greatly influenced on personality of adult survivors the personality development of abused adult is totally different from which are not abused. Mansor & Samah (2011) also find that female's adults who experienced abuse during their childhood tend to engage themselves in high-risk behaviors.

6. CONCLUSION

The study has analyzed the factors which have long term effect of child abuse on adult survivor. Several factors which are used to conduct the study are trust issue, personality, emotion process, physical health, and societal issue. It is concluded that females who have abused experience in childhood face its long term negative effects. Descriptive statistics is used to check the current age and childhood age of female's survivors. The study show that in rural area with nuclear family system of respondents experienced most abused. In the study emotional abused are most as compare to sexual abused. It is concluded that in Jutt cast abused happened more than other casts. The study shows that relatives are more involved in doing abused with females as compare to strangers or friends. Furthermore the study conclude the significant association between type of abused with family type, adult survivor and all other factors which are trust issue, personality, emotion process, physical health and societal issues. Degree of risk tells that female who are abused in childhood have low trust on others. The results of CFA shows all the variables are confirmed and their behavior as best in the Questionnaire. The results of SEM show that adult survivor of child abuse relate to each other. And all the relations are significant. Goodness of fit is tested across various indices, all the indices shows that the model is good fitted.

7. RECOMMENDATION

This research further can support to improve the life style of females survivors who are abused in childhood. Working with survivors of child abuse may be achieved by to disclose any instances of trauma, including child abuse in childhood. Parents should protect their child and build trust that in any case of abuse female tells her parents and disclose her abuse incident immediately. If a female disclose her abuse experience on that time with her any close relation or parents he fear that people not believe her. So parents not blame their child for abused experience they should try to established therapeutic relationship with their child that help them in forgetting abused incident to spend a normal life

According to this study the relatives are more involved in doing abused with females. So the parents should careful and not trust their relatives in their child matter. It is also recommended in a transparent and collaborative manner with the survivor, in order to reduce feelings of betrayal or abandonment.

REFERENCES

1. Ahrens, C.E., Stansell, J. and Jennings, A. (2010). To tell or not to tell: The impact of disclosure on sexual assault survivors' recovery. *Journal of Violence and Victims*, 25(5), 631-648.
2. Briere, J., Runtz, M., Eadia, E., Bigras, N. and Godbout, N. (2017). Disengaged Parenting: Structural Equation Modeling With Child Abuse, Insecure Attachment and Adult Symptomatology. *Journal of Child Abuse & Neglect*, 6(7), 260-270.
3. Corsentinoa, E., Tackett, K.A.K., Shefflera, J., Arcea, D., Rushinga, C.N. and Ericssona, S.N. (2014). The Influence of Prior Rape on the Psychological and Physical Health Functioning of Older Adults. *Journal of Aging & Mental Health*, 6(3), 1-14.
4. Ericsson, S.N., Cromer, K. and Hernandez, A. (2009). A Review of Childhood Abuse, Health, and Pain-Related Problems: The Role of Psychiatric Disorders and Current Life Stress. *Journal of Trauma & Dissociation*, 10(3), 170-188.
5. Johnson, E.J. and James, C. (2016). Effects of Child Abuse and Neglect on Adult Survivors. *Journal of Early Child Development and Care*, 186(11), 1836-1845.
6. Jonzon, E., and Linblad, F. (2004). Disclosure, reactions, and social support: Findings from a sample of adult victims of child sexual abuse. *Journal of Child Maltreatment*, 9(2), 190-200.
7. Long, L.L., Burnett, J.A. and Thomas, R.V. (2006). *Sexuality counseling: An integrative approach*. Upper Saddle River, NJ: Pearson.
8. Mansor, M. and Samah, A.A. (2011). A descriptive analysis on the personality of child physical abuse victims. *International Journal of Business and Social Science*, 2(14), 114-131.
9. Mcsherry, D. (2011). Understanding and addressing the "neglect of neglect": Why are we making a mole-hill out of a mountain? *Journal of Child Abuse & Neglect*, 3(1), 607-614.
10. Mcdonagh, A., Friedman, M., McHugo, G., Ford, J., Sengupta, A., Mueser, K., Demment, C.C., Fournier, D., Schnurr, P.P. and Descamps, M. (2005). Randomized Trial of Cognitive–Behavioral Therapy for Chronic Posttraumatic Stress Disorder in Adult Female Survivors of Childhood Sexual Abuse. *Journal of Consulting and Clinical Psychology*, 73(3), 515-524.
11. Norman, R.E., Byambaa, M., De, R., Butchart, A., Scott, J. and Vos, T. (2012). The long-term health consequences of child physical abuse, emotional abuse, and neglect: a systematic review and meta-analysis. *PLoS Medicine*, 9(11). doi: 10.1371/journal.pmed.1001349
12. Razak, M., Ibrahim, H., Mokhtar, S. and Nuraya, R.R. (2017). The Personality Development of Child Abuse Survivor. *Journal of Humanities, Language, Culture and Business*, 1(2), 1-9.
13. Ratican, K. (1992). Sexual abuse survivors: Identifying symptoms and special treatment considerations. *Journal of Counseling & Development*, 71(1), 33-38.

14. Springer, W.K., Sheridan, J. Kuo, K. and Carnes, M. (2003). The Long-term Health Outcomes of Childhood Abuse. *Journal of General Internal Medicine*, 18(10), 864-870.
15. Tackett, K.K., Cong, Z. and Hale, W.T. (2013). Depression, Sleep Quality, and Maternal Well-Being in Postpartum Women with a History of Sexual Assault: A Comparison of Breastfeeding, Mixed-Feeding, and Formula-Feeding Mothers. Journal *of Breast feeding Medicine*, 8(1), 16-22.
16. Young, J.C. and Widom, C.S. (2014).Long Term Effects of Child Abuse and Neglect on Emotion Processing in Adulthood. *Psychology Department*, *John Jay College of Criminal Justice*, 22(4), 1-13.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 291-307

A NOTE ON THE TRANSMUTED POWER LINDLEY DISTRIBUTION

Shakila Bashir and **Noor Waseem**
[1] Forman Christian College, Lahore, Pakistan
Email: shakilabashir@fccollege.edu.pk
[2] Kinnaird College for Women, Lahore, Pakistan
Email: noor.waseem@hotmail.com

ABSTRACT

In this article, we generalized the power Lindley distribution using the quadratic rank transmutation (QRT) to develop the transmuted power Lindley distribution (TPLD). We provide the comprehensive description of the mathematical properties of the proposed distribution along with its reliability behavior. The usefulness of the transmuted power Lindley distribution is illustrated using real data sets.

1. INTRODUCTION

In 1958 Lindley developed a one Parameter, lifetime Probability distribution named as Lindley distribution. Lifetime distributions are applied to explain, statistically the length of the life of a device/system. In many applied sciences such as medicine, engineering and finance, amongst others, modeling and analyzing lifetime data are crucial. Lindley distribution is applied to examine the failure time of data Cakmakyapan and Ozel (2016). The probability density function (pdf) of Lindley Distribution (LD) with scale parameter θ, is

$$f(x;\theta)=\frac{\theta^2}{\theta+1}(1+x)e^{-\theta x};x>0,\theta>0 \tag{1.1}$$

Nadarajah (2011) introduced the two-parameter generalized Lindley distribution. Shanker and Mishra (2013) worked on two parameter quasi Lindley distribution and discussed the monotonically increasing behavior hazard function. Some statistical properties and reliability analysis for the extended New Generalized Lindley Distribution have been studied by Shibu and Irshad (2016).

Ghitany et al. (2013) developed the Power Lindley (PL) distribution by using the transformation $X=Y^{1/\beta}$, the probability density function pdf of the PL distribution is given by

$$g(x)=\frac{\beta\theta^2}{\theta+1}x^{\beta-1}\left(1+x^{\beta}\right)e^{-\theta x^{\beta}};x>0\quad \theta,\beta>0 \tag{1.2}$$

and the cumulative distribution function (cdf) of PL distribution is given as

$$G(x)=1-e^{-\theta x^{\beta}}\left[1+\frac{\theta x^{\beta}}{\theta+1}\right] \tag{1.3}$$

Jan et al. (2018) introduced the generalization form of three parameter Lindley distribution named as Exponentiated Inverse Power Lindley Distribution (EIPLD) and discussed various mathematical properties including moments, hazard rate. For non-monotone survival data Alkarni (2015) proposed a new model which is an extension of power Lindley distribution presented as "Extended power Lindley distribution".

Shaw and Bukley (2009) suggested Quadratic Rank Transmutation Map (QRTM) to develop a new family of non-Gaussian distribution. Let $g(x)$ and $G(x)$ are the pdf and cdf of base distribution. The QRTM distribution is

$$F(x) = (1+\lambda)G(x) - \lambda G^2(x), |\lambda| \leq 1 \tag{1.4}$$

$$f(x) = g(x)\left[(1+\lambda) - 2\lambda G(x)\right], |\lambda| \leq 1 \tag{1.5}$$

where $F(x)$ and $f(x)$ are the cdf and pdf of corresponding Quadratic Rank Transmutation Map.

Arshad et al. (2018) constructed a life model names as Transmuted Exponentiated Moment Pareto Distribution and discussed various properties along with its reliability behavior and 3 special cases are developed. Alpha power transformed power Lindley (APTPL) is proposed by Hassan (2019). Kemaloglu and Yilmaz (2017) Introduced a two-parameter distribution by using QRTM presented as "Transmuted two-parameter Lindley distribution" and applied it to guinea pigs and bladder cancer data. A new generalization of Lindley distribution by using Quadratic Rank Transmuted Map as Transmuted Lindley distribution by Merovci (2013). Ekhosuehi (2018) Consider a new class of lifetime distribution named as "Three Parameter Generalized Lindley distribution" and has used (TPGL) distribution to model Strength of glass fibers data.

2. PROPOSED DISTRIBUTION

A three-parameter distribution named as Transmuted Power Lindley Distribution (TPLD) was developed by using the Quadratic Rank Transmutation Map (QRTM). By using equation (1.3) in equation (1.4) we get the cdf of TPLD

$$F(x) = (1+\lambda)\left[1 - e^{-\theta x^{\beta}}\left(1 + \frac{\theta x^{\beta}}{\theta+1}\right)\right] - \lambda\left[1 - e^{-\theta x^{\beta}}\left(1 + \frac{\theta x^{\beta}}{\theta+1}\right)\right]^2 \tag{2.1}$$

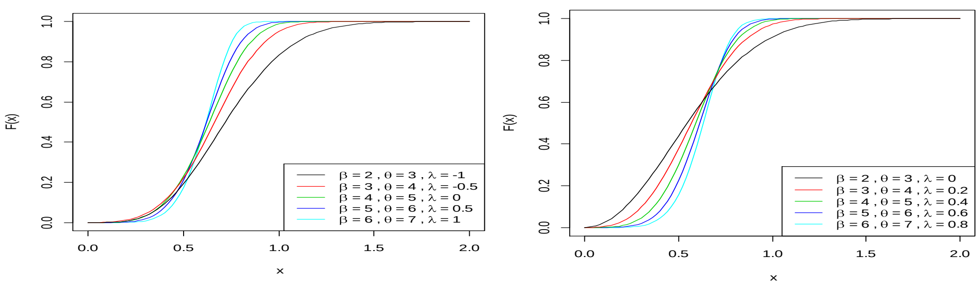


Figure 1: Cumulative Distribution Plot of TPLD

By using equation (1.2) and (1.3) in equation (1.5) we get the pdf of TPLD

$$f(x)=\frac{\beta\theta^2}{(\theta+1)}x^{\beta-1}\left(1+x^\beta\right)e^{-\theta x^\beta}\left[1-\lambda+2\lambda e^{-\theta x^\beta}\left(1+\frac{\theta x^\beta}{\theta+1}\right)\right] \tag{2.2}$$

where $x,\theta,\beta>0$ and $|\lambda|\le 1$. For $\lambda=0$, TPLD reduces to PLD and $\lambda=0$ & $\beta=1$, TPLD reduces to LD.

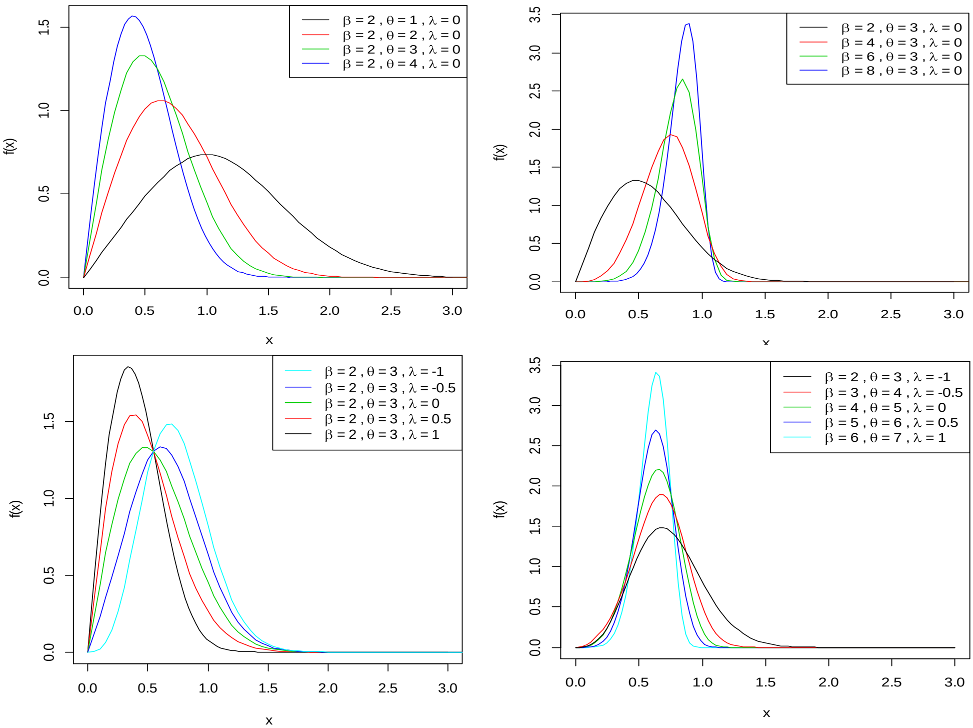


Figure 2: Probability Density Plots of TPLD

3. PROPERTIES OF TRANSMUTED POWER LINDLEY DISTRIBUTION

In this section the reliability measures for the TPLD are derived and graphical shapes are shown.

3.1 Survival Function

The survival function of TPLD is given by

$$S(x)=1-(1+\lambda)\left[1-e^{-\theta x^\beta}\left(1+\frac{\theta x^\beta}{\theta+1}\right)\right]+\lambda\left[1-e^{-\theta x^\beta}\left(1+\frac{\theta x^\beta}{\theta+1}\right)\right]^2 \tag{3.1}$$

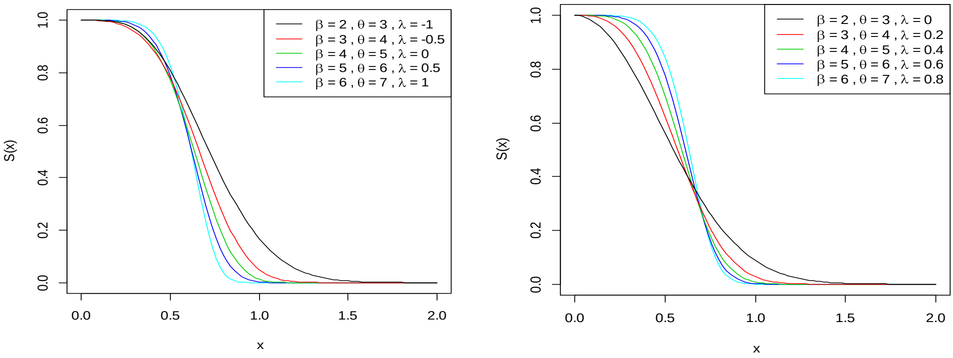


Figure 3: Survival Plots of TPLD

3.2 Hazard Function

The hazard function of TPLD is as follows

$$h(x)=\frac{\beta\theta^2 x^{\beta-1}\left(1+x^{\beta}\right)e^{-\theta x^{\beta}}\left[1-\lambda+2\lambda e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]}{(\theta+1)\left[1-(1+\lambda)\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}+\lambda\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}^2\right]} \quad (3.2)$$

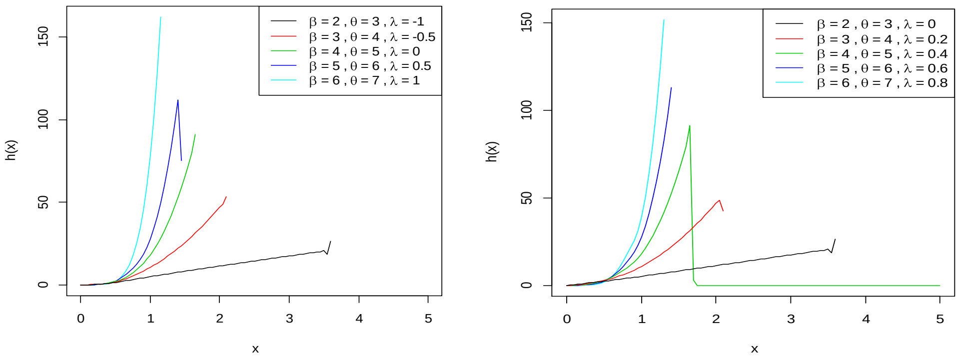


Figure 4: Hazard Plots of TPLD

From Figure 4, it can be seen that the shape of hazard function for TPLD is increasing and then decreasing.

3.3 Cumulative Hazard Function

The Cumulative hazard function of TPLD is given by

$$H(x) = -\ln\left[1-(1+\lambda)\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\} + \lambda\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}^{2}\right] \tag{3.3}$$

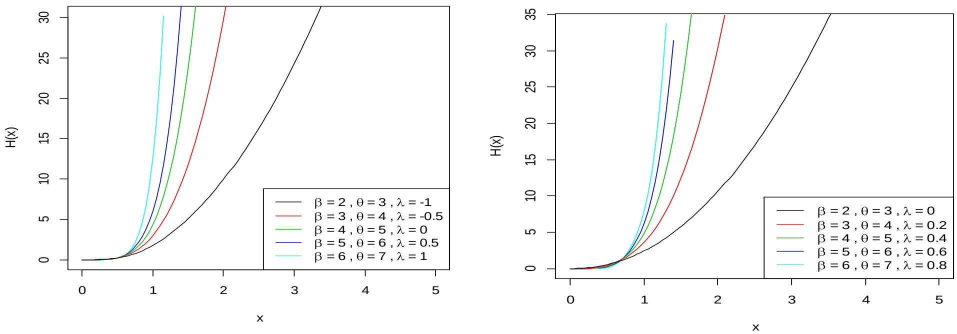


Figure 5: Cumulative Hazard Plots of TPLD

3.4 Reversed Hazard Function of TPLD

The reversed hazard function of TPLD is given by

$$r(x) = \frac{\theta^{2} x^{\beta-1}\left(1+x^{\beta}\right)e^{-\theta x^{\beta}}\left[1-\lambda+2\lambda e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]}{(\theta+1)\left[(1+\lambda)\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}-\lambda\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}^{2}\right]} \tag{3.4}$$

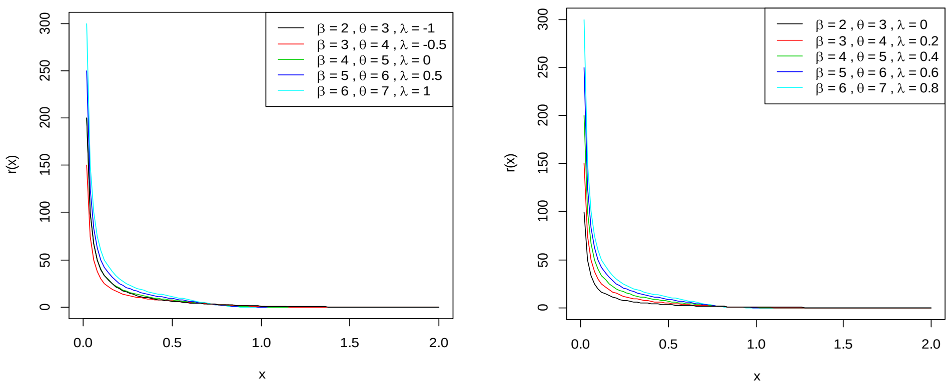


Figure 6: Reversed Hazard Plots of TPLD

3.5 Mills Ratio of TPLD

By definition mills ratio is defined as

$$M(x) = \frac{S(x)}{f(x)}$$

From equation (3.1) and (2.1), mills ratio of TPLD is

$$M(x) = \frac{(\theta+1)\left[1-(1+\lambda)\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}+\lambda\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}^{2}\right]}{\beta\theta^{2}x^{\beta-1}\left(1+x^{\beta}\right)e^{-\theta x^{\beta}}\left[1-\lambda+2\lambda e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]} \quad (3.5)$$

3.6 Odd Function of TPLD

By definition odd function is defined as

$$O(x) = \frac{F(x)}{S(x)}$$

From equation (2.1) and (3.1), odd function of TPLD is

$$O(x) = \frac{(1+\lambda)\left[1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]-\lambda\left[1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]^{2}}{1-(1+\lambda)\left[1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]+\lambda\left[1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]^{2}} \quad (3.6)$$

3.7 Elasticity of TPLD

By definition elasticity is defined as

$$e(x) = \frac{xf(x)}{F(x)}$$

From equation (2.1) and (2.2), elasticity of TPLD is

$$e(x) = \frac{\beta\theta^{2}x^{\beta}\left(1+x^{\beta}\right)e^{-\theta x^{\beta}}\left[1-\lambda+2\lambda e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right]}{(\theta+1)\left[(1+\lambda)\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}-\lambda\left\{1-e^{-\theta x^{\beta}}\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right\}^{2}\right]} \quad (3.7)$$

4. MOMENTS

Several characteristics of the distribution can be calculated by moments. With the help of r^{th} moment we can derive the higher moments, central moments, mean, variance and many other characteristics of the distribution.

4.1 Moments about Origin of TPLD

The rth moment about origin of TPLD is as follows

$$\mu_r' = \int_0^{\infty} x^r f(x)dx$$

$$\mu_r' = \frac{1}{\theta^{\frac{r}{\beta}}(\theta+1)}\left[\begin{array}{r}(1-\lambda)\left\{\theta\Gamma\left(\frac{r}{\beta}+1\right)+\Gamma\left(\frac{r}{\beta}+2\right)\right\}+\frac{\lambda}{2^{\frac{r}{\beta}+1}}\left\{\theta\Gamma\left(\frac{r}{\beta}+1\right)+\Gamma\left(\frac{r}{\beta}+2\right)\right\}\\ +\frac{\lambda}{(\theta+1)2^{\frac{r}{\beta}+1}}\left\{\theta\Gamma\left(\frac{r}{\beta}+2\right)+\frac{\Gamma\left(\frac{r}{\beta}+3\right)}{2}\right\}\end{array}\right] \tag{4.1.1}$$

For r =1, 2, 3 & 4 in equation, we get first four raw moment

$$\mu_1' = \frac{\Gamma\left(\frac{1}{\beta}\right)}{\theta^{\frac{1}{\beta}}\beta^2(\theta+1)}\left[\left(1-\lambda+\frac{\lambda}{2^{\frac{1}{\beta}+1}}\right)\left(\beta(\theta+1)+1\right)+\frac{\lambda}{2^{\frac{1+2\beta}{\beta}}\beta(\theta+1)}\left(\beta+1\right)\left(2\beta(\theta+1)+1\right)\right] \tag{4.1.2}$$

$$\mu_2' = \frac{2\Gamma\left(\frac{2}{\beta}\right)}{\theta^{\frac{2}{\beta}}\beta^2(\theta+1)}\left[\left(1-\lambda+\frac{\lambda}{2^{\frac{2}{\beta}+1}}\right)\left(\beta(\theta+1)+2\right)+\frac{\lambda}{2^{\frac{2(1+\beta)}{\beta}}\beta(\theta+1)}\left(\beta+2\right)\left(2\beta(\theta+1)+2\right)\right] \tag{4.1.3}$$

$$\mu_3' = \frac{3\Gamma\left(\frac{3}{\beta}\right)}{\theta^{\frac{3}{\beta}}\beta^2(\theta+1)}\left[\left(1-\lambda+\frac{\lambda}{2^{\frac{3}{\beta}+1}}\right)\left(\beta(\theta+1)+3\right)+\frac{\lambda}{2^{\frac{3+2\beta}{\beta}}\beta(\theta+1)}\left(\beta+3\right)\left(2\beta(\theta+1)+3\right)\right] \tag{4.1.4}$$

$$\mu_4' = \frac{4\Gamma\left(\frac{4}{\beta}\right)}{\theta^{\frac{4}{\beta}}\beta^2(\theta+1)}\left[\left(1-\lambda+\frac{\lambda}{2^{\frac{4}{\beta}+1}}\right)\left(\beta(\theta+1)+4\right)+\frac{\lambda}{2^{\frac{4+2\beta}{\beta}}\beta(\theta+1)}\left(\beta+4\right)\left(2\beta(\theta+1)+4\right)\right] \tag{4.1.5}$$

4.2 Fractional Positive Moments of TPLD

$$\mu'_{\frac{m}{n}} = \int_0^\infty x^{\frac{m}{n}} f(x)dx$$

$$\mu'_{\frac{m}{n}} = \frac{\beta\theta^2}{(\theta+1)} \int_0^\infty x^{\frac{m}{n}+\beta-1}\left(1+x^\beta\right)e^{-\theta x^\beta}\left[1-\lambda+2\lambda e^{-\theta x^\beta}\left(1+\frac{\theta x^\beta}{\theta+1}\right)\right]dx$$

$$\mu'_{\frac{m}{n}} = \frac{1}{\theta^{\frac{m}{n}\beta}(\theta+1)}\left[\begin{array}{r}(1-\lambda)\left\{\theta\Gamma\left(\frac{m}{n}\beta+1\right)+\Gamma\left(\frac{m}{n}\beta+2\right)\right\}+\frac{\lambda}{2^{\frac{m}{n}\beta+1}}\left\{2\theta\Gamma\left(\frac{m}{n}\beta+1\right)+\Gamma\left(\frac{m}{n}\beta+2\right)\right\}\\ +\frac{1}{(\theta+1)2^{\frac{m}{n}\beta+2}}\left\{2\Gamma\left(\frac{m}{n}\beta+2\right)+\Gamma\left(\frac{m}{n}\beta+3\right)\right\}\end{array}\right] \tag{4.2}$$

4.3. Fractional Negative Moments of TPLD

$$\mu'_{-\frac{m}{n}} = \int_0^\infty x^{-\frac{m}{n}} f(x)dx$$

$$\mu'_{-\frac{m}{n}} = \frac{\theta^{\frac{m}{n}\beta}}{(\theta+1)}\left[\begin{array}{r}(1-\lambda)\left\{-\theta\Gamma\left(1-\frac{m}{n}\beta\right)+\Gamma\left(2-\frac{m}{n}\beta\right)\right\}+\frac{\lambda}{2^{1-\frac{m}{n}\beta}}\left\{2\theta\Gamma\left(1-\frac{m}{n}\beta\right)+\Gamma\left(2-\frac{m}{n}\beta\right)\right\}\\ +\frac{1}{(\theta+1)2^{2-\frac{m}{n}\beta}}\left\{2\Gamma\left(2-\frac{m}{n}\beta\right)+\Gamma\left(3-\frac{m}{n}\beta\right)\right\}\end{array}\right] \tag{4.3}$$

4.4 Factorial Moments of TPLD

$$E[X]_n = \sum_{r=0}^{n} \varphi_r \mu_r'$$

$$E[X]_n = \sum_{r=0}^{n} \frac{1}{\theta^{\frac{r}{\beta}}(\theta+1)} \left[\begin{array}{l} (1-\lambda+\frac{\lambda}{2^{\frac{r}{\beta}+1}})\left\{\theta\Gamma\left(\frac{r}{\beta}+1\right)+\Gamma\left(\frac{r}{\beta}+2\right)\right\} \\ \qquad +\frac{\lambda}{(\theta+1)2^{\frac{r}{\beta}+2}}\left\{2\theta\Gamma\left(\frac{r}{\beta}+2\right)+\Gamma\left(\frac{r}{\beta}+3\right)\right\} \end{array} \right] \tag{4.4}$$

4.5 Moment Generation Function of TPLD

Moment generating function (MGF) of the TPL distribution is

$$M_x(t) = E(e^{tx}) = \int_0^{\infty} e^{tx} f(x)dx$$

$$M_x(t) = \sum_{r=0}^{\infty} \frac{t^r}{r!\,\theta^{\frac{r}{\beta}}(\theta+1)} \left[\begin{array}{l} (1-\lambda+\frac{\lambda}{2^{\frac{r}{\beta}+1}})\left\{\theta\Gamma\left(\frac{r}{\beta}+1\right)+\Gamma\left(\frac{r}{\beta}+2\right)\right\} \\ \qquad +\frac{\lambda}{(\theta+1)2^{\frac{r}{\beta}+2}}\left\{2\theta\Gamma\left(\frac{r}{\beta}+2\right)+\Gamma\left(\frac{r}{\beta}+3\right)\right\} \end{array} \right] \tag{4.5}$$

4.6 Central Moments of TPLD

Central Moments of the TPL distribution is

$$E(X-\mu)^r = \mu_r = \sum_{i=0}^{r} \frac{\binom{r}{i}(-1)^i(\mu_1')^i}{\theta^{\frac{r-i}{\beta}}(\theta+1)} \left[\begin{array}{l} (1-\lambda+\frac{\lambda}{2^{\frac{r-i}{\beta}+1}})\left\{\theta\Gamma\left(\frac{r-i}{\beta}+1\right)+\Gamma\left(\frac{r}{\beta}+2\right)\right\} \\ \qquad +\frac{\lambda}{(\theta+1)2^{\frac{r-i}{\beta}+2}}\left\{2\theta\Gamma\left(\frac{r-i}{\beta}+2\right)+\Gamma\left(\frac{r-i}{\beta}+3\right)\right\} \end{array} \right] \tag{4.6}$$

For r=2, 3 and 4 we get $\mu_2, \mu_3 \,\&\, \mu_4$.

4.7 The Mellin Transformation of TPLD

The Mellin transformation of TPLD is

$$M_x(m) = E(x^{m-1}) = \int_0^{\infty} x^{m-1} f(x)dx$$

$$M_x(m)=\frac{1}{\theta^{\frac{m-1}{\beta}}(\theta+1)}\left[\begin{array}{l}(1-\lambda+\frac{\lambda}{2^{\frac{m-1}{\beta}+1}})\left\{\theta\Gamma\left(\frac{m-1}{\beta}+1\right)+\Gamma\left(\frac{m-1}{\beta}+2\right)\right\}\\ +\frac{\lambda}{(\theta+1)2^{\frac{m-1}{\beta}+2}}\left\{2\theta\Gamma\left(\frac{m-1}{\beta}+2\right)+\Gamma\left(\frac{m-1}{\beta}+3\right)\right\}\end{array}\right] \tag{4.7}$$

4.8 Residual Life Function of TPLD

The residual life function of TPLD is

$$m_n(w)=E\left[\left(X-w\right)^n/X>w\right]=\frac{1}{S(w)}\int_w^\infty (x-w)^s f(x)dx$$

$$\therefore X\sim TPLD$$

$$m_n(w)=\frac{1}{S(w)}\sum_{s=0}^{n}\binom{n}{s}\left(-w\right)^{n-s}\int_w^\infty x^s f(x)dx$$

$$m_n(w)=\frac{1}{S(w)}\sum_{s=0}^{n}\binom{n}{s}\left(-w\right)^{n-s}\int_w^\infty x^s \frac{\beta\theta^2}{\left(\theta+1\right)}x^{\beta-1}\left(1+x^\beta\right)e^{-\theta x^\beta}\left[1-\lambda+2\lambda e^{-\theta x^\beta}\left(1+\frac{\theta x^\beta}{\theta+1}\right)\right]dx$$

$$m_n(w)=\frac{1}{S(w)}\sum_{s=0}^{n}\binom{n}{s}\frac{\left(-w\right)^{n-s}}{\left(\theta+1\right)\theta^{\frac{s}{\beta}}}\left[\begin{array}{l}(1-\lambda)\left\{\theta\Gamma\left(\frac{s}{\beta}+1,\theta w^\beta\right)+\Gamma\left(\frac{s}{\beta}+2,\theta w^\beta\right)\right\}\\ +\lambda\left\{\theta\Gamma\left(\frac{s}{\beta}+1,2\theta w^\beta\right)+\Gamma\left(\frac{s}{\beta}+2,2\theta w^\beta\right)\right\}\\ +\frac{\lambda}{\left(\theta+1\right)}\left\{\theta\Gamma\left(\frac{s}{\beta}+2,2\theta w^\beta\right)+\Gamma\left(\frac{s}{\beta}+3,2\theta w^\beta\right)\right\}\end{array}\right] \tag{4.8}$$

For life expectancy or mean residual life (MRL) function say $m_1(w)$ of TPLD put n=1 in equation (4.8)

$$m_1(w)=\frac{1}{S(w)}\sum_{s=0}^{1}\binom{1}{s}\frac{(-w)^{1-s}}{(\theta+1)\theta^{\frac{s}{\beta}}}\left[\begin{array}{l}(1-\lambda)\left\{\theta\Gamma\left(\frac{s}{\beta}+1,\theta w^{\beta}\right)+\Gamma\left(\frac{s}{\beta}+2,\theta w^{\beta}\right)\right\}\\ +\lambda\left\{\theta\Gamma\left(\frac{s}{\beta}+1,2\theta w^{\beta}\right)+\Gamma\left(\frac{s}{\beta}+2,2\theta w^{\beta}\right)\right\}\\ +\frac{\lambda}{(\theta+1)}\left\{\theta\Gamma\left(\frac{s}{\beta}+2,2\theta w^{\beta}\right)+\Gamma\left(\frac{s}{\beta}+3,2\theta w^{\beta}\right)\right\}\end{array}\right] \tag{4.9}$$

4.9 Reverse Residual Life Function of TPLD

The reverse residual life function of TPLD is

$$R_n(w)=E\left[(w-X)^n / X\le w\right]=\frac{1}{F(w)}\int_0^{\infty}(w-x)^n f(x)dx$$

$$R_n(w)=\frac{1}{F(w)}\sum_{t=0}^{n}\binom{n}{t}(-1)^t w^{n-t}\int_0^{\infty}x^t f(x)dx$$

$$R_n(w)=\frac{1}{F(w)}\sum_{t=0}^{n}\binom{n}{t}(-1)^t\frac{w^{n-t}}{(\theta+1)\theta^{\frac{t}{\beta}}}\left[\begin{array}{l}\left(1-\lambda+\frac{\lambda}{2^{\frac{t}{\beta}+1}}\right)\left\{\theta\Gamma\left(\frac{t}{\beta}+1\right)+\Gamma\left(\frac{t}{\beta}+2\right)\right\}\\ +\frac{\lambda}{2^{\frac{t}{\beta}+2}(\theta+1)}\left\{2\theta\Gamma\left(\frac{t}{\beta}+2\right)+\Gamma\left(\frac{t}{\beta}+3\right)\right\}\end{array}\right] \tag{4.10}$$

For mean waiting time or mean inactivity time of TPLD put n=1 in equation (4.10)

$$R_1(w)=\frac{1}{F(w)}\sum_{t=0}^{1}\binom{1}{t}(-1)^t\frac{w^{1-t}}{(\theta+1)\theta^{\frac{t}{\beta}}}\left[\begin{array}{l}\left(1-\lambda+\frac{\lambda}{2^{\frac{t}{\beta}+1}}\right)\left\{\theta\Gamma\left(\frac{t}{\beta}+1\right)+\Gamma\left(\frac{t}{\beta}+2\right)\right\}\\ +\frac{\lambda}{2^{\frac{t}{\beta}+2}(\theta+1)}\left\{2\theta\Gamma\left(\frac{t}{\beta}+2\right)+\Gamma\left(\frac{t}{\beta}+3\right)\right\}\end{array}\right] \tag{4.11}$$

5. QUANTILE FUNCTION OF TPLD

The q^{th} quantile function of Transmuted Power Lindley Distribution is as follows

$$x_q = \left[\frac{1}{\theta} \left\{ -\ln \left(\frac{1-q}{\left(1-\lambda+\lambda e^{-\theta x^{\beta}} \left(1+\frac{\theta x^{\beta}}{\theta+1}\right)\right)\left(1+\frac{\theta x^{\beta}}{\theta+1}\right)} \right) \right\} \right]^{\frac{1}{\beta}} \qquad (5.1)$$

For q=0.25, 0.5 and 0.75 we can find first, second and third quantiles of the TPL distribution.

6. MAXIMUM LIKELIHOOD ESTIMATION

The Maximum Likelihood Estimation is used for estimating parameters. Let $X_1, X_2, X_3, \ldots, X_n$ be a n random samples from Transmuted Power Lindley distribution. Then the Likelihood function (L) of equation (2.2) is given by

$$L(\theta,\beta,\lambda) = \left(\frac{\beta\theta^2}{(\theta+1)}\right)^n \prod_{i=1}^{n} x_i^{\beta-1}(1+x_i^{\beta}) + e^{-\theta\sum_{i=1}^{n} x_i^{\beta}} \prod_{i=1}^{n}\left[1-\lambda+2\lambda e^{-\theta x_i^{\beta}}\left(1+\frac{\theta x_i^{\beta}}{\theta+1}\right)\right]$$

$$\ln L = n\ln\beta + 2n\ln\theta - n\ln(\theta+1) + (\beta-1)\sum_{i=1}^{n}\ln x_i$$
$$+\sum_{i=1}^{n}\ln(1+x_i^{\beta}) - \theta\sum_{i=1}^{n} x_i^{\beta} + \sum_{i=1}^{n}\ln\left[1-\lambda+2\lambda e^{-\theta x_i^{\beta}}\left(1+\frac{\theta x_i^{\beta}}{\theta+1}\right)\right] \qquad (6.1)$$

7. SIMULATION STUDY OF TPLD

To assess the behavior of estimates derived by the method of MLE from TPLD a small scaled experiment carried our based on simulations study. The performance of the parameters estimated by MLE is evaluated on the basis of mean square errors (MESs). For this purpose, we generate $n = 25, 50, 100, 200, 500$ samples by 1000 simulations from TPLD pdf in eq. (2.2) by using Mathematica 11. Table 1 and 2 are showing the consistent and efficient performance of the estimates produced by MLE and these estimated values of the parameters are close to the trues values of the parameters for all n from TPLD. The decreasing behavior of MSEs shows that the MLE works well for TPLD (Table 3-6).

8. ESTIMATION OF PARAMETERS AND APPLICATION OF TPLD

The parameters of the TPLD are estimated by MLE and incorporated by Mathematica 11.

8.1 Estimation of Parameters of TPLD

Table 1
MLE Estimates and MSEs in Parenthesis are Calculated at Various Sample Sizes for α=3, β=2 and λ=0.5

α=3, β=2 and λ=0.5					
Parameters	**n=25**	**n=50**	**n=100**	**n=200**	**n=500**
$\hat{\alpha}$	4.727	4.021	3.699	3.975	4.163
	(1.727)	(1.021)	(0.699)	(0.975)	(1.163)
$\hat{\beta}$	1.552	1.709	1.632	1.669	1.455
	(-0.448)	(-0.291)	(-0.368)	(-0.331)	(-0.545)
$\hat{\lambda}$	-0.718	-0.418	-0.278	-0.384	-0.727
	(-1.218)	(-0.918)	(-0.778)	(-0.884)	(-1.227)

Table 2
MLE Estimates and MSEs in Parenthesis are Calculated at Various Sample Sizes for α=2, β=5 and λ=0.7

α=2, β=5 and λ=0.7					
Parameters	n=25	n=50	n=100	n=200	n=500
$\hat{\alpha}$	2.578	2.648	2.461	2.483	2.419
	(0.578)	(0.648)	(0.461)	(0.483)	(0.419)
$\hat{\beta}$	5.055	4.572	4.663	4.940	4.911
	(0.055)	(-0.428)	(-0.337)	(-0.060)	(-0.089)
$\hat{\lambda}$	0.278	0.218	0.272	0.308	0.370
	(-0.422)	(-0.482)	(-0.428)	(-0.392)	(-0.330)

8.2 Applications of TPLD

In this section, the Transmuted Power Lindley Distribution is applied on real life data set to evaluate the flexibility and applicability. The applications and parameters estimation is done with R, package.

Data Set 1

Consider the dataset of failure times of devices. There are 50 observations. This data is taken from Hassan et al. (2019). The data are given as follows.

0.1, 0.2, 1, 1, 1, 1, 1, 2, 3, 6, 7, 11, 12, 18, 18, 18, 18, 18, 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85,85, 86, 86

Table 3
Parameter Estimates and Information Criterion for the Dataset of Failure Times of Devices

Model	ML Estimates	-Log L	AIC	BIC	CAIC	KS
Transmuted Power Lindley (TPL)	$\hat{\boldsymbol{\theta}} = \mathbf{0.1999}$ $\hat{\boldsymbol{\beta}} = \mathbf{0.6366}$ $\hat{\boldsymbol{\lambda}} = \mathbf{-0.3522}$	241.3918	488.784	494.520	489.305	0.1664
Alpha Power Transformed Power Lindley (APTPL)	$\hat{\boldsymbol{\alpha}} = \mathbf{5.205}$ $\hat{\boldsymbol{\beta}} = \mathbf{0.605}$ $\hat{\boldsymbol{\theta}} = \mathbf{0.258}$	282.257	570.515	569.612	571.037	0.1745
Alpha Power Transformation Lindley (APTL)	$\hat{\boldsymbol{\alpha}} = \mathbf{4.359 \times 10^{-6}}$ $\hat{\boldsymbol{\theta}} = \mathbf{8.754 \times 10^{-3}}$	296.747	597.495	596.892	597.939	0.2186
Power Lindley (PL)	$\hat{\boldsymbol{\beta}} = \mathbf{1.753}$ $\hat{\boldsymbol{\theta}} = \mathbf{1.825 \times 10^{-3}}$	305.67	615.34	614.738	615.595	0.3249
Extended Lindley (EL)	$\hat{\boldsymbol{\alpha}} = \mathbf{1.068}$ $\hat{\boldsymbol{\beta}} = \mathbf{0.217}$ $\hat{\boldsymbol{\theta}} = \mathbf{168.806}$	348.441	702.883	701.98	703.405	0.8607
Exponential (E)	$\hat{\boldsymbol{\theta}} = \mathbf{4.204 \times 10^{-4}}$	389.68	781.36	781.059	781.503	0.9645
Lindley (L)	$\hat{\boldsymbol{\theta}} = \mathbf{2.538 \times 10^{-3}}$	440.025	882.05	881.749	882.133	0.9789

Data Set 2

The data set active repair times (hr) for an airborne communication transceiver from Jan et al. (2018) consists of 40 observations. The data are:

0.1, 0.2, 1, 1, 1, 1, 1, 2, 3, 6, 7, 11, 12, 18, 18, 18, 18, 18, 21, 32, 36, 40, 45, 46, 47, 50, 55, 60, 63, 63, 67, 67, 67, 67, 72, 75, 79, 82, 82, 83, 84, 84, 84, 85, 85, 85, 85,85, 86, 86

Table 4
Parameter Estimates and Information Criterion for Active Repair Times (hr) for An Airborne Communication Transceiver

Model	-Log L	AIC	BIC	K-S statistic
Transmuted Power Lindley Distribution (TPLD)	**95.1068**	**196.2136**	**201.276**	**0.1227**
Generalized Lindley Distribution (GLD)	97.0109	199.8218	203.1995	0.1410
Lindley Distribution (LD)	97.9109	199.8218	203.1995	0.1907

Data Set 3

We study the data set corresponding to remission times (in months) of a random sample of 128 bladder cancer patients. This data is taken from Merovci (2013).

0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69, 5.49

Table 5
Parameter Estimates and Information Criterion for the Remission Times (in months)

Model	Parameter Estimates	-2LL	AIC	AICC	BIC
Exponential	$\hat{\alpha} = 0.097$	853.52	855.53	855.5617	858.37
Lindley	$\hat{\theta} = 0.196$	839.04	841.06	841.0917	843.91
Transmuted Lindley	$\hat{\theta} = 0.156$ $\hat{\beta} = 0.617$	830.3	834.31	834.406	840.01
Transmuted Power Lindley	$\hat{\theta} = 0.64$ $\hat{\beta} = 0.62$ $\hat{\lambda} = -1.0$	**821.3**	**827.3**	**827.493**	**835.86**

Data Set 4

We analyze the real data set consists of the strengths of 1.5 cm glass fibers given in Ekhosuehi and Opone (2018). The data holds 63 observations.

0.55, 0.93, 1.25, 1.36, 1.49, 1.52, 1.58, 1.61, 1.64, 1.68, 1.73, 1.81, 2.00, 0.74, 1.04, 1.27, 1.39, 1.49, 1.53, 1.59, 1.61, 1.66, 1.68, 1.76, 1.82, 2.01, 0.77, 1.11, 1.28, 1.42, 1.50, 1.54, 1.60, 1.62, 1.66, 1.69, 1.76, 1.84, 2.24, 0.81, 1.13, 1.29, 1.48, 1.50, 1.55, 1.61, 1.62, 1.66, 1.70, 1.77, 1.84, 0.84, 1.24, 1.30, 1.48, 1.51, 1.55, 1.61, 1.63, 1.67, 1.70, 1.78, 1.89

Table 6
Parameter Estimates and Information Criterion for the Strengths of 1.5 cm Glass Fibers

Model	ML Estimates	-2Log L	AIC	BIC
Transmuted Power Lindley Distribution (TPLD)	$\hat{\beta} = 4.119$ $\hat{\lambda} = -0.467$ $\hat{\theta} = 0.307$	**28.0547**	**34.05**	**40.484**
Three Parameter Generalized Lindley distribution (TPGLD)	$\hat{\alpha} = 4.944$ $\hat{\beta} = 3.429$ $\hat{\lambda} = 0.156$	28.421	34.421	40.851
Exponentiated Lindley Geometric Distribution (EGLD)	$\hat{\alpha} = 9.521$ $\hat{\beta} = 6.217$ $\hat{\lambda} = -653.220$	31.663	37.663	44.093
Power Lindley Distribution (PLD)	$\hat{\alpha} = 4.458$ $\hat{\lambda} = 0.222$	29.380	33.380	37.666
Lindley-exponential Distribution (LED)	$\hat{\alpha} = 2.612$ $\hat{\lambda} = 32.308$	62.816	66.816	71.102
Lindley	$\hat{\alpha} = 0.996$	162.557	164.557	166.700

9. COMMENTS AND CONCLUSION

In this article, we have developed a new three parameter model named Transmuted Power Lindley Distribution (TPLD). Various mathematical properties of the TPLD have been discussed. The distribution is right skewed and suitable for lifetime data sets. We calculate the values of –LL and AIC (Akaike Information Criterion), BIC (Bayesian information criterion), AICC (Akaike Information Criterion Corrected) and CAIC (corrected Akaike information criterion). Distribution with lower AIC, BIC, AICC and CAIC considered as a better model. For all the data set results suggest that the new distribution (TPL) is more flexible as compare to other existing distribution.

REFERENCES

1. Alkarni, S.H. (2015). Extended power Lindley distribution: A new statistical model for non-monotone survival data. *European Journal of Statistics and Probability*, 3(3), 19-34.
2. Arshad, M.Z., Iqbal, M.Z. and Ahmad, M. (2018). Transmuted Exponentiated Moment Pareto Distribution. *Open Journal of Statistics*, 8(6), 939-961.
3. Cakmakyapan, S. and Ozel, G. (2016). The Lindley family of distributions: Properties and applications. *Hacettepe Journal of Mathematics and Statistics*, 46(6), 1113-1137.
4. Ekhosuehi, N. and Opone, F. (2018). A Three Parameter Generalized Lindley Distribution: Properties and Application. *Statistica*, 78(3), 233-249.
5. Ghitany, M.E., Al-Mutairi, D.K., Balakrishnan, N. and Al-Enezi, L.J. (2013). Power Lindley distribution and associated inference. *Computational Statistics & Data Analysis*, 64, 20-33.

6. Hassan, A.S., Elgarhy, M., Mohamd, R.E. and Alrajhi, S. (2019). On the Alpha Power Transformed Power Lindley Distribution. *Journal of Probability and Statistics*, *2019*.
7. Jan, R., Jan, T.R. and Ahmad, P.B. (2018). Exponentiated Inverse Power Lindley Distribution and its Applications. *arXiv preprint arXiv:1808.07410*.
8. Kemaloglu, S.A. and Yilmaz, M. (2017). Transmuted two-parameter Lindley distribution. *Communications in Statistics-Theory and Methods*, 46(23), 11866-11879.
9. Merovci, F. (2013). Transmuted lindley distribution. *International Journal of Open Problems in Computer Science and Mathematics*, 238(1393), 1-20.
10. Nadarajah, S., Bakouch, H.S. and Tahmasbi, R. (2011). A generalized Lindley distribution. *Sankhya B*, 73(2), 331-359.
11. Shanker, R. and Mishra, A. (2013). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research*, 6(4), 64-71.
12. Shibu, D.S. and Irshad, M.R. (2016). Extended new generalized Lindley distribution. *Statistica*, 76(1), 41-56.
13. Shaw, W.T. and Buckley, I.R. (2009). The alchemy of probability distributions: Beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *arXiv preprint arXiv:0901.0434*.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 309-329

CUSTOMER SATISFACTION TOWARDS MOBILE NETWORK (MOBILINK, ZONG)

Faiza Kanwal and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Pakistan
Email: faizakanwal998@gmail.com
zaheerabbas@uog.edu.pk

ABSTRACT

For the last few years, telecom industry in Pakistan is growing day by day as a profitable industry with ambition of maximum profit by providing high services (especially internet) quality according to needs and requirements of the users. Competition between two mobile network (Mobilink, Zong) has also extended as companies are increasingly focusing on maintaining existing and new customer. In which we investigate major determinants of customer needs towards mobile network service operators. For the present study, the university students are target. A well-organized questionnaire was used to collect information from 200 respondents by using simple random sampling. Descriptive statistics, correlation and Ordinal logistic regression model was used as data analysis tool. Results shows that fair prices, providing good services, loyalty towards customer and coverage are main factors which highly affects the customer satisfaction. Different facilities of service (internet packages, call packages, SMS packages) according to customer needs highly affect the customer satisfaction.

KEYWORDS

Customer Satisfaction, quality of service, customer loyalty, telecom industry, self-assessment, comparative analysis, Regression model, profitability.

1. INTRODUCTION

For each individual mobile network administration has turned into, to contact with their relatives, family, companions and with the agents. In Pakistan, there are numerous Mobile network organizations which offer paid ahead of time and postpaid administrations. The fulfillment level of mobile network administration clients' changes with their desires. This can likewise be co-related with the capability of administrations a mobile network organization can offer to its clients. Mobile network is normally a general term for a huge swath including Technologies that send data over long separations. It very well may be characterized in straightforward words "Mobile network administrations interface individuals and give the preferred position to convey their messages. As such it is remote association through which message can be sent from sender to beneficiary. In 1962 Telecommunication created without precedent for Pakistan. In 1991 there was just a single Telecommunication supplier organization which was working by the name of Pakistan Telecommunication Company Limited (PTCL). Be that as it may, in 1994 the circumstance changed on the grounds that Multi-national organizations began business in Pakistan and proposed first mobile network administration Global System for Mobile communications (GSM) in different areas of country.

In Pakistan 1994 Mobilink starts its activity in Pakistan. Mobilink was first Service provider organization to work on hundred percent advanced GSM innovation in Pakistan. Mobilink entered first and high piece of growth and market development that is the reason Mobilink moved toward becoming business sector pioneer in portable organizations in Pakistan. Today Mobilink driving with 36 users. Zong was proposed in 2008 as claimed auxiliary of China Mobile Communications Corporation. It as of now holds 20% of the piece of the overall industry in the nearby telecom industry (Jazz represents 38%, Ufone 14% and Telenor 28%). To break into the market, Zong retain at first itself as a service provider for the 'majority' with a value of price focused strategy. By knowing the factors, they create satisfaction among the users and they can take measures to retain the existing customers and also to gain new customers. In this aspect, the present study focuses on customer’s satisfaction with special reference to mobile network in University of Gujrat psychological state that express the appraisal of relationship between the customers and company services. Satisfaction use the following psychological elements: Subjective/thinking, impressive/emotional and behavioral.

Martensen et al. (2000) conducted a study. The main focus of this study was which drivers for customer satisfaction and loyalty? The methodology of this paper is based on the recently introduced European customer satisfaction index. Telephone interviews were conducted to obtain a representative sample of the respective companies’ current customers. For each identified company 250 customers were interviewed, adding up to more than 8000 interviews in total. The study results reveal that customers must be satisfied with services.

Wallin and Leijon (2000) conducted a study. In this research 15 major telecom operators had a surveyed regarding their state and future plans for the network management. This survey covers more than 100 million of the customers. The well-structured questionnaires are distributed to 20 operators by email.

(Gerpott el at, 2001) in coming future time, the company having chances to make more profit by keeping their customers satisfied. Emotional value refers to utility derived from feelings or affective states that a service provider engenders (Sweeney and Souter, 2001). This may include feeling good, as well as enjoyment and happiness during the use of the mobile network. In a retailing context, Sweeney and Souter (2001) found that emotional value is the strongest predictor of consumers’ purchase intention.

Liu (2002) found that the choice of a cellular phone is characterized by two attitudes: attitude towards the mobile phone brand on one hand and attitude towards the network on the other. Samuel (2002) observed that most of the respondents consider size, quality, price, instrument servicing is an important factor for selecting the handset while majority of the respondents are satisfied over the payment system, quality of services, coverage area and the process of attending the complaints regarding their mobile service provider.

Based on a study entitled Wireless network quality assessment, J.D. Power and Associates (2003) concluded that: "Carriers that offer superior network quality will increase their likelihood of attracting new customers and retain more of their existing base" This satisfaction may influence the concerned company by repurchase, purchase of more products, positive word of mouth and willingness of customer to pay more for the particular brand. Any business is likely to lose market share, customers and investors if it fails to satisfy customers as effectively and efficiently as its competitors is doing (Anderson, Fornell and Mazvancheryl, 2004).

Serkanaydin and Ozer (2005) conducted a study. This study was designed to build National Consumer Satisfaction Index for Turkey, drawing on models already in existence in Sweden, the USA, Norway and the Europe Union. Data was collected from a sample size of more than 1500 subscribers. Interview and questionnaire were used as data collection method. Structural equation modeling was applied.

The study of Ahn, Han and Lee (2006) shows that when the customers, do not get their complaints considered properly, they start looking for other brands. Sometimes, telecom service providers take considerably longer time to resolve the problems like network coverage or call quality, the customers do not wait for long and hence they lose satisfaction with that particular brand (Ahn, Han and Lee, 2006).

Roos & Gustafson, (2007) argued that there are two different terms to describe customers during the switching process; active and passive customers Passive switchers on the other hand are customers who get influenced by a third party and may not be aware that switching could be an option. Mohsen (2008) conducted a study. The aim of this research is finding the factor which contributes to customer satisfaction in Pakistan in mobile cellular services. The research explains the knowledge by explaining factors, which contribute to satisfying the customers of Pakistani mobile cellular service. The results suggest that customer satisfaction of mobile cellular users in Pakistan constitutes four factors price, transmission quality, usage ease and service support.

Mohan and Aranganathan (2009) in their study "Conceptual framework of Mobile Marketing Spamming the consumer around the world" found that, Mobile phones can also be an extremely cost effective communication channel as well as an efficient way of delivering a marketing message. The study has pointed out the success of mobile advertising will directly depend upon the market penetration and the success of Mobile Internet users.

Mallikarjuna et al., (2010) in their study "Customer switching in mobile industry - an analysis of pre-paid mobile customers in AP circle of India" found that, Switching is quite high in the pre-paid customer segment due to low switching costs and competitive tariff plans. Hence, investment in network and technology should go on to improve the geographic coverage, seamless connectivity and speed, quality of service such as voice call for enhancing customer loyalty.

Anand Shankar Raja M (2011) his study entitled on "A study on customer satisfaction towards BSNL with special reference to the city of Coimbatore" focuses to know whether customer receive the service in time and is it fulfilling their needs to desired levels. A simple statistical tool has been employed for the purpose of analysis of data. The study concludes that most of the customers are satisfied by the service provided by the BSNL. The level of dissatisfaction is mainly due to poor signal and low network coverage. The present scenario of the service provider must be focused with to resolve the problem of customer retention and satisfy the customers.

Raj Kumar Paulrajan and Harish Rajkumar (2011) in their study examine the perception choice in selecting the cellular mobile telecommunication service providers. The result of the study indicates that in selecting the telecommunication service provider, communication and price has a significant impact in choice of the consumer preference nevertheless the product quality and availability has a significant impact in selecting the mobile telecommunication service provide

Chintan Shah (2012) in his study entitled “Consumer Preferences for Mobile Service Provider” investigates the factor considered by the customers to shape their preference for the mobile service provider. For the study 150 mobile users of Bardoli city are surveyed. The researcher finds that the service quality and brand image, service charges and plan, and network quality plays a very important role in preferring mobile service provider. Jegan and Sudalaiyandi (2012) in their study on the consumer preference and their satisfaction towards the mobile phone service providers find that the call tariffs, network coverage and brand image encourage the consumer’s preference and satisfaction

Myilswamy, Ratheesh Kumar (2013) find that the postpaid and prepaid customer prefer to use Airtel mobile service because of the features, Rent, Tower coverage, Talk time, Advertisement and Corporate schemes. Finally, the study suggests that by paying special attention on these factors develop the business by satisfying the consumers. Zafar (2013) examines the impact of the mobile service attributes like call rates, service quality, service availability; promotion and brand image of both the gender’s purchasing decision. The result reveals that the male and female consumers have different preference in making the purchase decision to the avail mobile service. The study also discovers that the female consumers are tough to satisfy than the male consumers.

Zohaib Ahmad and Junaid Ahmad (2014) in their study titled “Consumer Purchase Behavior in Cellular Service Sector” says that the dominant factors quality, price, promotions, and social factors reflects the latest buying behavior of people or not. The study reveals that the social factor is the most dominating factor which determines the purchase behavior and basically reflects the societal image of the consumers.

A.K. Antony (2016) “A study on consumer satisfaction towards reliance jio connection palakad area kerala state” The present study aims to know the level of satisfaction towards the reliance jio net connect. He concluded most of the respondents are highly satisfied with the reliance net connect for its network coverage.

Hematherpatan (2016)2 “A study on customer’s perception towards jio sims” The study was made to know customer perception towards jio sim. During the study it was found that network level is very low, respondents had closenn jio for its attractive schemes.

2. SIGNIFICANCE AND OBJECTIVE OF STUDY

Customer satisfaction is a measure of how products and services supplied by a company meet or surpass the customer’s expectations. Customer expectation is the needs, wants, and preconceived ideas of a customer about a product or a service. If customer expectations are meet, then the customer is satisfied.

It has been observed through the literature review that network connectivity, coverage, quality of service, facilities, price, promotions and customer care and many other factors plays an important role for high customer satisfaction.

Following objectives have been formulated to achieve the above sited goals:

- To explore the basic features of data.
- To describe the customer satisfaction about Mobilink and Zong service with the statistical summaries.
- The main objective of the research is to identify those factors which are used into customer's satisfaction in mobile communication
- To check whether the customers are satisfied with the operator service.
- To evaluate the difference between the opinion of Mobilink and Zong service users.

3. FACTORS IN STUDY

There are the following factors in my study are as follows:

Customer Satisfaction:

Satisfaction defined as characteristics that fulfill the either the needs or wants of customer's ingredient way than competitors. Different researchers explained satisfaction in different ways. Customer satisfaction is a measure of how well services provided by the company and received by the customers or company meets the requirements of customers. Now-a-days, customers have more expectations from its service providers. Satisfaction of a consumer depends upon the quality of brand characteristics that provided by a company. For the intention of customers, it is important to satisfy the customers. The unsatisfied customers of a company do not take time to change that service. Low quality services can also bring dissatisfaction. When examined as a whole, three general components can be identified:

- Consumer satisfaction is an emotional response.
- The response refers to a particular focus (expectations, product, consumption experience etc.)
- The response occurs at a particular time (after consumption, after choice, based on past experience etc.)

Network Connectivity:

Network connectivity is the ability to access the network in a timely fashion. Timely network access is a crucial component to reliability. Thus, when someone needs to make a call promptly, getting access to the network and setting up a call becomes a critical factor in determining satisfaction with reliability. Every customer having different expectation from the services provider firm when they get more and better services the customer may become satisfied and trust of the customer built on the services provider too.

Coverage:

Coverage is the extent of the space in which a customer could use the service. It is a convenience that many customers find key to their satisfaction. Several studies mentioned in this research revealed that size of coverage is one of the main factors customers use in selecting a provider. It is more probable that many wireless networks will be available, and all of them will provide services over a variety of geographical coverage areas at various speeds (Hills & Johnson, 1996).

Quality of Service:

Perceived quality depends upon the past experience, word of mouth and the future anticipation of quality of the cellular service. Every cellular user requires best quality accordingly. Since companies are going on continuous improvement in quality through latest technology, installing costly equipment's, trying to improve call clarity and coverage, according to customer perception (Kim et al, 2004). For users as well as cellular service providers (CSP) it is necessary to know the relationships between objective and subjective application quality of service and network quality in order to identify technical reasons for user perceived quality problems (Information Society Technologies (IST), 2005).

Facilities:

Perceived entertainment refers to "the extent to which the activity of using the computer (technology) is perceived to be enjoyable in its own right, apart from any performance consequences that may be anticipated." (Davis et al., 1992) Davis and his colleagues also examined the impact of entertainment on usage intention in an extension to TAM (Davis et al., 1989). Mobile network service can make users feel entertainment; we propose that maybe gender can make differences in the entertainment.

Affordability/Prices:

Price is one of the factors in the marketing mix and is used in order to stimulate the consumer. Pagani (2004) shows that price is one of the most important determinants of adoption of multimedia mobile services in his works. Man and women have different reactions for the factors of price. It has showed a lot in the works that gender make difference in price. Cellular service charges depend on many factors like technology, services, maintenance and taxes etc. Users are willing to pay certain amount of money for using cellular services, and service providers want a profit from their share. In order to increase the share, the call rate and connection fee should be reduced (Parvez, 2005) or should be put adjusted according to customer mental accounts terms (Pirc, 2005) Literature explored showed substantive role of price and quality service with customer satisfaction, therefore, this study also effects at customer satisfaction in the mobile network of Pakistan were evaluated.

Promotions:

Promotion is a conceptualized as an emotional response which results from surprising and positive levels of performance. According to Sundhar [19], customer satisfaction is meet the customer expectations, but delighting customer is involved surpassing it and taking the entire experience to an emotional plane. It can be said that delighting customer will attract the customer to become loyal to the business that they attached, and increase the value of lifetime. Most of the mobile network operator will give promotion or bonus to their customers in delighting them. The promotion usually not last long. It was due to festive season or it being given to the customers which are loyal to the company. It also can be known as promotion because when the company delighting the customer they will share their experiences through different channels such as social media, websites, apps and as word of mouth. Throughout this approach it will indirectly branding the companies and give big positive impact due to customers' satisfaction. In this context, it is critical to gauge factors influencing customer satisfaction with the mobile data services delivery.

Customer Care:

Customer service is the direct one-on-one interaction between a consumer making a purchase and a representative of the company that is selling it. Most retailers see this direct interaction as a critical factor in ensuring buyer satisfaction. Customer service representatives are fully informed and have the latest information and the company's products and policies.

4. METHODOLOGY

Source of Data:

The population of this research is consisted of all students who are the customers of Mobilink and Zong they are enrolled in 2019 of University of Gujrat Hafiz Hayat Campus. Where the entire cellular services companies providing the mobile services. Taking Mobilink and Zong for the research, because these two networks are pioneer in the cellular industries in region Gujrat.

$$n = \frac{N}{1+Ne^2} \qquad \text{(Yamne, 1967)}$$

where n and N are sample and population size respectively and 'e' is margin of error. Let $e = 0.05$ and $N = 506$ then our required sample is 200. A sample of 200 respondents selected from both networks. In which 100 respondents were from the Mobilink and the rest of 100 respondents were from the Zong. The student of social sciences (Statistics, Sociology, Computer Science, Information Technology, Business Administration, and English) are 506 students who are the customers of Mobilink and Zong network. We have simple random sampling with proportional allocation.

Survey Instrument:

A well-structured questionnaire will be used as research instrument for data collection. There are 41 items using 5-point Likert-scale.

Data Analysis Techniques:

Descriptive statistics is used to get summary of the data. It provides the features and picture of data. Measures of central tendency, measure of dispersion used in quantitative and percentage is used to get results about qualitative.

For the purpose to fulfill the objective of the study, we have to used Discriminant Function analysis, Ordinal Logistic Regression and Neural network.

Ordinal Logistic Regression:

Data was also analyzed by ordinal logistic regression. Ordinal logistic regression is statistical tool used when the response variable is categorical. The outcome of the model provides predicted probabilities for each level of the response. Ordinal data found in the area such as psychometrics, customer satisfaction and the taste testing where the outcomes are related to latent variable but we can only observe relative levels. Logistic regression sometimes called “Logistic Model or Logit Model” is used for prediction of probability of occurrence of an event by fitting data to a logistic curve. It is a generalized linear model used for binomial regression. Like many forms of regression analysis, it makes use of several predictor variables that may be either numerical or categorical. The ordinal logistic regression model is a simple extension of the binomial logistic regression model.

Artificial neural network Artificial neural networks (ANNs) are distributed and parallel information systems which simulate the human brain to process information. ANNs simulate human cognition by modeling the inherent parallelism of neural circuits in the brain using mathematical models of how the circuits function (Spangler, May, & Vargas, 1999). They could be used to learn complex patterns of information and generalize the learned information (Venugopal & Baets, 1994). ANNs are massively parallel interconnections of simple neurons as a collective system and consist of many non-linear computational elements called nodes which are interconnected through direct links. One or more input values are took and combined into a single value, and transform into an output value. It can be used in applications where a model of a system is required based on an input set of training data. ANNs are widely used to examine the complex relationship between input variables and output variables (Nelson & Illingworth, 1994). Since early 1980, there has been an explosive growth in pure and applied research related to neural networks. During this period, the multilayer feed-forward neural networks was introduced and immediately found wide application in many fields (Rumelhart & McClelland, 1986).

Discriminant function analysis a linear discriminant function, to discriminate two mobile networks (Mobilink and Zong) according to related factors, was formed. Network type was taken as a dependent variable whereas, Coverage, Quality of service, Facilities, prices, Promotions and Customer Care were taken as independent variables. Data was coded into SPSS 21.0 (version). Discriminant analysis was carried out for testing the hypothesis that the selection of network (Mobilink and Zong) is independent of the factor like Coverage, Quality of service, Facilities, prices, Promotions and Customer Care.

5. RESULTS AND DISCUSSION

Reliability of the questionnaire is a key factor for statistical results. Table 1 consists of the test of reliability; the value of Cronbach’s alpha based on standardized item is 0.908 which indicates that the data is reliable for further statistical analysis.

Descriptive statistics are considered primarily to explain the basic features of the data. It provides simple summaries about the sample measures. Table 2 contains the mean and standard deviation of age of the respondents. Results show that the minimum age of the respondent is 18 years and maximum is 45 years. The mean age is 23.358 years with S.D 4.3247; it shows that most of the respondents are within 19 to 28 years of age.

Table 3 contains the sample information; there were 46.5 and 53.5 percent male and female respectively and 40.9 percent male from Rural, 44.1 percent male from Urban and other from suburban 31.8 percent female respondent from Rural, 56.1 percent female from Urban and other from suburban.

Table 4 describes the percentage of ordinal scale variables. Table 4 shows that average rank of respondents on the statement that coverage is better in your area overall decision is 38.0, it means on the average respondents are agree with that statement. Average rank of the respondents on the statement that quality of service decision is 35.6, it means on the average respondents are agree on the above statement. Average rank of respondent on the statement that facilities influence your decision is 35.9, it means on the average respondents are agree with that statement. Average rank of the respondents on the statement that fair Prices of your mobile network is 32.2, it means on the average respondents are agree with that statement. Average rank of respondents on the statement that Promotions influence on your decision to choose this mobile network is 37.5, it means on the average respondents are agree with that statement. Average rank of the respondents on the statement that give priority to customer care satisfaction in selection of mobile network decision is 37.0. 6.9% respondent strongly disagrees, 14.6% respondents disagree, 18.3% neutral, 38.0% agree and 22.2% strongly agree about the Coverage factor. Table 4 contains the percentages and average rank of other variables which can be interpreted in similar manner.

Results on test of equality of group means for the predictive variables are presented in Table 5. Table shows that there is a significant difference between the two mobile network (Mobilink and Zong). P-values are greater than 0.05 therefore we conclude that there is no discriminate the customer satisfaction significantly in two networks.

In Table 6 to examine whether there is an equal variance between or among the Mobilink and Zong network, Box's M test has been worked out. We reject to the null hypothesis and Table shows that there is an equal variance between the Mobilink and Zong with Box's M value 53.684 and p value 0.000.

Strength of Relationship between Predictive variables and Outcome Groups:

To measure the strength of relationship between predictive variables and mobile networks, canonical discriminant functions were used in the analysis; results are presented in the Table 7. Results show that there is a high degree positive canonical correlation (0.694) exists between predictive variables and two mobile networks.

Statistical Significance and Discrimination Power of the Model:

Table 8 shows a Wilk's Lambda value 0.519 which is close to 0.5. Though it ranges from 0-1 but, value close to 0 is considered well which indicates better discriminating power of the model. The chi-square test shows the highly significant discrimination between the two mobile networks (Mobilink and Zong) at 5 % level of significance with Chi-square value = 63.70 and p = 0.000.

Identification of the Variables Discriminating the Groups:

In Table 9 to identify the variables significantly discriminating Mobilink with Zong, this study includes six independent variables namely Coverage, Quality of Service, Facilities, Prices, Promotions, and Customer care in the discriminant function. Results of the function are presented in table 5. Results show that variable Quality of Service is best predictor with the coefficient value 1.016 followed by Coverage (0.373) Facilities (0.152), Prices (0.129), and Customer care (0.055). Only one variable ‘Promotions’ was found to have negative coefficient (-0.045) which shows a negative contribution towards the discriminant function.

In Table 10 Since this study aims to identify the variable that has best prediction power for discriminating the students, it is very important to know how these new students could be discriminated/grouped in any of the predefined group. The output shown in table 7 helps to answer this question by applying the following unstandardized linear discriminant function:

$$Y = \alpha + \beta_1 * \text{Coverage} + \beta_2 * \text{Quality of service} + \beta_3 * \text{Facilities}$$
$$+ \beta_4 * \text{Prices} + \beta_5 * \text{Promotions} + \beta_6 * \text{Customer care}$$

where, Y = Discriminant Score, α = Constant; β_1, β_2, β_3, β_4, β_5 and β_6 are the coefficient of the variables namely Coverage, Quality of Service, Facilities, Prices, Promotions, and Customer care

Ordinal logistic regression is used to explore and examine the relationship between the satisfaction of customers and explanatory variables. Ordinal regression is used to enhance the exploratory variables that effect the satisfaction of the customers.

In Table 11 “Intercept Only” shows a model that simply fits an intercept to predict the dependent variable while “Final” indicates a model that contains specified independent variables. Chi-square value statistic is the difference in the -2 log-likelihoods between the initial and final models. The null hypothesis is that all of the regression coefficients in the model are equal to zero. The p-value is less than 0.05 so we conclude that at least one of the regression coefficients in the model is not equal to zero.

In Table 12 “Goodness of Fit” table is also at test of the model and contains Pearson’s chi-square statistic and Deviance chi-square statistic. These statistics test whether observed data is consistent with the fitted model. As p-value is large than 0.05 so we conclude that the model is good fitted.

In Table 13 Pseudo R-square table the R-statistic cannot be exactly computed for ordinal regression model so these approximations are computed instead. Large Pseudo R-square statistic indicates that more of the variation in the response is explained by the model, to a maximum of 1. The Pseudo R-square table shows three measures of overall effect of size, which Negelkerke R-square is 0.739 expressed as high effect size

The parameter estimation table, is the core of the analysis. The coefficients, their standard errors, wald test, p-values and 95% confidence interval have been given in this

Table 14. Threshold are the cut points for dependent variable, it is like intercept in ordinary linear regression. The facilities factor on network services is plying a significant role as well as other variables are insignificant.

Table 15 checked the proportional odds assumption. This leads to the test of parallel line. Our significant of Chi-square statistic is less than 0.05 so we reject the null hypothesis. We conclude that the assumption is not hold.

Table 16 shows the distribution of mobile network customers according to satisfaction. Here the low satisfaction means the customers who have low level of satisfaction are 18%. The medium satisfaction means the customers who have medium level of satisfaction are 48%. The high satisfaction means the customers who have high level of satisfaction are 33%.

Table 17 shows the case processing summary in which training sample includes 135 customers and testing sample involve 65 customers.

In Table 18 classification matrix shows that in the tanning sample the correct classification customers have low satisfaction is 16.3%, the correct classification in the medium category means the customers that have the average level of satisfaction is 52.6% and the correct classification in high category is 31.1%. In the testing sample percentage of correct classification in low category is 18.5%, the percentage of correct classification in medium category is 47.7% and percentage of correct classification in high category is 33.8%. The overall correct percentage of classification in the training sample is 74.8%.

Figure 1 shows graphical representation of classification table. It plots observed vs. predicted pseudo probability of categories. The portion of the box plot above the 0.5 mark on the X-axis represents correct predictions as there are only three categories in the target variable. In the first box plot the observed low and medium category is below 0.5, the high category is above 0.5. In the next box plot only the middle category is above 0.5, that shows the correct classification. In the next plot only observed high category is above 0.5 that shows correct classification. So we can say that the medium category gives the highest percentage of correct classification.

Table 19 shows the normalized importance percentage of factors of customer satisfaction. Among all independent factors facilities factor is most important with percentage of 100, coverage is the second most important factor with 18.9% and then quality of service, promotions, prices and customer care with 17.2%, 15.6%, 12.6%, 12.1% respectively.

Figure 3 shows the graphical representation of three layered model. Model diagram of the neural network shows the relationship of hidden layers, input layers and output layers. The colored lines that connect the input, hidden and output layer have different interpretation. Grey lines show the weights regarding particular relationship are positive. Blue lines are the indication of the negative weights which are given in the Table 20.

6. CONCLUSION

- Customer satisfaction is totally concerned with the services provided to and perceived by the customers, if there is a high matching between seeming and provided services, then customer satisfaction level is very high, that directly leads to high customer loyalty.
- We suggest that Mobilink and Zong compete at the same level therefore there is no discriminant the customer satisfaction.
- And these six factors of customer satisfaction affect to the satisfaction of customers but only facilities are significant.

Suggestions:

- It is suggested that the sample should be taken large to get more precise results.
- Sincere efforts should be adopted to measure the customer's satisfaction about mobile network.
- The study was conducted in the limited sources & time. It can be improved by expanding it to national level.

REFERENCES

1. Lee, J. and Mills, J.E. (2007). Exploring tourist satisfaction with mobile technology. In *Information and Communication Technologies in Tourism 2007* (141-152). Springer, Vienna.
2. Goode, M.M., Davies, F., Moutinho, L. and Jamal, A. (2005). Determining customer satisfaction from mobile phones: A neural network approach. *Journal of Marketing Management*, 21(7-8), 755-778.
3. Hanif, M., Hafeez, S. and Riaz, A. (2010). Factors affecting customer satisfaction. *International Research Journal of Finance and Economics*, 60(1), 44-52.
4. Nimako, S.G. and Mensah, A.F. (2012). Motivation for customer complaining and non-complaining behaviour towards mobile telecommunication services. *Asian Journal of Business Management*, 4(3), 310-320.
5. Gyasi, S. (2012). Customer dissatisfaction and complaining responses towards mobile telephony services. *The African Journal of Information Systems*, *4*(3), Article-1., 84-99.
6. Gao, L., Waechter, K.A. and Bai, X. (2015). Understanding consumers' continuance intention towards mobile purchase: A theoretical framework and empirical study–A case of China. *Computers in Human Behavior*, 53, 249-262.
7. Mensah, A.F. (2012). Customer complaint behaviour towards mobile telephony services–cases of mtn and tigo Ghana. *International Journal of Business and Social Research (IJBSR)*, 2(4), p182.
8. Sivasubramanian, M. and Balaji, M.S. (2014). Service Quality and Its Impact on Customer satisfaction towards mobile sector of Tamilnadu. *Asia Pacific Journal of Research*, I(XVII), 139-147.
9. Kavitha, T.N.R. and Murugan, S.D.A Study on Customer Satisfaction towards Vodafone Sim Card Special Reference In Madurai City. *IOSR Journal of Business and Management*, 29-32.

10. Mohsan, F., Nawaz, M.M., Khan, M.S., Shaukat, Z. and Aslam, N. (2011). Impact of customer satisfaction on customer loyalty and intentions to switch: Evidence from banking sector of Pakistan. *International Journal of Business and Social Science*, 2(16), 263-270.
11. Kang, D. and Park, Y. (2014). Based measurement of customer satisfaction in mobile service: Sentiment analysis and VIKOR approach. *Expert Systems with Applications*, 41(4), 1041-1050.
12. Hafeez, S. and Hasnu, S.A.F. (2010). Customer satisfaction for cellular phones in Pakistan: A case study of Mobilink. *Business and Economics Research Journal*, 1(3), 35-44.
13. Hanif, M., Hafeez, S. and Riaz, A. (2010). Factors affecting customer satisfaction. *International Research Journal of Finance and Economics*, 60(1), 44-52.
14. Vijay, P. and Krishnaveni, V. (2016). Customer Preferences towards the mobile network service provider-A study with the special reference to coimbatore City. *International Journal of Management Research and Reviews*, 6(10), 1386.
15. Munyanti, I. and Masrom, M. (2018). Customer Satisfaction Factors towards Mobile Network Services. *Journal of Advanced Research in Business and Management Studies*, 13(1), 9-18.
16. Khan, S.A., Khattak, M. and Ali, A. (2013). Comparison of service performance and customer satisfaction in telecom sector. *International Review of Basic and Applied Sciences*, *1*(2), 51-60.

APPENDIX

Table 1
Test of Reliability

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of items
0.907	**0.908**	**41**

Table 2
Descriptive Statistics of Quantitative Variable

Variables	Minimum	Maximum	Mean	Std.Deviation
Age in years	**18**	**45**	**23.3582**	**4.32477**

Table 3
Percentages of Demographic Variables

Variable	Categories					
Gender	**Male**			**Female**		
	46.5			**53.5**		
Region	**Rural**	**Urban**	**Suburban**	**Rural**	**Urban**	**Suburban**
	40.9	**44.1**	**15.1**	**31.8**	**56.1**	**12.1**

Table 4
Percentages, Mean and Standard Deviation of all Ordinal Scale Variables

Factors/Variables	SD%	D%	N%	A%	S.A%	Mean	S.D
Coverage							
Network coverage is reliable.	**6.1**	**16.2**	**17.2**	**37.4**	**23.2**	**3.56**	**1.186**
Voice clarity is good.	**7.1**	**14.6**	**18.2**	**34.3**	**25.8**	**3.57**	**1.219**
Delivery of SMS, MMS, voice messages and other services are timely	**6.1**	**17.2**	**17.2**	**35.9**	**23.7**	**3.54**	**1.199**
The signals are strong even inside the buildings include basement.	**7.6**	**14.1**	**18.2**	**40.4**	**19.7**	**3.51**	**1.178**
The signals are strong in all cities covered by this mobile network.	**7.6**	**9.6**	**21.2**	**42.4**	**19.2**	**3.56**	**1.133**
Overall	**6.9**	**14.6**	**18.3**	**38.0**	**22.2**	**3.44**	**1.166**
Quality of Service							
Do you think mobile network fulfill the demand/requirements of the customer?	**9.6**	**13.6**	**24.2**	**34.8**	**17.2**	**3.38**	**1.210**
Do you think this mobile network has fastest internet speed to other.	**4.5**	**28.8**	**22.7**	**28.3**	**15.7**	**3.22**	**1.157**
It is usually easy to get through while sending or receiving calls.	**7.1**	**15.7**	**16.2**	**43.4**	**17.7**	**3.49**	**1.161**
Do you think this mobile network provide fastest 4G service through.	**7.1**	**13.1**	**22.2**	**38.9**	**18.7**	**3.49**	**1.148**
Do you think this service is equipped with modern technology.	**6.1**	**17.2**	**18.2**	**37.9**	**20.7**	**3.50**	**1.174**
Would you like to change your current mobile network in future.	**12.1**	**16.7**	**21.7**	**30.3**	**19.2**	**3.28**	**1.286**
Overall	**7.7**	**17.4**	**20.8**	**35.6**	**18.4**	**3.36**	**1.133**
Facilities							
Do you think that the loan service is good?	**14.1**	**12.6**	**21.7**	**35.9**	**15.7**	**3.26**	**1.271**
Do you think the balance sharing is good?	**10.1**	**17.2**	**21.7**	**32.3**	**18.7**	**3.32**	**1.245**
Do you think news service of entertainment menu is good?	**9.6**	**13.1**	**26.3**	**34.3**	**16.7**	**3.35**	**1.186**
Do you think the joke service of entertainment menu is good?	**8.1**	**17.2**	**22.2**	**33.3**	**19.2**	**3.38**	**1.218**
Do you think mobile network app is helpful for users?	**7.1**	**14.1**	**18.7**	**33.8**	**26.3**	**3.58**	**1.218**
Mobile app of this service that you can easily understand.	**6.1**	**16.2**	**17.2**	**37.4**	**23.2**	**3.56**	**1.186**
Service provider are always willing to help you?	**7.6**	**14.1**	**18.2**	**40.9**	**19.2**	**3.50**	**1.174**
Service provider are giving attention to your problems?	**7.6**	**10.1**	**21.2**	**42.9**	**18.2**	**3.54**	**1.129**
Are they getting your feedback.	**7.1**	**16.7**	**25.8**	**33.3**	**17.2**	**3.37**	**1.158**
Overall	**8.7**	**14.6**	**21.4**	**35.9**	**19.3**	**3.36**	**1.22**

Factors/Variables	SD%	D%	N%	A%	S.A%	Mean	S.D
Prices							
Are you satisfied with the call rates of mobile network?	**16.2**	**14.1**	**21.7**	**32.8**	**15.2**	**3.17**	**1.305**
Do you think the call rates of this mobile network on the other networks are affordable?	**10.6**	**15.7**	**27.8**	**27.8**	**17.7**	**3.20**	**1.324**
Do you think the call rates of this mobile network for international calls are affordable?	**15.2**	**13.1**	**28.8**	**23.2**	**19.2**	**3.20**	**1.324**
Do you think that the charges of this mobile network is not expensive?	**11.1**	**17.2**	**20.2**	**34.3**	**17.2**	**3.29**	**1.252**
Do you thinks this mobile network provide reasonable internet packages?	**6.1**	**13.6**	**19.2**	**41.4**	**19.7**	**3.55**	**1.133**
Overall	**11.7**	**14.8**	**23.5**	**32.2**	**17.8**	**3.34**	**1.302**
Promotions							
Information/message in Top-up card and useful for customers.	**9.6**	**13.6**	**24.2**	**34.8**	**17.7**	**3.38**	**1.210**
Promotional offers are communicated properly.	**5.6**	**15.2**	**23.7**	**41.4**	**14.1**	**3.43**	**1.082**
Billing methods are transparent and simple through online system.	**4.5**	**12.1**	**27.3**	**35.4**	**20.7**	**3.56**	**1.087**
Retail outlets are available in enough numbers and are easily accessible.	**8.1**	**15.7**	**20.2**	**37.9**	**18.2**	**3.42**	**1.189**
Do you think mobile network help users by promote mobile app.	**5.1**	**14.1**	**21.7**	**37.4**	**21.7**	**3.57**	**1.137**
Overall	**6.5**	**14.1**	**23.3**	**37.5**	**18.4**	**3.44**	**1.079**
Customer Care							
service provider give attention to your complaints?	**12.6**	**11.6**	**26.3**	**30.3**	**19.2**	**3.32**	**1.264**
Operators communicate in a language that you can easily understand.	**6.6**	**10.1**	**14.1**	**47.5**	**21.7**	**3.68**	**1.120**
Service providers are capable of answering your queries?	**9.1**	**14.6**	**17.7**	**31.8**	**26.8**	**3.53**	**1.220**
Service providers have required skills and knowledge to answer your queries?	**9.6**	**12.6**	**20.2**	**39.4**	**18.2**	**3.44**	**1.202**
Service providers are sincere and patient in resolving your problems.	**6.1**	**11.6**	**24.2**	**33.8**	**24.2**	**3.59**	**1.153**
Service providers provide security and maintains customer confidentiality.	**4.4**	**14.1**	**23.2**	**38.9**	**19.2**	**3.54**	**1.093**
Overall	**8.0**	**12.5**	**21.0**	**37.0**	**21.5**	**3.46**	**1.224**

Factors/Variables	SD%	D%	N%	A%	S.A%	Mean	S.D
Customer Satisfaction							
Are you satisfied with the packages of mobile network	**14.1**	**12.6**	**21.7**	**35.9**	**15.7**	**3.26**	**1.271**
Are you satisfied with operators services of mobile network	**10.1**	**17.2**	**21.7**	**32.3**	**18.7**	**3.32**	**1.245**
Are you satisfied with the general services of mobile network	**9.6**	**13.1**	**26.3**	**34.3**	**16.7**	**3.35**	**1.186**
Are you satisfied with the prices of mobile network?	**8.1**	**17.2**	**22.2**	**33.3**	**19.2**	**3.38**	**1.207**
Are you satisfied with the entertainment menu of mobile network	**7.1**	**14.1**	**18.7**	**33.8**	**26.3**	**3.58**	**1.218**
Overall	**9.9**	**14.8**	**22.1**	**33.9**	**19.3**	**3.37**	**1.210**

Table 5
Wilk's Lambda and F Statistics for Test of Equality

Tests of Equality of Group Means					
	Wilks' Lambda	**F**	**df1**	**df2**	**Sig.**
Coverage	.997	.677	1	198	.412
Quality of service	.985	3.051	1	198	.082
Facilities	1.000	.040	1	198	.842
Prices	1.000	.007	1	198	.933
Promotions	1.000	.069	1	198	.793
Customer care	1.000	.008	1	198	.929

Table 6
Box's M Test Results

Box's M		53.684
F	**Approx.**	2.474
	df1	21
	df2	144192.248
	Sig.	.000

Table 7
Canonical Discriminant Function and Canonical Correlation

Eigenvalues				
Function	**Eigenvalue**	**% of Variance**	**Cumulative%**	**Canonical Correlation**
1	0.928	100.0	100.0	.694

Table 8
Wilks' Lambda and Chi-square Test

Test of Function(s)	Wilks' Lambda	Chi-square	df	Sig.
1	0.519	63.70	6	0.000

Table 9
Standardized Canonical Discriminant Function Coefficients

Independent Variables	Function
	1
Coverage	0.373
Quality of service	1.016
Facilities	0.152
Prices	0.129
Promotions	-0.045
Customer care	0.055

Table 10
Unstandardized Canonical Discriminant Function Coefficients

Independent Variables	Function	Symbol
	Coefficient value	
Coverage	0.034	β_1
Quality of service	-0.097	β_2
Facilities	-0.380	β_3
Prices	-0.279	β_4
Promotions	-0.078	β_5
Customer care	0.119	β_6
Constant (α)	-8.538	(α)

Table 11
Model Fitting information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	410.391			
Final	203.920	206.471	7	.000

Table 12
Goodness-of-Fit

	Chi-Square	df	Sig.
Pearson	218.786	321	1.000
Deviance	203.920	321	1.000

Table 13
Pseudo R-Square

Cox and Snell	.644
Nagelkerke	.739
McFadden	.503

Table 14
Parameter Estimation

		Estimate	Std. Error	Wald	df	Sig.	95% Confidence Interval	
							Lower Bound	Upper Bound
Threshold	[Satisfaction = 1]	12.340	1.815	46.224	1	.000	8.783	15.898
	[Satisfaction = 2]	18.104	2.296	62.162	1	.000	13.604	22.605
Location	Coverage	-.370	.270	1.878	1	.171	-.900	.159
	Quality of service	.535	.336	2.535	1	.111	-.124	1.194
	Facilities	5.230	.650	64.676	1	.000	3.956	6.505
	Prices	-.333	.281	1.406	1	.236	-.883	.217
	Promotions	-.228	.350	.425	1	.515	-.913	.457
	Customer care	-.154	.326	.223	1	.637	-.792	.485

Table 15
Test of Parallel Lines

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	203.920			
General	164.793	39.127	7	.000

Table 16
Distribution customers According to satisfaction

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Low	36	18.0	18.0	18.0
	Medium	97	48.5	48.5	66.5
	High	67	33.5	33.5	100.0
	Total	200	100.0	100.0	

Table 17
Case Processing Summary

		N	Percent
Sample	Training	135	67.5%
	Testing	65	32.5%
Valid		200	100.0%
Excluded		0	
Total		200	

Table 18
Classification Matrix

Sample	Observed	Predicted			
		Low	Medium	High	Percent Correct
Training	Low	15	8	0	65.2%
	Medium	7	49	5	80.3%
	High	0	14	37	72.5%
	Overall Percent	16.3%	52.6%	31.1%	74.8%
Testing	Low	10	3	0	76.9%
	Medium	2	27	7	75.0%
	High	0	1	15	93.8%
	Overall Percent	18.5%	47.7%	33.8%	80.0%

Table 19
Importance of Independent Variables

	Importance	Normalized Importance
Coverage	.107	18.9%
Quality of service	.098	17.2%
Facilities	.567	100.0%
Prices	.071	12.6%
Promotions	.088	15.6%
Customer care	.069	12.1%

Table 20
Parameter Estimates (Synaptic weights) of Neural Network Model

Predictor		Predicted					
		Hidden Layer 1			Output Layer		
		H(1:1)	H(1:2)	H(1:3)	[Satisfaction =1]	[Satisfaction =2]	[Satisfaction =3]
Input Layer	(Bias)	2.155	-.523	-.760			
	Coverage	-.484	.062	-1.155			
	Quality of service	-.234	.113	.763			
	Facilities	-3.574	1.721	1.451			
	Prices	.260	-.164	-.310			
	Promotions	.334	.420	1.574			
	Customer care	.175	.574	1.191			
Hidden Layer 1	(Bias)				.527	.588	-.539
	H(1:1)				3.816	.725	-4.250
	H(1:2)				.806	.325	-.156
	H(1:3)				-2.357	1.527	1.389

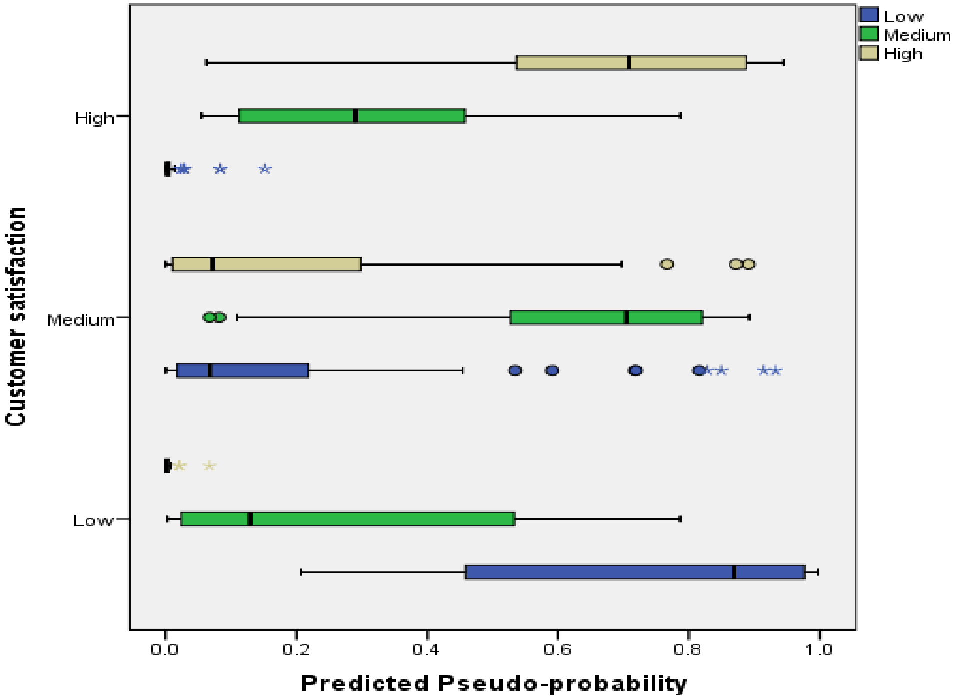


Figure 1:

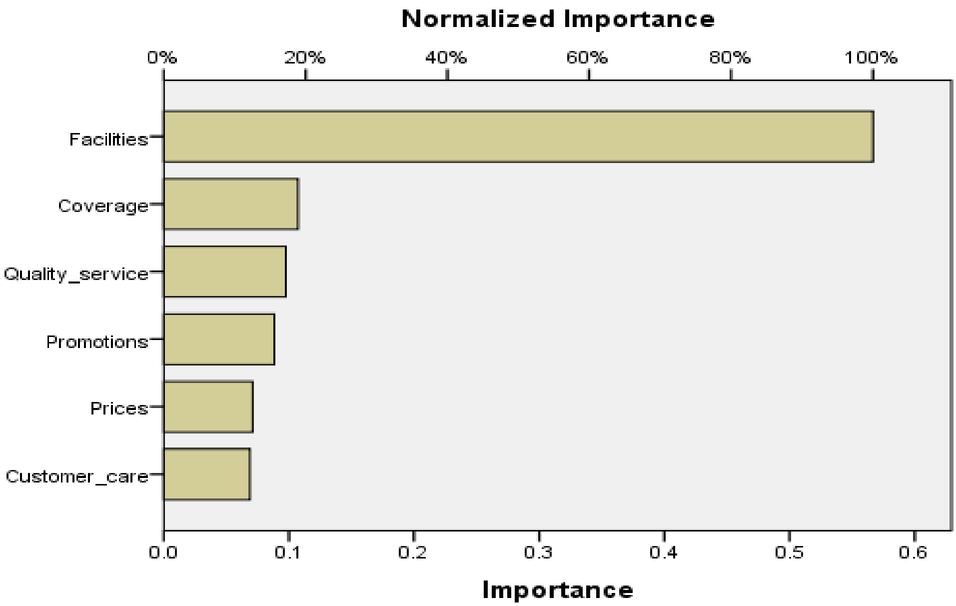


Figure 2

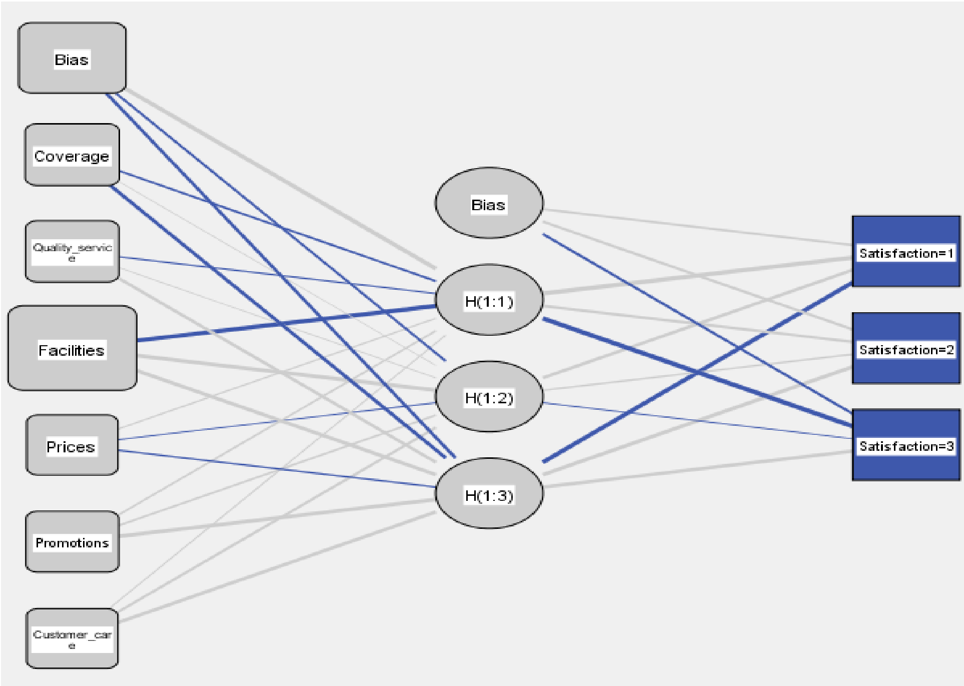

Bias
Coverage
Quality_service
Facilities
Prices
Promotions
Customer_care
Bias
H(1:1)
H(1:2)
H(1:3)
Satisfaction=1
Satisfaction=2
Satisfaction=3

Figure 3

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 331-344

IDENTIFICATION OF GOOD TEACHING CHARACTERISTICS (FACTORS)

Madiha Parveen[§] and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: 7868madihaparveen@gmail.com

ABSTRACT

A good society is based on the quality of their citizen's education. Teaching and learning is a process that provides formative effect on the individual mind or on the physical abilities. Here several components that administer the teaching and learning process including teachers, learners, environment and content. In the teaching process a teacher is the most important factor have a role as content provider, assessment learning and classroom management etc. The quality of teaching directly affects the achievement of learners. To obtain this, teacher should know the good teaching characteristics. This study examines to identify good teaching characteristics (factors) and also to know about effect of these good teaching characteristics on student's academic performance (CGPA). The total amount of sample units consisting of 150 students were selected using sample proportion of stratified random sampling. Neural network multilayer perception technique used for analysis. The results indicate that teaching characteristics have significant effect on student's academic performance (CGPA) and also have impact on teaching process.

KEYWORDS

Student academic performance (CGPA), Good teaching characteristics, Neural network multilayer perception.

1. INTRODUCTION

Schools, colleges and universities have no worth without students. Students are most essential asset for any educational institute. The students' performance (academic achievement) plays an important role in producing the best quality graduates who will become great leader and manpower for the country.

Students' performance at the level of secondary has the strong impact on other levels of higher and tertiary education. The teacher is the number two factor that influences students' academic performance. Teachers are facilitators of the learning process. Teachers should possess all qualities that a student or a society considers for a best teacher. And teachers should know the characteristics of good teaching. So teachers be concerned with different learning styles of learners and improve the quality of teaching by perceive their own strength and weaknesses.

2. FACTORS AFFECTING GOOD TEACHING CHARACTERISTICS AND STUDENT'S PERFORMANCE

2.1 Teaching Method

Teaching technique is an important factor affecting good teaching characteristics. Teaching techniques is a well-defined procedure used to achieve a particular task and teacher's personal style.

2.2 Interaction with Students

Teacher-interaction is an important factor of teaching process. The teacher-student relationship has great impact on student's self-esteem and to increase their skills.

2.3 Personality

Personality traits are important, helps the teachers and students to succeed. Teacher's personality traits pay important role to manage classroom and in learning of students. Positive attitude, high confidence, creativity, punctuality and maturity are the personality traits of good teacher's.

2.4 Teaching Practice, Beliefs and Attitudes

Teachers' beliefs, practices and attitudes are important for understanding and improving educational processes. They are closely linked to teachers' strategies for coping with challenges in their daily professional life and to their general well-being, and they shape students' learning environment and influence student motivation and achievement.

2.5 Teaching Experience

Experience gained over time increase the skills of knowledge and efficiency of workers. When students taught by more experienced teachers attain higher level of achievements. Teachers those have no experience are not much effective than those that have more than 20 years' experience.

3. OBJECTIVES

The Objective of this research paper is to identify good teaching characteristics (factors) that effect the performance of students and which factors that affect the quality of teaching.

4. LITERATURE REVIEW

Haider and Jalal in March (2018) studied. "Good Teacher and Teaching Through the lens of Students" in New Delhi, India. The purpose of this study was to know about the qualities of a good teacher and teaching. The sample consist of 30 students of class 11th were selected randomly from two schools of New Delhi, India. Responses of students were gathered through open ended questionnaire. The results reveal that nine traits of a good teacher find. And these traits have significant effect on teacher's and student's academic performance.

The study was conducted in Department of Environmental Science at Chinghai University of technology Zimbabwe in (2017) to investigate student's evaluation on teaching performance of lecturers in an academic department. The aim of this to investigating perception of students with respect to their lecturer's practice. A quantitative approach used and perspectives of students were measured by weighted mean and standard deviation. A questionnaire of student’s evaluation was prepared. The results of lecturer’s performance rate to be satisfied but they need to improve lecture pacing and to enhance students learning so improve the environment from outside the class.

Jani et al. in April (2015) studied the predictors of Lecturers "Teaching Effectiveness for public and private universities in Malaysia. “The objective of this study is to distinguished predictors of lecturers' teaching effectiveness for private and public universities of Malaysia’s sample of 316 lecturers from two public and two private universities in Perak. This research uses a self-administered questionnaire. A stepwise regression analysis used to identify which component of emotional intelligence skills have affected on teaching effectiveness. This study concludes that the personal leadership and self-management skills have significant impact on private universities teacher's effectiveness rather than public universities.

Musili conducted the related study in (2015) “Influence of teacher related factors on students’ performance in Kenya certificate of secondary education in public secondary schools in Kibwezi sub-county, Kenya. The objectives of this study was to identify the influence of job satisfaction of teacher’s on students’ performance, to determine the influence of teacher’s motivation on performance of students, in KCSE to assess the impact of teacher professional qualification on students’ academic performance and the last objective was to establish the impact of teacher experience on student’s academic performance. A sample consists of 18 principles, 90 teachers and 180 students were selected. And the results reveal that all factors including job satisfaction, teacher’s motivation, teacher professional qualification and teacher experience have great impact on student’s academic performance.

Thomas conducted the study in February (2014) to examine the effect of qualification of the teachers on student’s performance in mathematics. The main purpose of this study is that there is any difference between students’ academic performance taught by professionals and non-professionals in mathematics holding Nigeria Certificate Education (NCE). From sixteen secondary schools ten schools were purposively selected from these ten schools a sample of three hundred students were selected and the data were analyzed using Mean, Standard deviation and t-statistic. The results showed that professional qualification is the main variable affecting students’ performance in mathematics.

Rockstroh conducted the related study in (2013) “Effect of Teacher’s characteristics on students’ performance”. The main purpose of this study is to answer whether teacher's characteristics have effect on student’s achievements and whether there are differences in student achievements regarding school location. This study used data from the Ohio Department of Education and from Census Bureau. Data collected at school level. The study developed two models. So this study was conducted at school level they cannot make assumptions about the student teacher relationship.

Luntungan conducted the related study in (2012) 'Effects of teaching Methods and Students attitude on academic performance’. The main purpose this study was to investigate the effects of teaching methods in business instruction and attitude of students toward class and student on academic performance. A sample of 135 college students was selected from an Indonesian university. The methodology adopted in this study, Two-way ANCOVA and t-test. The results conclude that teaching methods affect student’s academic performance and attitude of students toward class.

Akiri in (2013) conducted this study “Effects of Teachers effectiveness on Students academic performance". The objective of this study is to investigate the effects of teacher’s effectiveness on student’s academic performance and also to find the effect of teacher's experience and qualification. The target population was all teachers and students in public secondary school in Delta State of Nigeria. Using stratified random sampling technique a sample of 350 including teachers and students was selected in public secondary schools. Three hypotheses used to determine the influence of teaching effectiveness on students’ academic performance. To test this hypothesis t-Test, Pearson Product Moment Correlation and One-Way ANOVA were used. Results showed that teacher’s effectiveness has significant effect on student’s academic performance and may be responsible for the low performance of students.

Abudu and Gbadamosi (2014) “Relationship between teacher's attitude and student's academic performance. The main purpose of this study is to examine the relationship among attitude of teachers and performance of students in secondary school chemistry. The study also looked at methods to improve academic performance of students in chemistry examinations. A simple random sampling technique used to collect data from senior secondary schools in Ijebuode and Odogbolu government area. Independent sample t-test, Pearson product moment correlation and analysis of variance used to analyze collected data. The results showed that achievement of students will be better if teachers show positive attitude toward teaching of chemistry.

Ganyaupfu (2013) studied “Teacher-centered, student-centered, teacher-student interactive and performance”. The main purpose of this study was to identify whether there are significant differences between different teaching methods on academic performance of students. A sample consists of 109 students of undergraduate from college’s Department of Economic and Business sciences were selected for the study. For analysis purpose, General Linear Model based univariate ANOVA technique used. The results conclude that when students engaged to solve problems during class activities then they build better understandings the main concepts.

5. METHODOLOGY

The methodology of this study was to find good teaching factors, in this phase core factors of good teaching characteristics that effect the student’s academic performance.

5.1 Research Design

The research design for this investigation is a survey. The independent variables are factors which affecting good teaching characteristics, dependent variables are teachers and students.

5.2 Population

The population is divided into two groups (1) The Teachers: full time instructors at UOG (2) Students: who are studying at the undergraduate level at UOG. The total amount of sample consisting of 50 teachers and 150 students were selected from university of Gujrat using cluster sampling technique.

5.3 Sample Size

The total amount of sample consisting of 150 students were selected from university of Gujrat using sample proportion of stratified sampling.

5.4 Statistical Technique

Neural network model technique is used to identify the core factors that affect the good teaching characteristics.

5.5 Sampling Technique

In this study we use Simple random sampling on teachers and Proportion allocation on students.

6. RESULTS AND DISCUSSION

6.1 Descriptive Statistic

The Table 1 consists of the test of reliability; the value of Cronbach's alpha based on standardized item is 0.867 which indicates that the data is reliable for further statistical analysis.

The Table 2 shows that we have selected 54 % females and 46 % males in our sample. The female respondent's percentage higher than the male respondent percentage.

Table 3 shows that average rank of respondents on the statement that teaching method greatly influence student's academic performance and quality of teaching process is 2.7250, it means that on average respondents are agree with that statement. Average rank of the respondents on the statement that teaching practice, beliefs and attitude also effect student's academic performance (CGPA) and the average mean is 2. 6470. And then personality of the teacher is also play major role in the academic performance is 2. 6424. It means that on the average respondents are agree with that statement. Average rank of the respondents on the statement that teaching experience is also influence student's performance is 2.6387, it means on average respondents are agree on that statement. Average rank of the respondents on the statement interaction with students greatly impact student's performance is 2.6347, it means that on average respondents are agree with that statement.

6.2 Neural Network

In Table 4 showing that by using multilayer perception model we want to predict that whether the given factors "Teaching method"" Teaching Experience"" Teaching practice, Beliefs and attitude"" Personality"" interaction with students" effect the student's performance. In the following analysis four independent variables and the dependent variable CGPA with three categories "1.5-2.5" ,2.6-3.5", "3.6-4". By using Neural we conclude these results as Table shows, in case processing analysis 68.2 % data is used for training purpose and the 31.8 % data is used for testing purpose. The table shows that we have 100 % valid cases without any excluded cases from the 150.

In Table 5 the classification table showing the correct classification. The table shows that classification categories of CGPA are correctly classified. The training sample by taking the 42.6 % correct percentage of all categories. In similar way in case testing sample by taking 53.2 % correct percentage of all categories.

The Table 6 shows that Area under the curve for different values. The area under the model useful summary statistics to detect the accuracy of model. If the area under the curve is close to 0.5 then model will be less accurate and if the under the curve is close to 1 then model will be consider best. Here area under the curve is close to 1, so the model is best.

Table 7 shows the model summary of training sample and testing sample which connected to error and incorrect prediction of response variable student performance. The value of cross Entropy Error in this model is 48.026 which is low as compare to overall sample data and this shows that our estimated model is better. Training contains 57.4 % incorrect predictions and in testing incorrect prediction is 46.8 %.

Table 8 shows that variable importance contribution d the independents variable predicting the dependent categories.The first table shows the first factor containing the 100% contribution in the dependent predictions, the next varables containing 64.9 % which is for second factor and so on.

In Figure 1 the relationship of input, hidden and the output layers are determined as the grey lines are for the positive relationship and the blue lines are for the negative relationship. The thickness of line shows the flaws and strength of relationship. The dark blue dark bold lines show the strong relationship either negatively or positively, while the thin lines show the weak relationship also either positively or negatively.

The ROC curve is given in Figure 2. It gives the sensitivity and specificity. The sensitivity is the true prediction of true results. While the specificity is the number of negative cases predicted line in the graph is the cut point for classification value above the lines are correctly classified. In this figure the three lines are given blue line in 1.5-2.5 category, the green is for 2.6-3.5 category and gray is for 3.6-4. All the lines are at 1 on the upper right corner so we say our model is best fitted.

Figure 3 Shows the cumulative gain explains the percentage the whole number of cases in given category can also be achieved by using the gain chart which targets the % of total of cases. In instance, the first point on the equal is at (10%, 20%) which shows the by using pseudo probability. The baseline for these curves is the sloping lines, the model provides gains, if curve falls above the base line. Here all above the line.

Figure 4 Shows the lift chart that is derived from the cumulative gain which shows the value on the y-axis correspond to ratio of the cumulative chart on each curve to the baseline. In instance, the first point on the equal is at (10%, 20%) which shows the by using pseudo probability.

Figure 5 shows the graphical representation of all variables. The most significant variable is “Teaching method” which is playing significant role in improving academic performance (CGPA). Secondly, teaching practice, beliefs and attitude towards student’s performance. Thirdly, personality playing significant role in the student’s performance (CGPA).

Table 1
Reliability Statistics

Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	N of Items
.869	.867	30

Table 2
Percentage of Nominal Scale Variables

	Percentages	Frequencies
Male	**46%**	**69**
Female	**54%**	**81**
Total	**100**	**150**

Table 3
Percentage, Mean and Standard deviation for All Ordinal Variables

Factors/Variables	SD%	D%	A%	S.A%	Mean	S.D
Teaching Method:						
Teacher preparation for the lecture before delivering effects student's performance.	2%	12.7	66.7	18.7	3.02	.629
Teacher prepares assignments individually for each student.	3.3	36.7	34	26	2.83	.857
Development or training plans related to teaching does not improve student's performance.	6	28	46	19.3	2.79	.824
Revise versions of their work and lectures improve student's performance.	2	31.3	51.3	15.3	2.80	.714
Providing the useful examples that can apply in real life improve student's performance.	7.3	22	54	16.7	2.80	.803
Having a good management of teaching time in each of topics.	18	26	40.7	15.3	2.53	.960
Student's result is based on the teaching method.	16.7	35.3	38.7	9.3	2.41	.875
Student provide good results depend on teaching method.	12.7	27.3	44.7	15.3	2.63	.894
Overall	8.5	27.5	47	17	2.7250	.47017
Teaching Experience:						
Highly experienced teacher has positive effect on student's performance.	2.7	2	55.3	40	3.33	.650
Experience of using technology in the class has effect on student's performance.	18.7	35.3	19.3	26.7	2.54	1.078
Experience of teacher help to easily convey lecture of student's mantel level.	7.3	69.3	11.3	12	2.28	.770
Experience of teacher helps the student to understand lecture deeply.	10	44	40	6	2.42	.753
Experience of teacher have complete grip on main subject to help students.	14.7	22.7	50	12.7	2.61	.889
Overall	10.7	34.7	35.2	19.5	2.6387	.60056

Factors/Variables	SD%	D%	A%	S.A%	Mean	S.D
Teaching Practice, Beliefs, Attitude:						
Good teachers demonstrate the correct way to solve a problem.	2.7	2.7	59.3	35.3	3.27	0.644
Teacher attitude does not affect the student's performance.	17.3	30	33.3	19.3	2.55	0.994
Teacher friendly behavior with student's effect the student performance.	6	55.3	30.7	8	2.41	0.724
Students learn best by finding solutions to problems by their own way.	7.3	40.7	44	8	2.53	0.748
Having of injustice improve the student's performance.	16	26.7	42.7	14	2.55	0.926
Teacher beliefs about their students improve the student's performance.	13.3	42.7	31.3	12.7	2.43	0.878
Teacher positive attitude give opportunities to students to improve their skills.	4.7	32	47.3	16	2.75	0.779
Overall	9.6	32.9	41.3	16.2	2.6470	.44363
Interaction with Students:						
Teachers supported and tried to help students when it was needed.	2.7	2.7	55.3	39.3	3.31	0.657
Having a friendly nature does not affect the student's performance.	18.7	34	21.3	26	2.55	1.072
Talking with students helps them about lecture.	8	66.7	14	11.3	2.29	0.771
Having of good interaction of teacher and student in classroom effect the student's performance.	10.7	44	38.7	6.7	2.41	0.770
Teacher rude behavior does not affect the student's performance.	13.3	23.3	50	13.3	2.63	0.878
Overall	10.7	34.1	35.9	19.3	2.6347	.60543
Personality:						
Providing of honor and respect the rights of students affects the student's performance.	3.3	5.3	63.3	28	3.16	0.666
Having of interest, willingness and responsibility in teaching.	15.3	30.7	31.3	22.7	2.61	1.002
Good interaction of teacher and student in classroom effects the student's performance.	6.7	52	28	13.3	2.48	0.809
Teacher's behavior effects the teaching method.	6	46.7	40.7	6.7	2.48	0.712
Having good sense of humor affects the student's activities in class.	18	28.7	39.3	14	2.49	0.947
Overall	8.5	27.5	47	17	2.6424	.56444

Neural Network

Table 4
Case Processing Summary

Sample	N	Percent
Training	101	68.2%
Testing	47	31.8%
Valid	148	100.0%
Excluded	2	
Total	150	

Table 5
Classification

Sample	Observed	Predicted			
		1.5-2.5	2.6-3.5	3.6-4	Percent Correct
Training	1.5-2.5	1	4	12	5.9%
	2.6-3.5	1	2	32	5.7%
	3.6-4	2	7	40	81.6%
	Overall Percent	4.0%	12.9%	83.2%	42.6%
Testing	1.5-2.5	0	1	8	0.0%
	2.6-3.5	0	4	11	26.7%
	3.6-4	1	1	21	91.3%
	Overall Percent	2.1%	12.8%	85.1%	53.2%
Dependent Variable: CGPA					

Table 6: Area under the Curve

		Area
CGPA	1.5-2.5	.605
	2.6-3.5	.513
	3.6-4	.499

Table 7
Model Summary

Training	Cross Entropy Error	105.179
	Percent Incorrect Predictions	57.4%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error
	Training Time	0:00:00.13
Testing	Cross Entropy Error	48.026
	Percent Incorrect Predictions	46.8%

Table 8
Independent Variable Importance

	Importance	Normalized Importance
Teaching Method	.304	100%
Teaching Practice, attitude, Beliefs	.197	64.9%
Personality	.169	55.6%
Teaching Experience	.166	54.7%
Interaction with students	.088	28.9%
Gender	.076	24.9%

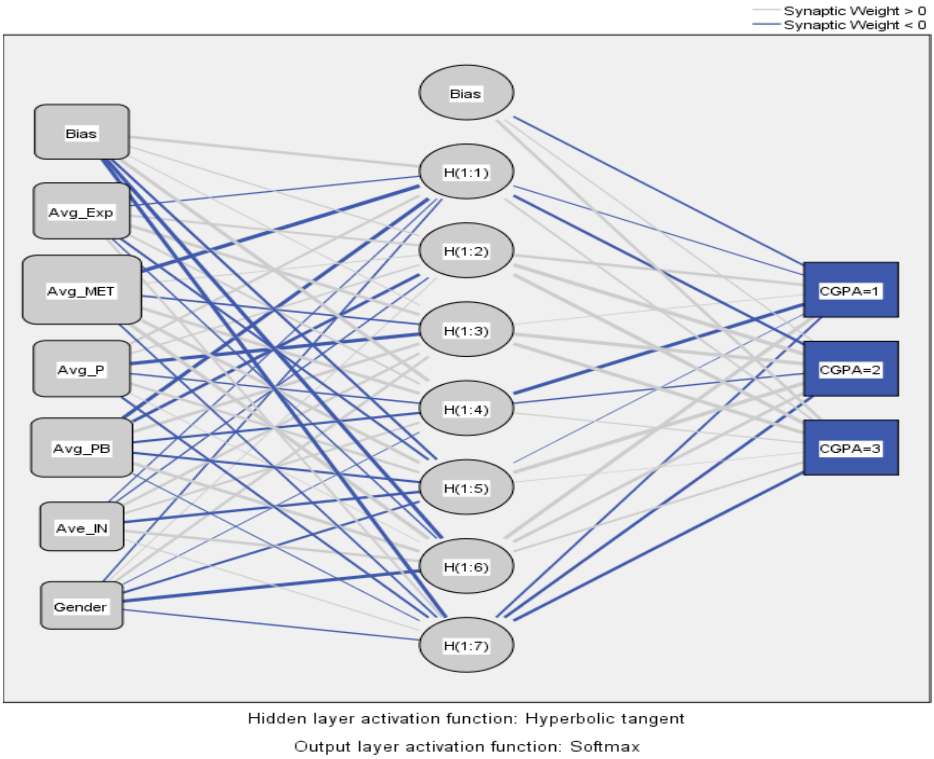


Figure 1: Network Diagram

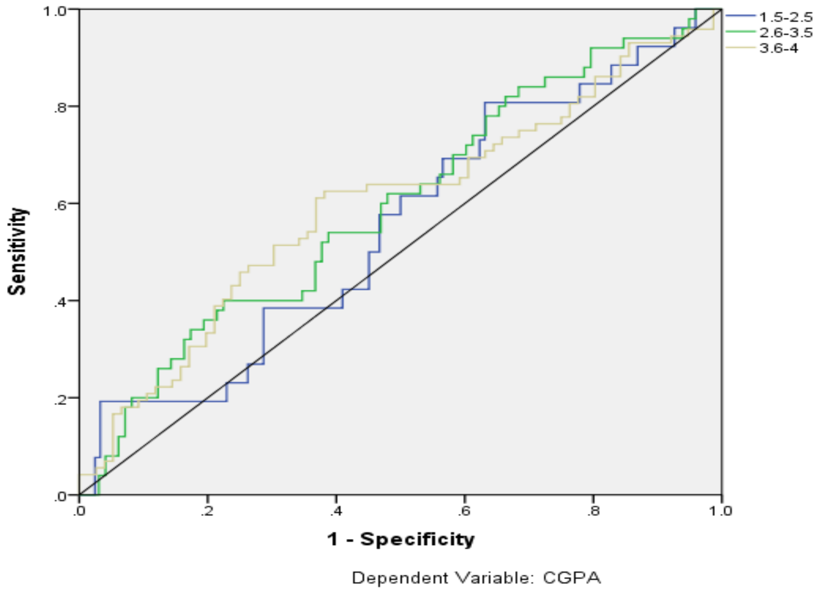


Figure 2: ROC Curve

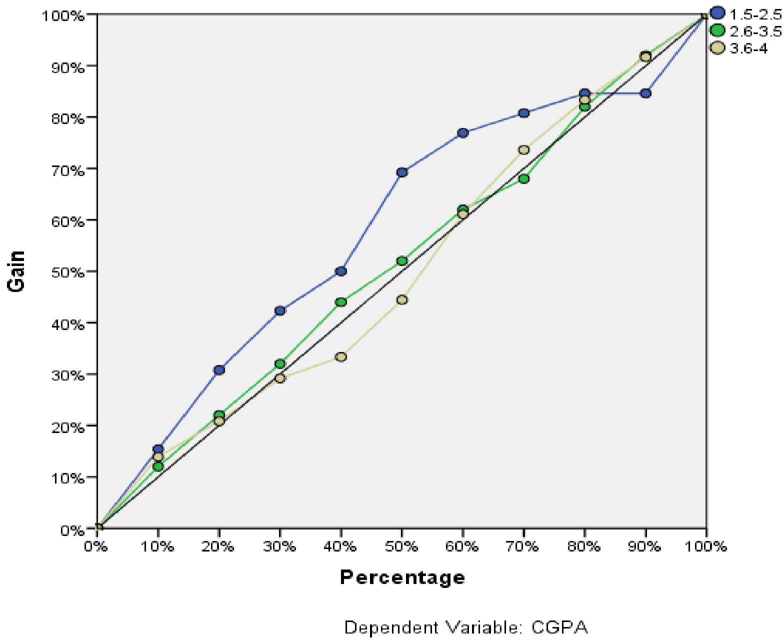


Figure 3

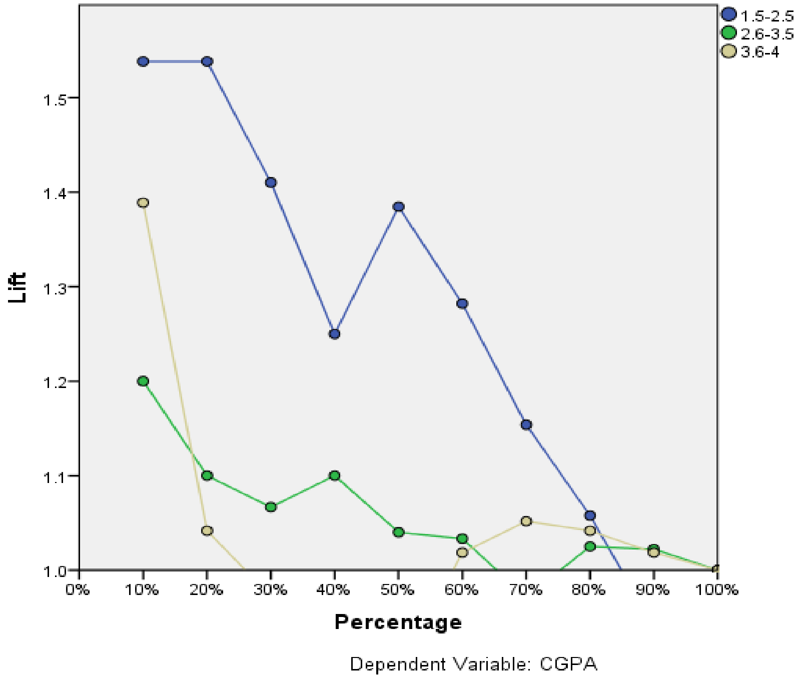


Figure 4

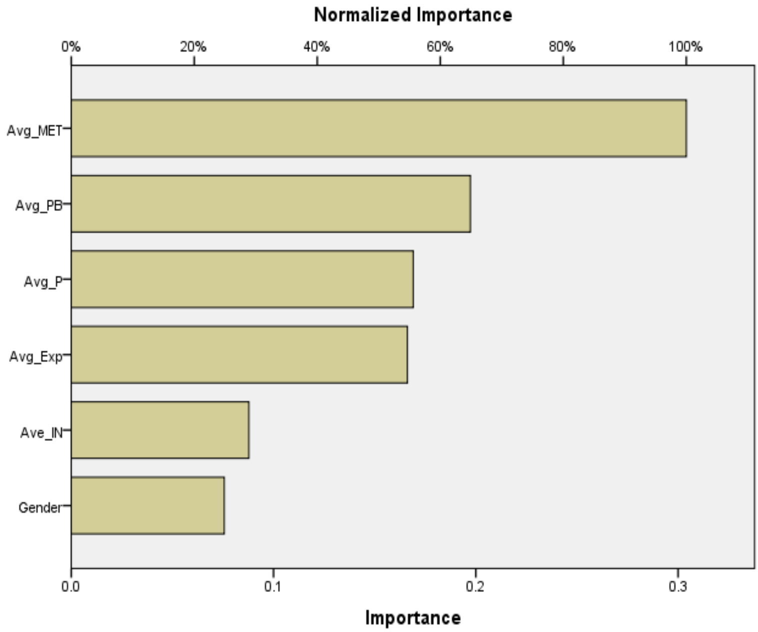


Figure 5

6.3 Conclusion

This study examines the overall satisfaction of respondents and those factors are also studied which have effect on the satisfaction of the respondent. Satisfaction of the students is an important task because the students have need and right to participate in education programs and receive satisfactory services. For this purpose of data collection, a questionnaire designed and the data collection from the University of Gujrat. A Sample of 150 students is selected for the analysis and using proportional sampling. For the analysis of data, we use Neural Network Multilayer Perception. Data can be analyzed using computer software SPSS 21. In this study 54% female and 46% are male. The reliability of the data is 0.867.

Result of this study emphasize the importance of evaluating factor such as teaching method, teaching experience, teaching practice, beliefs and attitude, interaction with students and personality. The percentage of females is higher than male who prefer teaching method which effect student's performance (CGPA).

The neural network has been used to predict the probabilities. The model demonstrates those teaching method, personality and teaching practice, beliefs and attitude have highly significant role in model building. In this study, teaching method, teaching practice, beliefs and personality are significantly associated with the student's performance.

6.4 Limitation

Due to limited time and resource we conduct this study only from the University of Gujrat students are considered as respondents.

REFERENCES

1. Haider, A. and Jalal, S. (2018). Good Teacher and Teaching through the Lens of students. *International Journal of Research*, 5(7), 1395-1409.
2. Phiakoksong, S., Niwattanakul, S. and Angskun, T. (2013). An Application of Structural Equation Modeling for Developing Good Teaching Characteristics Otology. *Information in Education*, 12(2), 253-272.
3. Akram, M.J. (2010). Factors affecting the performance of teachers at higher secondary level in Punjab, Doctoral dissertation, Pir Mehar Ali Shah Arid Agriculture University Rawalpindi, Pakistan.
4. Jani, S.H.M., Shahid, S.A.M., Thomas, M. and Francis, P. (2015). The Predictors of Lecturers' Teaching Effectiveness for Public and Private Universities in Malaysia. *International Journal of Social Science and Humanity*, 5(4), 384-388.
5. Musili, A.M. (2015). *Influence of teacher related factors on students' performance in Kenya certificate of secondary education in public secondary schools in Kibwezi sub-county Kenya*, Master Dissertation, South Eastern Kenya University.
6. Luntungan, R. (2012). Effects of teaching methods and students' attitude on academic performance. In *International Forum*, 15(2), 42-56.

7. Ganyaupfu, E.M. (2013). Teaching methods and students' academic performance. *International Journal of Humanities and Social Science Invention*, 2(9), 29-35.
8. Abudu, K.A. and Gbadamosi, M.R. (2014). Relationship between teacher's attitude and student's academic achievement in senior secondary school chemistry. A case study of Ijebu-Ode and Odogbolu Local Government Area of Ogun state. *Wudpecker Journal of Educational Research*, 3(3), 035-043.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 345-348

FEMALE EDUCATION AND TRADITIONAL ATTITUDE OF THEIR PARENTS: EVIDENCE FROM UNIVERSITY OF GUJRAT

Zainab Yaqub[§] and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: yzainab978@gmail.com

ABSTRACT

For the growth and buildup of any successful society, female education is considered as an unavoidable aspect. The conservative attitude of the parents is the main obstacle for the female education especially in rural areas.

In result, the females did not contribute in progress of Pakistan. Comparatively, the females of urban areas are more progressive. The purpose of this research is to find out side effects of different thoughts of parents on the education of the females in rural and urban areas.

The data was collected from the females who belong to rural and urban areas and currently studied in the University of Gujrat. A sample of size 200 females was collected using proportional allocation. Descriptive statistics and neural network accomplished.

Results of study reveal that the females from rural areas suffered the more barriers in education as compare to the females of urban areas. The negative behavior of parents about female education had destroyed the abilities of females. Further, the results show that the urban people liberalized as compare to the rural people.

KEYWORDS

Parental attitude, Female education, urban residence, rural residence.

1. INTRODUCTION

Education is necessary for all the human beings across of the masculine, tribe or country. For any state, to overlook the rights of womanish is not a good policy. Females are skilled and competent part in disciplined culture.

In the labor parties, the educated lady helps their illiterate fellows; so that implement the majority rules in any region. Knowledge gives the ideas to tackle the big restriction for the women Education.

The Western women are routinely worked in free trade areas without any restriction and contribute in progress of their countries. So this review examine that getting an education for "Female Education and Traditional Attitude of Their Parents: Evidence from University of Gujrat" a female proves progressive for future life. Attitude of Parents plays an important role in female education.

In rural areas mostly parents think that education is not much necessary for the females. The urban residences were more liberalized as compare to the rural residence.

Parental Attitude

Attitude of Parents plays an important role about female education. In backward areas mostly parents think that education is not much necessary for the females. Parents have a fear of sexual urge during the travel or in an educational institution.

Parents think that the life of female is restricted after marriage life. In rural areas, resources are not enough used for female education as compare to the male education. Females are played important role only in a house being a housewife.

Females could not contribute in progress of the country. So the parental attitude about female has a strong on a female education.

Female Education

Female education is the most serious issue in education field. Males play an important role for the progress of their country. So this was no mean that the female has no right to getting an education. From one female make a whole family in a future.

As a result, if mother was educated then she will also prefer the education of the children. Then the children become successful native of a country. In this way a female contributes in progress of country. Along this, an educated female financially support the family.

Educated female life did not spend her life to depend on others. The standards of selection of educational institutions were also very important.

Urban Residence

A large population lived in urban areas. Most of the people were well educated. People of different culture lived in urban areas. People have different thoughts on different aspects of life. People have positive thought about the education of the female.

They thought that the life of a female was not only limited in a house. The urban residences were more liberalized as compare to the rural residence.

They give freedom to their children. The decision about the females ware always finalized on the basis of the liking or disliking. In urban areas, males and females had equal rights.

“Female Education and Traditional Attitude of Their Parents: Evidence from University of Gujrat”

The selection of the technical subjects for females was specified as the subjects for the males pre specified. The marriage decision was also finalized according the female’s opinion.

So the female of urban residence was as successful as male. So that the gender equality plays a positive role to polish the females abilities and capabilities.

Rural Residence

A small population lived in rural areas. The occupation of the most people is agriculture. People of same culture lived in rural areas. People have same thought on different aspect of life. People have negative thought about the education of the female.

They thought that the life of a female is limited in a house. The decision about the females is always finalized by others (parents or husband).

But the males are fully independent about their decisions. In rural areas, the selection of educational institutions for the females is always inferior as compare to the male. Decision was also finalized about the parents' opinion.

The inequality is also hold in selection of subjects between males and females. Males selected more technical subjects as compare to the females. The rural residence did not tolerate that any person discussed about their mother, sister or wife.

Rural residence considers a male person as a helping tool for the family. So that the gender discrimination plays a negative role to polish the females abilities and capabilities. Pardah is the biggest issue for the females of the rural areas.

Lot of the barriers was removed but still the barriers are present for the females of rural areas.

2. STATEMENT OF THE PROBLEM

The problem under taken to study the education of females and traditional attitude of their parents towards the female education.

3. OBJECTIVES

There are following objectives are given below

1. To observe the parental attitudes towards the females education.
2. To expose the effect of addicted perspectives on female attitude.
3. To examine the thoughts of students in University Of Gujrat about the female education.

4. COMMENTS AND CONCLUSION

Education for females is as necessary as for males. Think beyond the barriers in education. Positive attitude of parents makes the path for the education of females.

Explore the solutions of the barriers. If a female is educated then whole race will be educated.

The requirements of today that the female education avails full attention specially the females from rural areas.

In today's life, every country has a great race to achieve the progress then the education for the females is compulsory.

If every female is educated then Pakistan will become more powerful in coming years.

REFERENCES

1. Onoyase, A. (2018). Attitude of Parents toward Female-Child Secondary Education in Sokoto State, Nigeria: Implications for Counselling. *Journal of Educational and Social Research*, 8(2), 21-27.
2. Chingtham, T. and Guite, T. (2017). Parental attitude towards girl’s education. *IOSR Journal of Research & Method in Education IOSR-JRME*, *7*(4), 1-6.
3. Breakwell, G.M. and Beardsell, S. (2016). Gender, parental and peer influences upon science attitudes and activities. *Public Understanding of Science*. https://doi.org/10.1088/0963-6625/1/2/003
4. Mor, K. and Sethia, S. (2015). Parent’s Attitude towards Girls Education in Haryana. *American International Journal of Research in Humanities, Arts and Social Sciences. ISSN (Online)*, 2328-3696.
5. Sarkar, R.K., Reza, M.M. and Hossain, M.E. (2014). Socio cultural barriers of girls’ educational attainment experiences from rural Bangladesh’. *Cultural Anthropology*, 10(2), 349-358.
6. Iqbal, S., Mohyuddin, A., Ali, Q. and Saeed, M. (2013). Female education and traditional attitude of parents in rural areas of Hafizabad-Pakistan. *Middle-East Journal of Scientific Research*, 18(1), 59-63.
7. Mohyuddin, A. and Ambreen, M. (2012). Changing Attitude towards Female Education (A Case Study of Village Zandra in Balochistan). *International Journal of Scientific and Engineering Research*, 3(12), 1-14.
8. Shahzad, S., Ali, R., Qadeer, M.Z. and Khan, M.S. (2011). Community Attitude towards Female Education. *International Journal of Academic Research*, 3(1), 370-373.
9. Raza, A. and Sohaib Murad, H. (2010). Gender gap in Pakistan: a socio-demographic analysis. *International Journal of Social Economics*, 37(7), 541-557.
10. Rogers, M.A., Theule, J., Ryan, B.A., Adams, G.R. and Keating, L. (2009). Parental involvement and children's school achievement: Evidence for mediating processes. *Canadian Journal of School Psychology*, 24(1), 34-57.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 349-362

IMPACT OF ORGANIZATIONAL CULTURE ON EMPLOYEES' JOB-SATISFACTION

Aiman Ijaz and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: 16131513-008@uog.edu.pk
zaheerabbas@uog.edu.pk

ABSTRACT

Organizations mainly demands to increase their productivity and efficacy by paying attention to the employees' needs and increasing their job satisfaction level. Faculty members are most significant factor for any education institution that establish the achievements of university. This study is design to examine that how much organizational culture of University of Gujrat (UOG) influence their employee's job satisfactions (faculty). A sample consisting of 200 respondents were selected from UOG faculty members thorough Stratified sampling with proportional allocation method. Descriptive statistics, correlation analysis, and structure equation modelling were used for analysis purpose. Results reveals that organization culture has significant effect on job satisfaction of faculty members. The findings also illustrate that gender has no effect on the relationship of organizational culture and job satisfaction.

KEYWORDS

Organizational Culture, Faculty Members, Job Satisfaction, Structure Equation Modelling.

1. INTRODUCTION

Employees are the fundamental part of an organization. The success and growth of an organization depend on effectiveness employees. Organizational culture is the collection of people behavior, Organization norms, values, system and symbols and also include the beliefs and attitude. Organization members shared belief, values and norms about organization culture. Motivation is the complex force to keeping work in an organization. Employees effort are effect on the job performance as modified by abilities, roles and perceptions.

1.1 Organizational Culture

Organizational culture defined the set of people behavior, organization norms, values, system and symbols and also include the beliefs and attitude. Organizational Culture influence the groups interact with colleagues. When people join the institution then takes the belief, norms and values of the organizational culture. Organizational culture is developed for the long period of time and culture are different from one organization to another. In organizational culture, language is most important factor that communicate member with each other. There are three dimension includes in organization culture. (1) Environment (2) Coworkers (3) Supervisor.

1.2 Job Satisfaction

Job satisfaction is defined as the extent to which an employee feels self-motivated, content & satisfied with his/her job. Job satisfaction happens when an employee feels he or she is having job stability, career growth and a comfortable work life balance. People have positive thinking and feeling whose people satisfices from organization job. Job satisfaction are effect on the employee's life. Job satisfaction term are not well defined but there are used some theories about job satisfaction. There are three dimension includes in job satisfaction. (1) Work Itself (2) Pay/fringe benefits (3) Promotion.

1.3 Job Satisfaction and Organizational Culture

Organization culture are play very important role in increasing job satisfaction. It is also known that organizational culture affected on the job satisfaction. Employees Job satisfaction refers to how content the employee is with the place they work in and the work they do (Kreitner and Kinicki, 2013). Furthermore, Moradi et al. (2013) observed that Employees Job satisfaction is an assortment of job features, personal characteristics, environment and feelings, which require components such as co-workers, supervision and the structure of the organization may change over time. As a consequence, an employee views his/her job favorable or unfavorable in the organization where whether job conditions meet his/her expectations or not (Werther and Davis, 1999). Likewise, the employee often develops multilayered combination of positive and negative feelings, values and perceptions about various features of job, such as the nature of work, supervision style, co-workers' relationships, conditions of employment, working conditions, workload, promotion, training, job security and career opportunities (Oshagbemi, 2000; Judge and Church, 2000; Garcia, 2005; Guest, 2017; Luthans, 2005; Khan, 2006; Chamberlain et al., 2016).

2. LITERATURE REVIEW

Wilkins et al., (2018), this study aims to investigate the effect of organizational identification on employee attitudes and behaviors. The objective of the research to identify the influence of organization on employee attitudes and behaviors. The total of 795 questionnaires were collected of academic and non-academic staff. Structure equation modeling was used for testing. In this study to construct the model, one way anaylsis of variance reveal that there is no significant difference between the mean score. The result showed that data is good fitted.

Sabir et al. (2018), each organization have different organization culture. This study aims to investigate the relationship between the four types organization culture based on contend value framework and job satisfaction among different universities. A sample consist 200 respondents by using the simple random sampling. Determine the objective of the study through regression anaylsis. The findings reveal that two variable have significant impact on job satisfaction. Independent variable is higher impact on dependent variable is by the market culture.

Körner et al., (2015), this study is to purpose to explore relationship of organization culture and job satisfaction. Organization Culture is explaining as the shared norms, beliefs and assumption within organization. A sample was collected 272 employees. Structural equation modeling is used in this study. The result showed that predict 35% of job satisfaction include organization culture and teamwork. There are highly correlated organization culture and job satisfaction.

Habib et al., (2014), the aims of this research to examine the impact of organization culture on job satisfaction, employee's commitment and turn over intention. Organization run through different organization values. In this study used the primary data. A sample was collected 235 employees of different organization. Find out the result of study through Correlation analysis. Result showed that organization nature is significant effect on the job satisfaction. Organization culture are dominant element that are highly correlated job satisfaction.

Selamat et al., (2013), this article investigates the impact of organizational climate on teacher job performance. The objective of this study to examine relationship between organizational climate and teacher performance. A sample consist 37 secondary school teacher district of Klang. Data were collected by using survey method. This study showed that organizational climate significant factor on teacher's job performance. This study was found that implementation of role in organization and do not focus on paper work.

Uddin et al., (2013), this research paper examine the Impact of Organizational Culture on Employee Performance and Productivity. Organization culture is helps to make the behavior pattern of employees. Performance is the ability to specific task in specific manners. A total sample of 34 interviews were conducted. The finding of this study that organization culture has significant effect on organization performance. Organization culture is an open system approach that is direct effect on organization performance.

3. MATERIALS AND METHODS

Population

In this study, all employees (faculty) who currently performing their services in university of Gujrat are consider as target population. Size of target population in this study is 150 faculty members.

Sample Selection

The faculty members of different department are not homogeneous with respect to environment, coworkers and supervisor. We have used stratified random sampling with proportional allocation method to select a sample.

$$n_i = n\frac{N_i}{N}$$

where n and N are sample and population size and N_i are department population size.

Research Instrument

Questionnaires is used as data collection instrument. It consists of two parts First part of the questionnaire is designed to obtain information on the demographic characteristics of faculty members, like gender, age, qualification, experience. Next part designed to obtain information on some quantitative variables related to organization culture and job satisfaction. Then there are 36 items using a 5-point Likert-scale. Items are designed to assess six dimensions associated organization culture and job satisfaction. (Environment, Co-Workers, Supervision, work itself, Pay/fringe benefits, Promotion Opportunities)

Data Analysis Techniques

Confirmatory Factor Analysis

CFA is used to provide a confirmatory test of our measurement theory. A measurement theory specifies how measured variables logically and systematically represent constructs involved in a theoretical model. In confirmatory factor analysis (CFA), theory is a systematic set of casual relationships that provide the comprehensive explanation of a phenomenon. In confirmatory factor analysis (CFA), model is a specified set of dependent relationships that can be used to test the theory. In confirmatory factor analysis (CFA), is used to test structural equations. The path diagram shows the graphical representation of cause and effect relationships of the theory. In confirmatory factor analysis (CFA), endogenous variables are the resulting variables that are a causal relationship and exogenous variables are the predictor variables.

In confirmatory factor analysis (CFA), identification is used to test whether or not there are a sufficient number of equations to solve the unknown coefficient. In confirmatory factor analysis (CFA) identifications are of three types: (1) under identified, (2) exact identified, and (3) over-identified. In confirmatory factor analysis (CFA), goodness of fit is the degree to which the observed input matrix is predicted by the estimated model.

Structural Equation Modeling

Structural equation modeling (SEM) is a relatively new analytical tool, but its roots extend back to the first half of the twentieth century. During the late 1960s and early 1970s, the work of Joreskog and Sorbom led to simultaneous maximum likelihood estimation of the relationship between constructs and measured indicator variables as well as among latent constructs.

Structural equation modeling (SEM) is a family of statistical models that seek to explain the relationship among multiple variables. In doing so, it examines the structure of interrelationships expressed in a series of equations, similar to a series of multiple regression equations. These equations depict all of the relationships among constructs (the dependent and independent variables) involved in the analysis. Constructs are unobservable or latent factors represented by a multiple variables (much like variables representing a factor in factor analysis). SEM can be thought of as a unique combination of both types of techniques (interdependence, dependence) because SEM's foundation lies in two familiar multivariate techniques: factor analysis and multiple regression analysis. SEM is the only multivariate technique that allows the simultaneous estimation of multiple equations.

With large sample sizes, the χ^2 test statistic is known to always reject in any formal test of significance (Byrne, 1998). Hilton et al. (2004), more focused on the Root Mean Square Error of Approximation (RMSEA), the Goodness-of-Fit Index (GFI), the Non-

Normed Fit Index (NNFI; termed Tucker-Lewis Index or TLI in Marsh & Yeung, 1996), the Comparative Fit Index (CFI) and the Relative Fit Index (RFI, termed Relative No centrality Index or RNI in Marsh & Yeung, 1996), and the normed version of the χ^2 test statistic: $\chi^2/d.f$. For the last index, no clear-cut guidelines exist; values in the range of 2.0 to 5.0 are acceptable, with lower values indicating better fit. For RMSEA, values ≤ 0.05 indicate good fit, values ≤ 0.08 indicate reasonable fit. The indices GFI, NNFI, CFI, and RFI, all normally lie in the range 0.0 – 1.0, with higher values indicating better fit. As a benchmark for good fit, the value 0.90 is often used (Kline, 2005).

4. RESULTS AND DISCUSSION

4.1 Descriptive Statistics

Reliability of the questionnaire is a key factor for statistical results. Table 1 consists of the test of reliability; the value of Cronbach's alpha based on standardized item is 0.879 which indicates that the data is reliable for further statistical analysis.

Descriptive statistics describe the basic properties of data. All items in each factor described then mean, standard deviation and frequency percentage in the study. Table 2 Mean is the central tendency of data and measure dispersion through standard deviation. We use frequency percentage in categorical data. Descriptive statistics of all the variables are given in Table 1 to table 3. Table shows that male percentage of faculty member (59.3) and female percentage of faculty member (40.7). Percentage of Age that interval 20-50 have 86.7. The percentage of qualification that Ph.D. (28.0), MPhil (61.3) and Master (10.7). The Percentage of experience that interval 1-10 have 78.7.

Table 3 Percentages of Departments that Business Management department has highest percentage because highest Population as compare to other department and Mechanical Engineer has lowest percentage because lowest population as compare to other department.

Table 4 shows that the overall percentage of environment factor mostly respondent are response Agree 44.1% with mean 15.833 and S.D 2.9592. The overall percentage of co-workers factors that mostly respondent are response Agree 54.1%, with mean 17.2267 and S.D 2.6090. The overall percentage of supervision factors that mostly respondent are response Agree 46.1% with mean 15.933 and S.D 2.7700. The overall percentage of Work itself factors that mostly respondent are response Agree 44.2% with mean 16.24 and S.D 2.9124. The overall percentage of pay factors that mostly respondent are response disagree 45.6% with mean 14.52 and S.D 3.91. The overall percentage of promotion factors that mostly respondent are response Agree 40.3% with mean 15.37 and S.D 3.26.

Table 5 shows that results of spearman's rho correlation analysis to see relationship strength between organization culture and job satisfaction. We compute the overall scores of these factors and apply the bivariate spearman's correlation. The value of correlation coefficient between coworkers and Supervision is 0.232 with p-value 0.004, it shows significant and positive relationship between variables. Next relationship between co-workers and promotion is 0.302 with p_value 0.000. It shows that significant positive relationship. Next supervision has also significant positive relationship with pay, it has correlation value is 0.183. Next pay has negative relationship with environment. Next work itself have significant relationship with environment, p-values is 0.004. The value of correlation coefficient between Pay and Promotion is 0.363 with p-value 0.000, it shows significant and positive relationship between variables.

4.2 Confirmatory Factor Analysis

Confirmatory factor analysis is a special type of factor analysis and is the first part of a complete set of a structural equation model. We confirm all the factors which we considered such as Environment, Coworkers, Supervisor, Work Itself, Pay and Promotion by using confirmatory factor analysis. Table C-1 shows the model estimates of confirmatory factor analysis of all the factors. Table C-1 contains the model estimates of confirmatory factor analysis of Job satisfaction. p-values of all the items are significant so we reject the null hypothesis that all items are not confirm for that factor. So we conclude that all items of that factor are confirmed for that factor. Individual parameter estimates exhibited that co-workers have a parameter estimate value of 1.602, which is high as compared to other variables in the factor; it means that variable is most important for the factor. In other words, friendly atmosphere among co-workers is an essential variable for job satisfaction. Environment has a parameter estimate value of which is low as compared to other variables in the factor; it means that variable is less important for the factor. Goodness of Fit measures is used to assess the model fitness. Almost all goodness of fit measures meets the recommendation level for this factor. So, goodness of fit measures supports our estimated model.

Table C-1 also contains the model estimates of other factors; which can be interpreted in similar manner. After Confirmatory Factor Analysis our next step is to fit the structure equation model on those factors (including items) that are confirmed by Confirmatory Factor Analysis.

4.3 Structure Equation Modeling

In this paper, we use Structure Equation Modeling to develop the job satisfaction model of faculty members at University of Gujrat. We use Environment, Co-Workers, Supervision, Pay, Promotion and job satisfaction as a constructs.

A structure model involves specifying Structural relationships between latent constructs. Table D-1 contains the Parameter estimates of Structure equation model. p-values of all the parameters are significant so we reject the null hypothesis that the coefficients are zero. So we conclude that all relations are significant and positive.

Table D-2 Goodness of Fit measures of is used to assess the fitness of structure equation model. p-value of Chi-Square test is significant. So our model is fit. Recommended value of ($\chi^2/d.f$) is less than 3. In this case, the value of ($\chi^2/d.f$) is 2.84 that is less than 3. So it also supports our estimated model. In this case, GFI value is .804 the value of AGFI is .726, which supports our estimated model. In this model, the value of RMSEA is .06 that is less than .08. So RMSEA is supported to fitted model. All the important Goodness of fit measures indicates that our estimated model is best fitted.

4.4 Conclusions

In this paper, we determine the different factors that are related to job satisfaction and develop an job satisfaction model of faculty members at University of Gujrat by using

Structure equation modeling. They should do focus on environment, Pay and promotion then they can enhance their academic Performance.

For high job satisfaction, we suggest that faculty basically improve their organization culture like improve the environment, friendly atmosphere among co-workers, and supportive supervision. Environment like give comfortable working condition, provide basic facilities, Co- workers like cooperate co-workers, friendly atmosphere and informal relation among co-workers and supportive supervision. We want high job satisfaction then increase Pay and promotion opportunity. Then we enhance the job satisfaction.

REFERENCES

1. Wilkins, S., Butt, M.M. and Annabi, C.A. (2018). The influence of organization identification on employee attitudes and behaviors in multination higher education institutions. *Journal of Higher Education Policy and Management*, 40(1), 48-66.
2. Danish, R., Draz, U. and Ali, H.Y. (2015). Impact of organizational climate on job satisfaction and organizational commitment in education sector of Pakistan. *American Journal of Mobile system, Application and Service,* 1(2), 102-109.
3. Korner, M., Wirtz, M.A., Bengel, J. and Goritz, A.S. (2015). Relationship of organization culture, teamwork and job satisfaction in interprofessional teams. *BMC Health Service Research*, 15(10), 243.
4. Habib, S., Aslam, S., Hussain, A., Yasmeen, S. and Ibrahim, M. (2014). The impact of organization culture on job satisfaction, employees commitment and turn over Intention. *Advances in Economics and Business*, 2(6), 215-222.
5. Uddin, M.J., Luva, R.H. and Hossian, S.M. (2013). Impact of organization culture on employees performance and productivity .A case study of telecommunication sector in Bangladesh. *International Journal of Business and Management,* 8(2), 63.
6. Erkutlu, H. (2012). The impact of organization culture on the relationship between shared leadership and team proactivity. Team Performance Management. *An International journal*, 18(1/2), 102-119.
7. Shah, S.A., Memon, M.S. and Laghari, K. (2017). The Impact of organizational culture on the Employees Job Satisfaction. A Study of Faculty Members of public Sector Universities of Pakistan. *Interdisciplinary Journal of Contemporary Research in Business,* 3(8), 847-859.

APPENDIX-A

Table A-1
Test of Reliability

Reliability Statistics		
Cronbach's Alpha	**Cronbach's Alpha Based on Standardized Items**	**N of Items**
.879	.887	36

Table A-2
Percentages of Nominal Scale Variables

Variable	Categories	
Gender	**Male**	59.3
	Female	40.7
Age	**20-50**	86.7
	50-70	13.3
Qualification	**PhD**	28.0
	MPhil	61.3
	Master	10.7
Experience	**1-10**	78.7
	10-20	21.3

Table A-3
Percentages of Nominal Scale Variables

Department	Percent	Department	Percent
Physics	6.0	Chemical Engineer	4.7
Economics	2.7	Electrical Engineer	4.0
Software. Engineer	2.0	Mechanical Engineer	1.3
IT	4.0	Geography	2.0
MATH	4.7	Education	2.7
Computer science	7.3	International relation	2.0
psychology	4.7	English	7.3
Sociology	2.7	Political science	4.7
Bio chemistry +Bio technology	4.0	Islamic	4.0
Mass com	3.3	Environment science	3.3
Botany	4.7	zoology	4.7
Chemistry	5.3	Business Management	8.0

Table A-4
Percentages, Mean and Standard Deviation of all Ordinal Scale Variables

Factor/Variables	S.D%	D%	A%	S.A%	Mean	S.D
Working environment in university is comfortable.	9.3	13.3	54.7	22.7	2.91	0.854
Basic facilities are available for the effective work.	9.3	32.0	52.7	6	2.55	0.747
Cleanliness environment provide in university.	2.7	32	56	9.3	2.72	0.667
The policies does not clearly define by administration.	14.7	59.3	15.3	10.7	2.22	0.826
I cannot decide about my self-decision.	4.0	26.0	41.3	28.7	2.95	0.842
University Top management safeguards us at risky situations.	22.0	19.3	44.7	14.0	2.51	0.988
Environment	**10.3**	**30.3**	**44.1**	**15.2**	**15.8**	**2.93**
I feel very comfortable with my co-workers.	12.0	0	74.7	13.3	3.01	0.505
The organization of my co-workers are cooperative.	33.3	0	42.0	24.7	2.91	0.759
My co-workers are source of inspiration for me.	6.7	17.3	55.3	20.7	2.90	0.801
Working style of my co-workers is different for mine.	14.7	24.0	43.3	18.0	2.65	0.942
A friendly atmosphere prevails among the co-workers in our department.	0	23.3	59.3	17.3	2.94	0.637
Co-workers spend enough time in informal relation.	4.7	27.3	50.0	18.0	2.81	0.781
Co-Workers	**4.3**	**22.9**	**54.1**	**18.7**	**17.2267**	**2.6090**
Supervisor is impolite.	18	26	38	18	2.56	0.986
Supervisor precise good work and is supportive.	7.3	23.3	60.0	9.3	2.71	0.736
Supervisor is competent in his work.	2	28.7	56.7	12.7	2.80	0.676
Supervisor does not delegate powers.	9.3	63.3	16	11.3	2.29	0.790
Supervisors are good relationship with their employees in department.	0.7	22	53.3	24	3.01	0.700
Supervisors always lend their help for the subordinates in university.	15.3	22.7	52.7	9.3	2.56	0.863
Supervision	**8.8**	**31**	**46.1**	**14.1**	**15.9333**	**2.7700**
My work gives me a sense of accomplishment.	8	18	47.3	26.7	2.93	0.875
Most of the time I am frustrated with my work.	8.7	40.0	49.3	2	2.45	0.681
My work is very challenging to me.	0.7	35.3	51.3	12.7	2.76	0.672
Most of the time I do routine work.	8	60	19.3	12.7	2.37	0.806
My work gives me opportunity to learn	0	23.3	43.3	33.3	3.10	0.749

something.						
You complete task (assigned by supervisor) with responsible and professional manner.	22	7.3	54.7	16	2.65	0.998
Work itself	**7.9**	**30.7**	**44.2**	**17.2**	**16.2467**	**2.91245**
My pay is adequate to meet my necessary expenses.	18.7	34.0	21.3	26	2.55	1.072
I receive reasonable annual increments.	8	66.7	14	11.3	2.29	0.771
My pay is commensurate with my qualification.	10.7	44	38.7	6.7	2.41	0.770
Pay at my level is less than as compared to other university.	13.3	23.3	50	13.3	2.63	0.878
I am well paid in proportion to my ability.	7.3	64.0	22.0	6.7	2.28	0.878
The old-age benefits are quite adequate.	16.7	41.3	31.3	10.7	2.36	0.884
Pay/fringe benefits	**12.4**	**45.6**	**29.6**	**12.4**	**14.5200**	**3.91301**
I have good opportunity for advancement in my job.	13.3	14.7	48.7	23.3	2.82	0.942
There is a chance of frequent promotion in my job.	11.3	36.7	44	8	2.49	0.800
Opportunity for promotion is some-what limited here.	4	39.3	48.7	8	2.61	0.694
My job is a dead-end job.	16.7	60.7	14.7	8	2.14	0.786
University provides retirement benefits to its employees.	0	35.3	40.7	24.0	2.89	0.764
University cares for the financial benefits stability of its employees.	23.3	20.7	45.3	10.7	2.43	0.965
Promotion Opportunities	**11.4**	**34.6**	**40.3**	**13.7**	**15.3733**	**3.26985**

D: Disagree **SD:** Strongly Disagree **SA:** Strongly Agree, **A:** Agree,

Table A-5
Bivariate Correlation of all Factors

			TS_EN	TS_CW	TS_SUP	TS_WORK	TS_PAY	TS_PRO
Spearman's rho	TS_EN	Correlation coefficient	1.000	0.098	0.140	0.235	-0.081	0.005
		Sig. (2-tailed)	-	0.234	0.088	0.004	0.325	0.948
		N	150	150	150	150	150	150
	TS_CW	Correlation coefficient	0.098	1.000	0.232	0.138	0.050	0.303
		Sig. (2-tailed)	0.234	-	0.004	0.093	0.547	0.000
		N	150	150	150	150	150	150
	TS_SUP	Correlation coefficient	0.140	0.232	1.000	0.163	0.183	0.075
		Sig. (2-tailed)	0.088	0.004	.	0.046	0.025	0.361
		N	150	150	150	150	150	150
	TS_WORK	Correlation coefficient	0.235	0.138	0.163	1.000	-0.027	0.041
		Sig. (2-tailed)	0.004	0.093	0.046	.	0.740	0.615
		N	150	150	150	150	150	150
	TS_PAY	Correlation coefficient	-0.081	0.050	0.183	-0.027	1.000	0.363
		Sig. (2-tailed)	0.325	0.547	0.025	0.740	.	0.000
		N	150	150	150	150	150	150
	TS_PRO	Correlation coefficient	0.005	0.303	0.075	0.041	0.363	1.000
		Sig. (2-tailed)	0.948	0.000	0.361	0.615	0.000	.
		N	150	150	150	150	150	150

APPENDIX-B

Confirmatory Factor Analysis

Table B-1
Model Estimates of Confirmatory Factor Analysis

	Parameter Estimate	Standard Error	T Statistic	Prob. Level
Environment				
Working environment in university is comfortable.	0.505	0.096	5.240	0.000
Basic facilities are available for the effective work.	0.509	0.054	9.492	0.000
Cleanliness environment provide in university.	0.230	0.070	3.295	0.000
The policies does not clearly define by administration.	0.261	0.067	3.884	0.000
I cannot decide about my self-decision.	0.655	0.108	6.066	0.000
Co-Workers				
My co-workers are source of inspiration for me.	0.401	0.061	6.547	0.000
A friendly atmosphere prevails among the co-workers in our department.	1.602	0.507	3.158	0.000
Co-workers spend enough time in informal relation.	0.265	0.063	3.252	0.001
Supervision				
Supervisor precise good work and is supportive.	0.289	0.097	2.989	0.003
Supervisor is competent in his work.	0.334	0.096	3.494	0.000
Supervisors are good relationship with their employees in department.	0.504	0.125	4.032	0.000
Supervisors always lend their help for the subordinates in university.	0.718	0.078	9.157	0.000
Pay/fringe benefits				
My pay is adequate to meet my necessary expenses.	0.682	0.082	8.362	0.000
My pay is commensurate with my qualification.	1.379	0.268	5.150	0.000
I am well paid in proportion to my ability.	1.425	0.372	3.833	0.000
The old-age benefits are quite adequate.	1.250	0.372	3.833	0.000
Promotion Opportunities				
I have good opportunity for advancement in my job.	0.590	0.079	7.486	0.000
There is a chance of frequent promotion in my job.	0.508	0.117	4.357	0.000
My job is a dead-end job.	0.712	0.193	3.695	0.000
University provides retirement benefits to its employees.	0.960	0.160	6.020	0.000
University cares for the financial benefits stability of its employees.	1.496	0.294	5.082	0.000

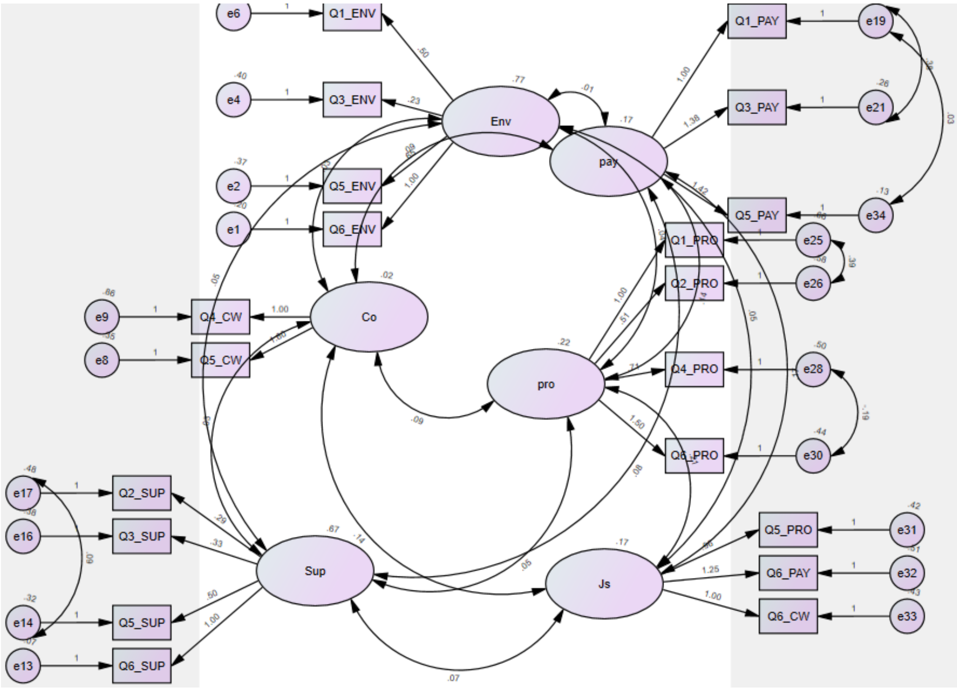
APPENDIX-C

Table C-1
SEM results for Job Satisfaction

Variables	Parameter Estimate	Prob. Level
Coworker↔ Work Itself	.049	.05
Pay↔ Promotion	.224	.000
Supervisor ↔ Coworker	.065	.013
Coworker ↔Promotion	.152	.000
Coworker ↔Pay	.123	.000
Job satisfaction ↔Pay	0.059	0.05

Table C-2
Measures of Goodness of Fit of Model

	$\chi 2$	d.f	p-value	$\chi 2$ /d.f	GFI	AGFI	RMSEA
SEM	426.380	150	0.000	2.84	0.804	0.726	0.06


e6
Q1_ENV
e4
Q3_ENV
e2
Q5_ENV
e1
Q6_ENV
Env
pay
Co
pro
Sup
Js
e9
Q4_CW
e8
Q5_CW
e17
Q2_SUP
e16
Q3_SUP
e14
Q5_SUP
e13
Q6_SUP
Q1_PAY
e19
Q3_PAY
e21
Q5_PAY
e34
Q1_PRO
e25
Q2_PRO
e26
Q4_PRO
e28
Q6_PRO
e30
Q5_PRO
e31
Q6_PAY
e32
Q6_CW
e33

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 363-374

EVALUATION OF LEARNING PROBLEMS OF UOG STUDENT

Zanib Jabeen and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: zaheerabbas@uog.edu.pk

ABSTRACT

Education is essential accomplishment in every person's life. It is frequently viewed as the methods by which the individual will have the option to verify productive work and accomplish throughout everyday life. Learning problems can effect a person's capacity to speak, read, recall, spell, listen, write, organize information and do mathematics. There are some factors in my study "interpersonal relationship" "visual processing" "Expressive language". In this study we evaluate the learning problems of UOG students. The data is collected from 200 students using simple random sampling from students of UOG (Hafiz Hayat Campus). A well-structured questionnaire was used to collect information regarding the subject. The methods used to analyze the data are Descriptive statistics, Confirmatory Factor Analysis. The Results show that students have learning problems and these problems effected on the students study habits.

KEYWORDS

Learning problem, Confirmatory Factor Analysis, Study skills, Expressive Language, Reading.

INTRODUCTION

Learning problem means student have problem in the learning process. Learning problems may be concern to the area of reading, writing, listening, speaking, mathematical skills. For example a student who can write everything correctly if someone dictate him what it show? It means student have problem in area reading, reader was completely unable to read either the printed or written characters properly. Students have learning problems but he had no apparent problem. The reading problem sometimes related to the language. Medical defined this phenomenon as individual illness. Problem is owned by the individual and needed to be treated. Therefore the researchers stated that learning problems exist in student's learning.

Learning Problems

Education is one of the most important accomplishments in an individual's life. It is considered to be the means through this person will be able to gain employment and achieve success in life. Consequently, Education is a valuable tool for every person. However, there are varying degrees of ability to learn. Not all students can learn with the same efficiency as other students. But research shows that the students with poor learning ability may have some type of problems. These students are able to learn with the proper training, instructional strategies and accommodations. These adults are referred to as having learning problem.

The term learning problem describe as neurobiological problem in which a person's brain work or is structured differently. These differences interfere with a person's abilities to utter, hear, peruse, spell, recall, organize information, and do mathematics.

Learning disabilities affect a person's ability in the areas of listening, speaking, reading, writing, and mathematics. A student has problem:

i. Interpersonal relationship
ii. Visual spatial organization
iii. Organization and planning skills
iv. Flexible concept formation
v. Study skills
vi. Specific academic areas
vii. Social judgment

Types of Learning Problems

Two major types of learning disabilities

Development Speech and Language Disorder

i. Articulation Disorder (problem in producing speech sounds)
ii. Expressive Language Disorder (disability in communicating with spoken language)
iii. Receptive Language Disorder (disability in understanding what other people say).

Effects of learning problems

Learning problems can effect throughout individuals and others life around them in many different ways:

- Students with learning problems often face failure in their life, misunderstanding from teachers resulting in humiliation in the classroom and ridicule from other students or peers, and low self-esteem.
- Parents of such adults often become stressed, and siblings may grow resentful or feel jealous of attention given to the learning disabled adults.

Success of Students with Learning Problems

Success means different things to different people. It may mean something different at different times in a person's life. However, success may differ according to person's expectations. This may include good friends, positive family relations, being loved, confident, job satisfaction, physical and mental health, financial comfort, spiritual contentment, and an overall sense of meaning in one's life. Different individuals may place lesser or greater importance on these various components of success.

OBJECTIVES

1. To investigate the significant variables for each factors of learning problems by using Confirmatory Factor Analysis.
2. To identify the factors of learning problems, this can highly discriminate the effect of learning problems on study habits.
3. To develop a discriminant function so that on can classify the respondent on which the effect of learning problems on their study habits is medium or high.

HYPOTHESIS

H0: There is relationship between students and their learning problems.
H1: There is no relationship between students and their learning problems.

OPERATIONAL DEFINITION

Learning problem exists when students have ability to learn is average or higher, but his skills in one or more academic areas are significantly weaker. Learning problem is not a life-long problem that affects the manner in which individuals select, retain, and expression information. It reflects difficulties in encoding and decoding information as it travels between the senses and the brain. Learning problem are also known as learning differences based on the fact that certain individuals learn differently they aren't unable to learn, but respond best to ways of learning that are different from traditional teaching methods. Learning problems vary in the level of severity and invariably interfere with the acquisition or use of one or more of the following important skills:

a) **Oral Language**
   i. Listening
   ii. Speaking
   iii. Understanding

b) **Reading**
   i. Decoding
   ii. Comprehension

c) **Written Language**
   i. Spelling
   ii. Written expression

d) **Mathematics**
   i. Computation
   ii. Problem solving

Learning problem are permanent-information processing disorder that effects individuals manners of learning with average to above average intelligence learn. Deficits in areas such as reading, mathematics and written language are presumed to be due to a central nervous system dysfunction. Learning problems occur regardless of gender, race, or ethnic origin and they are not the result of a poor academic background, mental retardation, or emotional disorders. The world health organization stated learning problems as: A state of arrested or incomplete development of mind learning problem is not a disease, nor a physical or mental illness. Unlike the physical or mental illness, so far as we know it is not treatable. Internationally three criteria's are identify the learning problems:

i. Intellectual impairment
ii. Social or adaptive dysfunction
iii. Early onset

DICTIONARY DEFINITION

Academic of Learning Disorder

- Dyslexia_ reading ability disorder (disorder in understanding the written word)
- Dysgraphia _ writing ability disorder (difficulty in forming written letters)
- Dyscalculia _ computing ability disorder (difficulty in understanding and Math calculation).
- Reasoning: difficulty in organizing and integrating thoughts
- Organization skills: difficulty in organizing all facets of learning

General strategies for working with students with learning problems

Teachers report that the following strategies have been effective with some students who have learning disabilities:

- Capitalize on the student's strengths.
- Provide high structure and clear expectations.
- Use short sentences and a simple vocabulary.
- Provide opportunities for success in a supportive atmosphere to help build self-esteem.
- Allow flexibility in classroom procedures (e.g., allowing the use of tape recorders for note-taking and test-taking when students have trouble with written language).
- Make use of self-correcting materials, which provide immediate feedback without embarrassment.

Adaptability Struggle

Switching from traditional classroom and face to face instructor training to computer-based training in a virtual classroom makes the learning experience entirely different for students. Their resistance to change doesn't allow them to adapt to the online learning environment, whereas it takes time for them to get accustomed to Course Management Systems (CMS) and the methods of computer-based education. While passive listening and notes taking are expected in a traditional classroom, online discussions or creating a web page demand springing into action. Students with a "traditional" mindset find it difficult to adapt; however, they need to accept the new learning circumstances with an open mind and heart. Understanding the benefits of eLearning and even discussing them with their peers may change this mindset and better prepare students for online classes.

Technical Issues

Many students are not provided with the high bandwidth or the strong internet connection that online courses require, and thus fail to catch up with their virtual classmates: Their weak monitors make it hard to follow the Course Management System and their learning experience becomes problematic. Moreover, most of them live off campus and find it difficult to keep in tune with the technical requirements of the chosen course. Some of them don't even own computers and seek help in Learning Resource Centers for technical assistance. The only solution to this problem is knowing exactly what kind of technological support they will need for a certain course before enrolling in it, as well as properly equipping themselves for the course's successful completion.

Computer Literacy

Although students are generally tech savvy, and thus able to manage computers well, lack of computer literacy is a major issue among students today. Many of them cannot operate basic programs such as Microsoft Word and PowerPoint and therefore are not able to handle their files. Furthermore, many students find fixing basic computer problems troublesome, as they have no knowledge in this area. However, technological proficiency is a must for following online courses, as it enables students to manage their assignments and courseware in an organized manner without struggling. Basic courses in computer literacy enhance students' knowledge in the field; having a fundamental knowledge of computer hardware would help them participate in online classes without interruptions and hindrances.

Time Management

Time management is a difficult task for learners, as online courses require a lot of time and intensive work. Furthermore, whereas it is mostly adults who prefer web-based learning programs for their place and time flexibility, they rarely have the time to take the courses due to their various everyday commitments. A regular schedule planner would be a significant help to these learners, as they could even set reminders for their courses and assignments.

Self-Motivation

Self-motivation is an eLearning essential requirement; however, many online learners lack it, much to their surprise. After enrolling in distance learning courses, many learners fall behind and nurture the idea of giving up, as difficulties in handling a technological medium also seem insurmountable. Students need to find the motivation to follow the new educational trends and also properly equip themselves for future challenges in their education and careers. Only a positive attitude will help them overcome the challenges in eLearning; though this is hard to practice, students need to understand that it is necessary in order to reap the eLearning's benefits in the future.

E-learning is good news, but at its initial stage it poses certain threats to students. Attitude change and technological literacy would help them gain confidence in order to succeed in their courses with a positive vibe.

FACTORS THAT AFFECT STUDENT LEARNING

Intellectual Factor

The term refers to the individual mental level. Success in school is generally closely related to level of the intellect. Pupils with low intelligence often encounter serious difficulty in mastering schoolwork. Sometimes pupils do not learn because of special intellectual disabilities. A low score in one subject and his scores in other subjects indicate the possible presence of a special deficiency. Psychology reveals to use that an individual possess different kinds to intelligence. Knowledge of the nature of the pupil's intellect is of considerable value in the guidance and the diagnosis of disability. The native capacity of the individual is of prime importance in determining the effectiveness of the, learning process.

Learning Factors

Factors owing to lack of mastery of what has been taught, faulty methods of work or study, and narrowness of experimental background may affect the learning process of any pupil. If the school proceeds too rapidly and does not constantly check up on the extent to which the pupil is mastering what is being taught, the pupil accumulates a number of deficiencies that interfere with successful progress, In arithmetic, for instance, knowledge of basic addition is essential to successful work in multiplication. Weakness in addition will contribute directly to the deficiency in multiplication. Likewise, failure in history may be due to low reading ability or weakness in English. Similarly, because of faulty instruction, the pupil may have learned inefficient methods of study. Many other kinds of difficulty which are directly related to learning factors may interfere with progress.

Physical Factors

Under this group are included such factors as health, physical development, nutrition, visual and physical defects, and glandular abnormality. It is generally recognized that ill health retards physical and motor development, and malnutrition interferes with learning and physical growth , Children suffering from visual, auditory, and other physical defects are seriously handicapped in developing skills such as reading and spelling. It has been demonstrated that various glands of internal secretion, such as the thyroid and pituitary glands, affect behavior. The health of the learner will likely affect his ability to learn and his power to concentrate.

Mental Factors

Attitude falls under mental factors attitudes are made up of organic and kinesthetic elements. They are not to be confused with emotions that are characterized by internal visceral disturbances. Attitudes are more or less of definite sort. They play a large part in the mental organization and general behavior of the individual

Attitudes are also important in the development of personality. Among these attitudes aw interest, cheerfulness, affection, prejudice, -open mindedness, and loyalty. Attitudes exercise a stimulating effect upon the rate of learning and teaching and upon the progress in school.

The efficiency of the work from day to day and the rapidity with which it is achieved are influenced by the attitude of the learner. A favorable mental attitude facilitates learning. The factor of interest is very closely related in nature to that of symbolic drive and reward

Emotional and Social Factors

Personal factors, such as instincts and emotions, and social factors, such as cooperation and rivalry, are directly related to a complex psychology of motivation. It is a recognized fact that the various responses of the individual to various kinds of stimuli are determined by a wide variety of tendencies Some of these innate tendencies are constructive and others are harmful. For some reason a pupil may have developed a dislike for some subject because he may fail to see its value, or may lack foundation. This dislike results in a bad emotional state.

Some pupils are in a continuing state of unhappiness because of their fear of being victims of the disapproval of their teachers and classmates. This is an unwholesome attitude and affects the learning process to a considerable degree. This is oftentimes the result of bad training.
Social discontent springs from the knowledge or delusion that one is below others in welfare.

Teacher's Personality

The teacher as an individual personality is an important element in the learning environment or in the failures and success of the learner. The way in which his personality interacts with the personalities of the pupils being taught helps to determine the kind of behavior which emerges from the learning situation The supreme value of a teacher is not in the regular performance of routine duties, but in his power to lead and to inspire his pupils through the influence of his moral personality and example. Strictly speaking, personality is made up of all the factors that make the individual what he is, the complex pattern of characteristics that distinguishes him from the others of his kind. Personality is the product of many integrating forces In other words, an individual's personality is a composite of his physical appearance, his mental capacity, his emotional behavior, and his attitudes towards others. Effective teaching and learning are the results of an integrated personality of the teacher, Generally speaking, pupils donot like a grouchy teacher who cannot control his temper before the class. It is impossible for a teacher with a temper to create enthusiasm and to radiate light and sunshine to those about him ,The teacher must therefore recognize that in all his activities in the classroom he is directly affecting the behavior of the growing and learning organism.

Material and Methods

Theoretical Framework for the Capacity Building Model for Mathematics Achievement. The roots of the XYZ Systems Approach to mathematics reform are grounded in both the literature in the STEM education field and in the extensive experiences of the mathematics educators and mathematicians who will be involved in implementing and researching the proposed systems research. Based on our own experiences and supported by Cohen and Hill's *Learning Policy* (2001), we believe that many educational reforms fail because teachers and the systems in which they work are not placed at the center of the reform

Target Population

The students of Masters and BS-Hons studying in university of Gujrat Pakistan.

Sampled Population

The study population consisted on all students of university of Gujrat (Hafiz Hayat Campus).

Sample Size

Sample size consists of 200 respondents from University of Gujrat

Sampling Unit

Every Student of Masters' and BS Hons in University of Gujrat was considered as sampling unit of the study.

Sampling Method

Simple Random Sampling was used for sample selection.

Data Analysis

The Following Analysis used in data analysis

- Descriptive Statistics
- Confirmatory Factor Analysis
- Multiple Discriminant Analysis

Confirmatory Factor Analysis

Confirmatory factor analysis (CFA) is a multivariate statistical procedure that is used to test how well the measured variables represent the number of constructs. Confirmatory factor analysis and exploratory factor analysis (EFA) are similar techniques, but in exploratory factor analysis, data is simply explored and provides information about the numbers of factors required to represent the data. In exploratory factor analysis, all measured variables are related to every latent variable. But in confirmatory factor analysis, researchers can specify the number of factors required in the data and which measured variable is related to which latent variable. Confirmatory factor analysis is a tool that is used to confirm or reject the measurement theory.

Descriptive Statistics

Descriptive statistics is used to get summary of the data. It provides the features and picture of data. Measures of central tendency, measure of dispersion used in quantitative and percentage is used to get results about qualitative.

Multiple Discriminant Analysis

Multiple discriminant analysis (MDA), also known as canonical variates analysis (CVA) or canonical discriminant analysis (CDA), constructs functions to maximally discriminate between *n* groups of objects. This is an extension of linear discriminant analysis (LDA) which - in its original form - is used to construct discriminant functions for objects assigned to two groups.

RESULT AND DISCUSSION

Reliability of the questionnaire is a key factor for statistical results. Table 1 consists of the test of reliability of the problems in Interpersonal Relationships, Visual processing, Effect of learning problems on study habits, Problems in Academic Skills, Mathematical skills value of Cronbach's alpha based on standardized item is 0.82 which indicates that the data is reliable for further statistical analysis.

Descriptive statistics are considered primarily to explain the basic features of the data. It provides simple summaries about the sample measures. Table 2 contains the mean and standard deviation of age of the in years, Gender and program of the respondent. Results show that the minimum age of the respondent is 18 years and maximum is 25 years. The mean age is 19.9350 years with S.D 0.9671; it shows that most of the respondents are within 18 to 22 years of age.

Table 3 contains the sample information; there were students 82.5 percent were from BS and remaining 17.5 percent were from MA/MSc Program. There were almost 71.5% of the respondents were Female and remaining 28.5% are Male in overall sample of students.

Table 5 with interpersonal relationship students are agree with this statement and all other factors are average by the student's statement.

Table 1

Cronbach's Alpha	N of Items
0.82	**36**

Table 2

	Age	Program
Mean	19.9350	1.1750
Std. Deviation	0.9671	0.3809

Table 3

	Percent
BS	**82.5**
MA/MSC	**17.5**

Table 4

<table>
<tr><th>Variable</th><th colspan="4">Categories</th></tr>
<tr><th>Gender</th><th colspan="2">Male</th><th colspan="2">Female</th></tr>
<tr><th rowspan="2">Program</th><th>BS</th><th>MA/MSC</th><th>BS</th><th>MA/MSC</th></tr>
<tr><td>20.5%</td><td>8%</td><td>62%</td><td>9.5%</td></tr>
</table>

Table 5

S#	Statement	SD%	D%	N%	A%	SA%	Mean	S.D
1. PROBLEM IN INTERPERSONAL RELATIONSHIP								
1	With other students influences your study.	9.0	13.0	19.5	49.5	9.0	3.365	1.103
2	Play an effective role in group work to overcome learning disability.	9.5	20.5	14.5	44.5	11.0	3.270	1.184
3	Contribute significantly in academic interactions.	5.0	11.5	15.5	40.0	28.0	3.545	0.970
4	Provide facilities to reduce your disabilities.	8.0	10.5	21.5	33.0	27.0	3.605	1.215
5	Develop interaction with teacher outside the class and it has positive influence on study.	4.5	12.0	22.0	31.0	30.5	3.710	1.154
6	Teacher behavior effect the student learning.	5.5	6.5	9.0	39.0	40.0	4.015	1.118
2. VISUAL PROCESSING								
1	Eyes become red after a short time of work.	12.5	22.0	26.0	29.0	10.5	3.030	1.198
2	Oral reading is choppy: words skipped, endings left off, frequent repetitions.	3.5	19.5	25.0	44.5	7.5	3.33	0.987
3	Often ask you to repeat yourself.	1.5	15.0	42.0	37.5	4.0	3.275	0.820
4	You often face headaches after a short time working at reading or writing.	7.5	26.5	32.5	25.5	8.0	3.140	2.372
5	Closes one eye while reading or writing.	22.5	28.5	8.5	25.0	15.5	2.825	1.426
6	Lifts eyes from page frequently to glance around.	5.5	21.5	30.5	35.0	7.5	3.175	1.029
3. EFFECT OF LEARNING PROBLEMS ON STUDY HABITS								
1	Your planned schedule.	13.0	10.5	28.5	39.5	8.5	3.200	1.151
2	Your goal of study.	5.0	10.0	26.5	42.5	20.5	3.725	0.918
3	To maintain education with learning disability.	12.0	20.0	31.0	27.0	10.0	3.030	1.164
4	Confidence level of education.	7.5	13.5	38.5	31.5	9.0	3.210	1.035
5	Class interaction with teachers and students.	2.5	18.0	27.0	29.5	23.0	3.525	1.107
6	In understanding a new concept and make you aggressive.	13.5	21.0	29.5	22.0	14.0	3.020	1.239
7	Your presentation skills.	5.0	5.0	24.0	50.0	16.0	3.670	0.975
4. PROBLEMS IN ACADEMIC SKILLS								
i. Reading								
1	The learner shows marked difficulty in oral and silent reading.	12.6	20.6	26.6	25.1	15.1	3.095	1.249
2	Pattern are slow and deliberate.	1.0	17.0	38.7	30.7	12.6	3.368	0.943
3	Skip words, re-reads lines in oral reading.	6.0	20.6	32.2	28.1	13.1	3.216	1.100
4	Lose place on page.	5.5	34.2	24.6	32.2	3.5	2.939	1.013
5	Style in halt and jerky.	20.6	23.6	23.1	26.1	6.5	2.743	1.234

ii. Expressive Language								
1	Problems with grammar and syntax.	19.1	14.1	17.1	42.7	7.0	3.045	1.272
2	Spells words differently in the same document.	13.6	25.6	17.6	33.2	10.1	3.005	1.241
3	Weak visual memory for spelling.	22.6	13.6	17.1	35.2	11.6	2.995	1.365
4	Reverses letters in spelling, e.g., Friday becomes Friday, girl becomes gril.	10.6	24.6	22.1	26.1	16.6	3.135	1.257
5	Punctuation errors are common.	7.5	19.1	23.6	38.7	11.1	3.266	1.121
6	Continuously whispers to self while writing.	10.6	15.6	23.1	34.2	16.6	3.306	1.223
5. MATHEMATICAL SKILLS								
31	Trouble remembering math facts and procedures.	22.6	14.1	20.1	30.2	13.1	2.969	1.370
32	Difficulty copying numbers and working with numbers in columns.	32.2	14.6	20.1	28.1	5.0	2.593	1.325
33	Trouble with left/right orientation.	19.6	24.1	17.6	27.6	11.1	2.864	1.316
34	Cannot remember in which direction to work in carrying out simple math.	9.5	17.1	40.7	26.1	6.5	3.030	1.039
35	Confuse in understanding similar numbers or transpose numbers.	7.5	24.1	33.7	20.6	14.1	3.095	1.144
36	Difficult to follow sequential (in order) procedures and directions with multiple steps.	5.0	19.1	41.2	23.6	11.1	3.165	1.023

Table 6

Factor	Wilks' Lambda	F-Value	d.f	df2	P-value
PIR	.902	21.506	1	197	.000
VIS	.978	4.499	1	197	.035
ELPS	.981	3.739	1	197	.055
R	.971	5.807	1	197	.017
EL	.988	2.383	1	197	.124
MS	.869	29.763	1	197	.000

Table-4
Canonical Discriminant Function

Function	Eigen Value	% of Variance/ Cumulative %	Wilks' Lambda	Chi-square	df	p-value
1	0.281	100.0	0.781	48.075	6	0.00

Functions at Group Centroids

Gender	Function
Female	-0.334
Male	0.833

Structure Matrix

	Function
MS	.733
PIR	-.623
R	.324
VIS	.285
ELPS	.260
EL	.207

CONCLUSION

In this study we address the problem of learning problems. We want to check learning problem on study habits at university of Gujrat. For the investigation of this purpose we have used discriminant analysis after confirmation of factors by using confirmatory factor analysis. on the basis of Discriminant Analysis we conclude that , problem in interpersonal relationship are the best discriminant for predicting the levels of study habits then other learning problem factors. The student have medium level of effect on study habits due to learning problem. It means that at UOG the learning problem are not much effecting the level of study habits.

REFERENCES

1. Statistics Solutions. (2013). Confirmatory Factor Analysis. Retrieved from https://www.statisticssolutions.com/wp-content/uploads/wp-post-to-pdf-enhanced-cache/1/confirmatory-factor-analysis.pdf
2. Ahdesmäki, M. and Strimmer, K. (2010). Feature selection in omics prediction problems using cat scores and false non-discovery rate control. *Ann. Appl. Stat.*, 4(1), 503-519.
3. Rao, C.R. (1948). The utilization of multiple measurements in problems of biological classification. *J. R. Stat. Soc.*, B. 10(2), 159-203.
4. White, J.W. and Ruttenberg, B.I. (2007). Discriminant function analysis in marine ecology: some oversights and their solutions. *Mar. Ecol. Prog. Ser.*, 329, 301-305.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 375-381

INTEGRATING SURVEY DATA WITH BIG DATA: POSSIBILITIES AND CHALLENGES

Naila Alam and **Tooba Khan**
Kinnaird College for Women, Lahore, Pakistan
Email: naila.alam1969@gmail.com
toobakhan1333@gmail.com

ABSTRACT

Combining information from two or more data sources is known as data integration and is an emerging area of research. There are many possibilities of data integration using many combinations of data sources. One of them is integration of survey data with Big Data that are typically combined from dissimilar sources and are integrated to form new data sets. The developed countries are focusing on Modernization of Official Statistics by incorporating new data sources in their statistical production processes. This involves linking records together and to produce more timely and more aggregated statistics at higher frequencies than traditional approaches. This paper explores the challenges in integration of survey data with big data that are faced by the researchers and statisticians. Some suggestions are proposed to improve and to modernize the official Statistics in Pakistan.

1. INTRODUCTION

Quality data generation is essential for evidence based policy and planning. The government of Pakistan has initiated the process of reviewing and reporting needs for data gap and to create robust data network that can contribute towards achievement of the Sustainable Development Goals ((SDG) 2015-30). The reliable and updated data on various social and economic indicators about Household, Children, Women and Men provide a basis for evidence based polices and to compare the progress over time in relation to various key socio-economic factors. Detailed and updated data of any type can be used in policy formulation by policymakers to improve e-Governance, health facilities and delivery and to improve management of places and cities.

The recent project of the government of Pakistan "digital Pakistan" emphases on improvement of the quality of life and well-being of citizens by ensuring availability of reachable internet services to all citizens around the country that will, in turn, promote e-Governance to ensure productivity, transparency and accountability. This digitalization will surely result in bulk of unstructured and big data. Exploring the potential of big data for producing meaningful official statistics will be definitely a very important issue in Pakistan in the coming days.

Recently, some of the studies have explored the potential and benefits in integration of Big Data with survey data. The Bureau of the Conference of European Statisticians (CES), in 2017, stated few experiments to integrate big data from smart sensors, social networks, mobile companies, satellite imagery, credit card transactions, web pages, and various other sources, with traditional sources of official statistics. Mass imputation approach was used to integrate big observational data with survey data (Yang & Kim, 2018). The policy makers and Social scientists have explored benefits from studies regarding integrating survey data and digital trace data. Benefits include the information of human behavior at a large scale instead of a small sample, the improvement and validation of measurements error and evidences to improve causal inference in experimental settings (Stier, Breuer, & Thorson, 2019). A market opportunity can be identified, target audience can be understood, and findings can be incorporated into creative execution. Infusing survey research with big data also reduces the volume of questions to be asked as big data provides more answers. So, the understanding of consumer behavior can grow exponentially overtime as the two are brought together. (Patton, 2018).

This paper looks into the possibilities and challenges of integrating data sets from different sources, specifically big data with survey data, to improve the quality of Government services and for evidence based policies through real-time integrated management and data analytics.

Big Data: The term "Big Data" can be described as a set of rich and complicated characteristics, texts, pictures and outcomes, all associated with data. Although, the definition of big data is complex and diverse but its main characteristics mutually accepted are:

Volume: (Big Data is large enough such that it cannot be handled with traditional tools and approaches),

Velocity: (Generated in Real-Time) and

Variety: (It consists of variety of complex datasets that can include unstructured text, logs, images and videos).

Sources of Big Data:

- Internet is the main source of big data and users that can connect with other devices using internet as the communication tool are the generators of this data. This data which is generated from connected devices is named as "The Internet of Things" (IOT), include behavioral, trends and humans data. Behavioral data is a subset of the IOT such as physical activities, location and health status of the consumers. This data is mostly recorded from personal devices such as smart watches and smartphones used by persons.
- The interactions and relationships of users in business and in management are known as transaction data. This data consists of data regarding online orders, credit card activities, payments and online billing. This data is mostly used to capture communication and satisfactions of a customer with a company and the provided services.
- Administrative data is another source of big data because these data sets are usually large in volume and are unstructured. The content of administrative data is collected by statistical or other agencies without using any design. It can be in the form of registers collected by public offices or online registration of citizens, such

as patient's admissions data in a national hospitals database, tax payers data, pensioners or driver licenses databases.

- Another source of big data is data brokers that collect and store data about consumers and customers. They integrate different big data sources including behavioral, transactional and attitudinal data to create a profile for each individual consumer regarding some relevant research problems. Acxiom, Epsilon, Experian Marketing Services, Aristotle, and Nation Builder are examples of some data broker companies ("Committee on Commerce, Science and Transportation 2013").

Survey Data: The structured data that is collected from a sample of respondents using a designed questionnaire and using a proper sampling design is named as survey data. This data is gathered from a target audience or sample about a particular topic of interest using a designed questionnaire to conduct research and to get evidences for future policies.

There are many methods used to gather survey data for statistical analysis in research from the desired sample of individuals. The mostly used techniques are online surveys, telephonic surveys and face-to-face surveys.

The New Scenario: The Pakistan Bureau of statistics is the custodian and responsible for producing official statistics and conducting household surveys at large scale for gathering data for official use in Pakistan. They have successfully carried out a large number of household surveys (structured data), such as "Multiple Indicator Cluster Survey" (MICS), "Pakistan Social and Living Standards Measurement Survey" (PSLM), "Pakistan Demographic & Health Survey" (PDHS) and "Pakistan Demographic Survey" (PDS) at the provincial and national level.

Recently, the launch of government of Pakistan's project "digital Pakistan" targets at availability of internet services to all Pakistanis and setting up integrated government databases. This data will enable to identify pattern on big data through data mining and data analysis to improve the quality of e-Government services.

This new scenario raises many questions such as "who will be responsible and in charge of data gathering and establishment of databases" and "Is that custodian department sufficient in handling and using necessary techniques related to big data?"

Other questions may be:

1. Will the traditional techniques and approaches to deal with structured data work with big data?
2. Will the survey research by traditional methods play a less dominant role in this era of big data? And
3. Can the big data and survey data be integrated?

Recent researches have discussed ideas for blending Big Data with traditional survey to support many research questions.

The classic statistical procedure is one in which a population frame, survey design and a sampling technique are used to collect data from a sample of respondents. The data generation and production is very different in a Big Data domain as compare to a traditional survey. The most apparent Big Data advantage is that it is significantly cheaper than surveys. The large scale household surveys are very expensive and time consuming as they require a team of enumerators, experts and supporting staff to collect data. In contrast, Big Data rely on computer software and electronic data gathering can be much more cost effective, but at the same time, it is expensive to clean and process the big data that require allocation of the expert team.

The new paradigm is to reconcile, aggregate, and correlate different data sources. The integration of big data with different data sources is a latest area of research and a very little work has been inspected the ways in which it may best be used with other types of data to provide richer datasets. The developed countries are focusing on innovation and to foster the most important types of skills for working with Big Data. The Scheveningen Memorandum (ESSC, 2013) in 2013 and the Big Data Action Plan and Roadmap (ESSC, 2014) in 2014 by the European Statistical System (ESS) has a vision to explore the potential of big data for producing official statistics. The International statistical offices such as the US Census Bureau, Statistics Canada and R&D Department at Statistics Sweden are working on innovation and exploration in Big Data integration techniques that can be used to improve explanatory power and reduce costs of surveys and Census operations (Bostic, 2013).

This paper specifically studies the possibilities and challenges in integrating big data with survey data that may encounter in attempting to integrate these different sources of data.

2. POSSIBILITIES

The exploration of high-dimensional and unstructured data is a great challenge for both of the survey statisticians and data analyst. The survey and big data experts can bring together the answers to many research questions that are not possible only from a survey or from some big data sources to advance the official statistics and science of measuring and exploring the human thoughts.

The administrative data is a form of big data and has been used for official and statistical purposes for a long time (Wallgren & Wallgren, 2014). Administrative data are structured and well defined as compare to other Big Data sources in the form of records collected by public offices, such as national department's data bases include health, tax or driver licenses databases. There are many studies that use Administrative data in survey data to reduce coverage error and non-response. Many international government agencies are incorporating augmentation of big data with survey data. The "National Center of Health Statistics" (NCHS) is working on the development of linkage program to link various NCHS surveys to available administrative records under an interagency agreement among different institutions to use these linked data files to support various research problems.

The most related to survey data as a source of Big Data is Paradata. "Paradata is data that comes in real time about the process of answering the survey questions" (Callegaro, 2013), including data collected by systems and enumerators before, during, and after the administration of a questionnaire. For example, number of attempts done to contact a person, number of halves in which questionnaire was completed, gender of interviewer, Day of interview, Month of interview, Year of interview, Start of interview (Hour), Start of interview (Minutes), End of interview (Hour), End of interview (Minutes). Paradata often come in real time and are in complex formats. This Paradata can be used to reduce nonresponse in big scale household surveys and by examining the patterns relating with answering some questions or for follow up studies (Olson, 2013).

Another way to use Big Data is to improve the quality of questionnaire and to ask better questions in surveys by Data mining and pattern detection through big data. For example the Big data available from IOT can be used to validate different question wording and to determine what is closer to the “true value” (Gibson et al. 2014).

The Augmentation of Big Data with survey data can be useful for detailed and structured information. For example, different websites ask their users to answer few “categorical” questions about their services. For example, users can be asked if they are satisfied or not with the restaurant they just ordered online food.

3. CHALLENGES

The modern world is relying on data for policy making and for predicting existing and future patterns. The culture of using available data and leveraging new analytic and visualization tools to extract information from this available data is increasing. In Pakistan, we are lagging behind in utilizing these modern techniques and opportunities. The question in this scenario is that “What are the operational and statistical challenges associated with the use of Big Data”?

The major reason is lack of collaboration between statisticians and computer experts and the challenge is to bring the statisticians and computer experts at the same forum. It will definitely not be feasible and possible for statisticians to become proficient in all new tools and skills related to integration of big data and survey data. Instead, the collaboration with others who have different expertise and skills can be fruitful.

The big challenges in this new era of big data are to explore the impact of switching to new data collection techniques and the use of this data to draw inferences integrating with existing surveys. The Development of new algorithms and modeling techniques which could enable to combine dissimilar data sets and to make them ready for meaningful analysis is an open area of research to improve the quality of data through looking into complex constructs, integrating a multitude of diverse resources and introducing new analytical techniques and visualization tools.

Another challenge will be to remove legal and administrative barriers for exchange of data within different public and private sector departments to develop a shared platform. This will support researchers, analytics, and decision making to blend Big Data with traditional surveys in a precise way depending on the demands of a given problem or situation.

The big data are aggregated to create individual profiles by data brokers such as a person's demographic and behavioral information that consists of thousands of individual pieces of information integrated. The Brokers then sell the profiles to other organizations so they can research people for various reasons. Data brokers also sell the profiles of these individuals to government agencies such as income tax or FBI for different purposes. One challenge in this world of social media and information is to develop laws to protect the user's basic rights and privacy.

4. SUGGESTIONS

Some of the suggestions to modernize and to update the official statistics in Pakistan based on guidelines used by international statistical agencies are given as under:

The custodian department should be responsible for establishing national data bases and must have strong linkages with other departments to avoid duplication.

Adequate IT infrastructure is required to establish national databases and to use it for data mining and for producing useful official statistics. Specific software and hardware are required to ensure the adequate collection, storage, analysis and reporting of information, for the IT implications of big data are complex.

It is imperative to foster the most important types of skills for working with Big Data, such as IT skills, statistical skills and other skills including ethics, data governance and creative problem solving. Determined data analytics and data scientists having affinity for IT are required to extract valuable knowledge from data. International collaborations may be very helpful and beneficial for the official statistical community.

Diffusion of innovations can increase adoption and utilization of Big Data despite the challenges faced by developing countries which can limit effective utilization and capability of these technologies. Strategies to accelerate technology diffusion are necessary, such as investment in human capital, increasing digitization levels, cultural change, and development of legal frameworks to deal with privacy concerns.

REFERENCES

1. Bostic, W.G. (2013). Big Data Projects at the Census Bureau. Paper presented at the *Council of Professional Associations on Federal Statistics*, Washington, DC, USA.
2. Callegaro, M. and Yang, Y. (2018). The Role of Surveys in the Era of "Big Data". In: Vannette D. and Krosnick J. (Eds.), *The Palgrave Handbook of Survey Research*. Palgrave (pp. 175-192). Cham: Macmillan.
3. Callegaro, M. (2013). Paradata in Web Surveys. In *Improving Surveys with Paradata: Analytic Use of Process Information*, edited by Kreuter, F. (pp. 261–79). Hoboken, NJ: Wiley.
4. Gibson, W.K., Cronin, H., Kenny, R.A. and Setti, A. (2014). Validation of the self-reported hearing questions in the Irish Longitudinal Study on Ageing against the Whispered Voice Test. *BMC Research Notes*, 7(1), 361.
5. Japec, L., Kreuter, F., Berg, M., Biemer, P., Decker, P., Lampe, C. and Usher, A. (2015). Big data in survey research: AAPOR task force report. *Public Opinion Quarterly*, 79(4), 839-880.
6. Government of Pakistan (2017). *Data Reporting Gap, Draft Report - For Review.* Ministry of Planning and Development Pakistan.
7. Olson, K. (2013). Paradata for Nonresponse Adjustment, *The Annals of the American Academy of Political and Social Science*, 65(1), 142-170.

8. Patton, B. (2018). Understanding and Leveraging Big Data for Audience Insights. Retrieved from https://www.targetmarketingmag.com/post/understanding-leveraging-big-data-audience-insights/
9. Porter, S. and Lazaro, C.G. (2014). Adding Big Data Booster Packs to Survey Data. Paper presented at the *Council of American Survey Research Organizations Digital Research Conference*, San Antonio, TX, USA.
10. Stier, S., Breuer, J., Siegers, P. and Thorson, K. (2019). Integrating Survey Data and Digital Trace Data: Key Issues in Developing an Emerging Field. *Social Science Computer Review*, 1-14.
11. Tazi, L. (2016). *Big data and official statistics in Morocco: opportunities and challenges*. Retrieved from http://www.iaos-isi.org/images/IAOS2017-19/2016 Conference/P6D/Lina_Tazi_Paper.pdf
12. Wallgren, A. and Wallgren, B. (2007). *Register-Based Statistics: Administrative Data for Statistical Purposes*. New York: Wiley & Sons.
13. Wallgren, A. and Wallgren, B. (2014). *Register-Based Statistics: Statistical Methods for Administrative Data* (2nd ed.). New York: Wiley & Sons.
14. Yang, S. and Kim, J.K. (2018). Integration of survey data and big observational data for finite population inference using mass imputation. arXiv preprint arXiv:1807.02817.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 383-388

A STUDY ON DIALYSIS PATIENT TO CHECK THE EFFECT OF PSYCHOLOGICAL SOCIAL AND ECONOMIC PROBLEM| ON THEIR HEALTH

Maryam Arzoo[1], Erum Shahzadi[2] and **Sania Naeem**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: [1]pervaiza713@gmail.com
[2]erum.shahzadi@uog.edu.pk

ABSTRACT

Chronic kidney disease is very common day by day. The most frequent method to treat these diseases is dialysis. As it leads to the mortality so the cross sectional study has been carried out to check the psychological stressor by the patient. We hear this sentence commonly that Health is Wealth, the person can do everything when he is healthy. Psychologically, if the patient cannot diagnose their disease, these patient depressed more than normal people and the person who have negative thinking suffer from high blood pressure.

The main purpose of this study to explore the psychological economic and social issues faced by the patient of dialysis. In this study we used purposive sampling. Sample size is 270 including 17 years to 81. According to our result mostly females of middle class family are affected. The classification table of MLR shows that 98.4% model is better to predict that the changes in independent variables (psychological, social and economic) also made the same unit change in dependent variable (dialysis). The correspondence analysis shows the relationship between dialysis and drinking capacity. It reveals that the dialysis is effected the drinking capacity of the patient. The patients of 81:120 are agree and strongly agree that the drinking capacity has influence on dialysis patient.

INTRODUCTION

Dialysis treatment major causes on change of patient's life. The major causes of having dialysis are high blood pressure and high sugar level. Haemodialysis and Peritoneal dialysis are types of dialysis. Dialysis removes of waste food like creatinine and urea, and free blood from the water. While the kidneys are in condition of kidney crash. Haemodialysis is single of three renal alternate therapies (the other two are kidney relocate and peritoneal dialysis). Haemodialysis can be an outpatient or inpatient therapy. Routine haemodialysis is conducted in a dialysis outpatient facility, either a purpose built room in a hospital or a dedicated, stand-alone clinic. Less frequently hemodialysis is done at home. Dialysis treatments in a health center are initiate and manage by specific and expert staff prepared of nurses and technicians. There are many factors which affect the life of dialysis patient the one which is chronic renal failure (headache, weakness, BP etc.). CRF affect signifycally the societal living of patients undergoes haemodialysis. Psychological burden along the dialysis patient impact on quality of life of patient. CRF is a nonstop psychological procedure for patient and their family in array to agree to their latest image and used to with the new situation of haemodialysis. The QOL of patients require dialysis is affect a lot since it is related with the change in their day by day habits and daily life for equally themselves and their family. At the similar era, their material health, practical status, private relations and societal & financial position are significantly affected.

Psychological problem is the major cause for dialysis patients. The patients which are in depression was found that they are three time disobedient with treatment recommendations as no depressed patients. [DiMatoo, 2003]. In this disease most of the things change in the patient like mood disorder which means they shout and angry with their family member without any reason. There hobbies are also change because of weakness, change in personality.

In social, patients affect their self-esteem due to the dialysis. Because they do not have any hope that one day they will get rid of that disease. At the starting time they take too much stress and depression but after some duration they take it as habit because they know it can't end. Patients that deal with dialysis disease having many living style, diet and other restriction according to their sickness. This living style restriction has important contact on social function and with family presentation balance to make sure the protection of protein and other vitamin levels. These restrictions can impact on patient sickness, self-control hopelessness and regulation. [Paul.L kemmel, 1998].

Health is most important factor. Because if a person is healthy can do everything which want to do. Water capacity is most important role in dialysis patient's life. They need to use less consumption of water to drink. If they drink water more that will not release and stay in body. It is because that patient does not breathe normal. They face difficulty to breathe. In summer it is very difficult to control their drinking capacity. But they have to control it. If they don’t that will effect on their health. Patient has anemia disorder because of dialysis. During dialysis waste water separate from blood effect the blood level. That’s why they feel weakness and they need fresh blood to complete the HB level. In every one week they need fresh blood.

Research Objectives:

- Find out that how much psychological problems affect the health of patient.
- Find out how much social problems effect the health of patient.
- Find out that how much economic problems affect the health of patient

METHODS

The study population consist of three hospital of Gujranwala city. The both government and private hospital are our target population. We used purposive sampling in this survey because time was short and also sampling frame is not easily available. By the use of questionnaire, data was collected in 16 days. It was not easy to understand this disease and its scientific terms but it become feasible by consulting the doctor and by internet. Sample of 250 patients was collected 123 males and 127 females.

Instrument of the Study:

Data was collected through well-structured questionnaire. Questionnaire contained 2 sections. The first section contains information about age, gender, Faculty, Semester, Area, Birth order and no. of siblings. Second section of questionnaire contained questions associated to various factors on 5-point Likert scale.

Data Analysis Techniques

For the analysis of data, different statistical software is used such as Statistical packages for social sciences (SPSS), used for the analysis are as following:

- Reliability Analysis
- Descriptive Statistics
- Correlation
- Multiple Logistics Regression

Correlation:

Having dialysis	Sig. value	Correlation variants
Psychological	0.005	0.175**
Social	0.005	0.177**
Economic	0.688	0.26
Health	0.000	0.224**
Religious	0.267	-0.070

As we can see the above table the correlation of dependent variable with all independent variable. We can see that the factor psychological, social and health have significant correlation with dependent variable (having dialysis). The factor economic and religious has insignificant correlation with dependent variable.

Model Fitting Information				
Model	**Model Fitting Criteria**	**Likelihood Ratio Tests**		
	2-Log Likelihood	**Chi-Square**	**df**	**Sig.**
Intercept Only	329.083			
Final	26.206	302.877	214	.000

Likelihood Ratio Tests				
Effect	**Model Fitting Criteria**	**Likelihood Ratio Tests**		
	-2 Log Likelihood of Reduced Model	**Chi-Square**	**df**	**Sig.**
Intercept	26.206[a]	0.000	0	
Health	329.906[b]	303.700	60	.000
Psychological	128.227[c]	102.021	64	.002
Economic	83.769[c]	57.563	40	.036
Social	98.353[c]	72.147	46	.008

The likelihood model is used to test that all the coefficients are zero or not in model. In LR test that one predictors RC is not zero in model. The logistic regression Chi-Square value is finding by 329.083 – 26.206 = 302.877. The null hypothesis is that all of the regression coefficients in the model are equal to zero. The less p-value from test tells us that at least one of the RC in model is not zero.

The LR Chi-Square statistic of psychological find by
-2* health) – (-2* (psychological)) = 329.906 – 128.227=102.02
with 64 degree of freedom.

The LR Chi-Square statistic of economic find by
-2* (psychological) – (-2*(economic)) = 128.227-102.02 =57.563
with 40 degree of freedom.

The LR Chi-Square statistic of social find by
-2 (social) – (-2*(intercept)) = 98.353-26.606
with 46 degree of freedom.

We can see the sig column all the values are less than 0.05 tells that at least one of the RC in the model is not zero.

Classification				
Observed	**Predicted**			
	1-40	**41-80**	**81-120**	**Percent Correct**
1-40	183	1	0	99.5%
41-80	2	56	1	94.9%
81-120	0	0	7	100.0%
Overall Percentage	74.0%	22.8%	3.2%	98.4%

The classification table for having dialysis in months. This table shows that 183 cases out of 184 are correctly specified in category of “1-40”. With 99.5% percentage while there are 56 cases out of 59 are correctly specified in category "41-80" with percentage of 94.9% and there are 7 cases out of 7 are correctly specified in category " 81-120" with percentage of 100% overall percentage of the correctly classification for model is 98.4% it means model is best fit for prediction.

RECOMMENDATIONS

The reason of this research is to find out that how much psychological, economic and social problems effect on dialysis patient. It is useful for general public. Four analysis technique used in this study first one is descriptive statistics, second one is correlation and the third is multiple logistic regression and forth one is correspondence analysis.

Correlation the highest correlation is between psychological and age of patient. The effect the psychological level of patient, health of patient and social effect are highly correlated with dependent variable (having dialysis).

Multinomial logistic regression analysis. According to the results of analysis having dialysis (months) depend on the psychological, economic and public phase of patient. As one unit change in psychological, financial and public phase of patient create 1 unit change in the dependent variable.

REFERENCES

1. Gerogianni, S., Babatsikou, F., Gerogianni, G., Koutis, C., Panagiotou, M. and Psimenou, E. (2016). Social life of patients undergoing haemodialysis. *Int. J. Caring Sci.*, 9, 122-134.
2. Stavroula, G. (2014). Psychological aspects in chronic renal failure. *Health Science Journal*, 8(4), 205-214.
3. Vafaei, A.A. and Nobahar, M. (2017). The care preferences of patients under hemodialysis. *Journal of Renal Injury Prevention*, 6(3), 210-215.
4. Kimmel, P.L. (2001). Psychosocial factors in dialysis patients. *Kidney International*, 59(4), 1599-1613.
5. Johnson, J.P., McCauley, C.R. and Copley, J.B. (1982). The quality of life of Hemodialysis and transplant patients. *Kidney International*, 22(3), 286-291.
6. Sullivan, J.D. (2010). End stage renal disease economics and the balance of treatment modalities. *Journal of Service Science and Management*, 3(01), 45-50.
7. Walker, R.C., Howard, K., Tong, A., Palmer, S.C., Marshall, M.R. and Morton, R.L. (2016). The economic considerations of patients and caregivers in choice of dialysis modality. *Hemodialysis International*, 20(4), 634-642.
8. Stillman, M.A. and Gonzalez, E.A. (2009). The incidence of pica in a sample of dialysis patients. *International Journal of Psychology and Counselling*, 1(5), 066-093.
9. Frazão, C.M.F.D.Q., Sá, J.D.D., Medeiros, A.B.D.A., Fernandes, M.I.D.C.D., Lira, A.L.B.D.C. and Lopes, M.V.D.O. (2014). The adaptation problems of patients undergoing hemodialysis: socio-economic and clinical aspects. *Revista latino-americana de enfermagem*, 22(6), 966-972.
10. Ginieri-Coccossis, M., Theofilou, P., Synodinou, C., Tomaras, V. and Soldatos, C. (2008). Quality of life, mental health and health beliefs in haemodialysis and peritoneal dialysis patients: investigating differences in early and later years of current treatment. *BMC nephrology*, 9(1), 14.
11. Song, K.K., Zhao, D.L., Wang, Y.D., Wang, Y., Sun, X.F., Miao, L.N., Ni, Z.H., Lin, H.L., Liu, F.Y., Li, Y. and He, Y.N. (2017). Analysis of factors associated with death in maintenance hemodialysis patients: A multicenter study in China. *Chinese Medical Journal*, 130(8), 885-891.
12. Tong, J., Liu, M., Li, H., Luo, Z., Zhong, X., Huang, J., Liu, R., He, F. and Fu, J. (2016). Mortality and associated risk factors in dialysis patients with cardiovascular disease. *Kidney and Blood Pressure Research*, 41(4), 479-487.
13. Muehrer, R.J., Schatell, D., Witten, B., Gangnon, R., Becker, B.N. and Hofmann, R.M. (2011). Factors affecting employment at initiation of dialysis. *Clinical Journal of the American Society of Nephrology*, 6(3), 489-496.
14. Fischer, M.J., Porter, A.C. and Lash, J.P. (2013). Treatment of depression and poor mental health among patients receiving maintenance dialysis. *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 61(5), 694-697.
15. Dziubek, W., Kowalska, J., Kusztal, M., Rogowski, Ł., Gołębiowski, T., Nikifur, M. and Woźniewski, M. (2016). The level of anxiety and depression in dialysis patients undertaking regular physical exercise training-a preliminary study. *Kidney and Blood Pressure Research*, 41(1), 86-98.

16. López Revuelta, K., García López, F.J., de Alvaro Moreno, F. and Alonso, J. (2004). Perceived mental health at the start of dialysis as a predictor of morbidity and mortality in patients with end-stage renal disease (CALVIDIA Study). *Nephrology Dialysis Transplantation*, 19(9), 2347-2353.
17. Kessler, R.C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K.R. and Wang, P.S. (2003). The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Jama*, 289(23), 3095-3105.
18. DiMatteo, M.R., Lepper, H.S. anad Croghan, T.W. (2000). Depression is a risk factor for noncompliance with medical treatment: meta-analysis of the effects of anxiety and depression on patient adherence. *Archives of Internal Medicine,* 160*(*14), 2101-2107.
19. Kimmel, P.L., Peterson, R.A., Weihs, K.L., Simmens, S.J., Alleyne, S., Cruz, I. AND Veis, J.H. (1998). Psychosocial factors, behavioral compliance and survival in urban hemodialysis patients. *Kidney International*, 54(1), 245-254.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 389-392

RELATIONSHIP BETWEEN BIRTH ORDER AND PERSONALITY

Kinza Ijaz, Erum Shahzadi[§] and **Ayesha Mansoor**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: erum.shahzadi@uog.edu.pk

ABSTRACT

The birth order of a child in a family has a great influence on his career choices, commitment and self-respect in the individual's life extent. First born children and so students, seek more attention, care and love from parents as compare to the later born. This research focuses on factors like Social Attitude, Family Relationship, Self-esteem, Intellectual Ability, Life satisfaction, Mental Health and Impact of Birth Order. The main objective of this research is to assess the relation between birth order and personality traits of students at University of Gujrat. A cross sectional study is conducted at University of Gujrat (Hafiz Hayat Campus). Data is collected through well-structured questionnaire from 300 students, questionnaire consists of two sections and total 49 items. Population of the students is selected through stratified random sampling with proportional allocation.

Results of this thesis conclude that all the personality traits are highly correlated to each other. In the model of first born respondents the personality traits are relate to each other such as family relation has an impact of the achievements of first born. Also their mental health is related to their academic performance. Self-esteem has a relation with life satisfaction which means high self- esteem leads to satisfaction with life. The model of middle born shows that the personality traits are associated to each other such as their family relation affects their self -esteem, Also their self-esteem is connected with their life satisfaction and their life satisfaction has an impact on their academic performance. The model of last born respondents indicate that their personality traits are connected to each other such that their mental health influences their life satisfaction.

1. INTRODUCTION

People are attracted by the aspect that children of a family act differently although they were raised in the similar environment, such as neighborhood, and share the same genetic pools from both of their parents. On top of behaviors, siblings vary in terms of personality characteristics such as intelligence, domestic sentiment and others. First born are always defined to be responsible, hard workers and perfectionists whereas last born and only child are always called as the dear of the house and are frequently spoiled kids. As a result, these changes between siblings have involved the courtesy of researchers over the past decades. Generally, parents are excited and expected about their first child. They tend to be very caring and pay more consideration as well as investment on this child. However, their attention and investment fluctuate across children. As the second child comes in the family, the first born may experience less attention and the same occurs to the second born once a third child comes and so on. Also, it is suggested that parental resources that a child received decrease as the sib ship size grows larger.

Birth order and family style moves many development difficulties in children. Studies have revealed that firstborn and the last-born children in the family are subject of unusual impacts and are more weak to behavioral harms. Parent's expectancy from first child is mostly high and when the younger child is born, they expect the older child to be more lenient, kind and supportive. It results in pressure and annoyed feelings. The middle children marked extreme depending behavior than the first born and the later born child and seek more mature help and support than other children and are generally more talkative and express more negative affect than others.

Intellectual Ability and Achievement defined as the academic performance and achievement of participants. There are empirical research findings that demonstrated that birth order effect on academic achievement. Nuttall et al. reported that firstborn girls had better academic performance than the later-born girls. Every child born in a family has own personality and hence self-esteem dissimilar from other siblings. This is associated with Adler's birth order theory which retains that firstborns have the highest self-esteem, loyal and tolerate by the rules of a place unlike the later born siblings who tend to be less dedicated, fail to follow instructions and morals of a place or career and are rebels [Adler, 1964]. The severe cases of bribery, language misuse and low self-esteem occur due to low attention given to student's sib ship relations and the birth order effect. These singularities may have made the health and careers have insufficient and unprincipled facility of services [Nhandi, 2017].

2. MATERIAL AND METHODS

The study population consists of all those students who are studying at University of Gujrat (Hafiz Hayat Campus). Students studying 2nd, 4th, 6th and 8th semester of BS program are our target population. Sample of students is selected using stratified random sampling with proportional allocation. The study is cross sectional study; data is collected from students who are studying at University of Gujrat. Sample of 300 students is collected, from which 30 belong to Faculty of Arts, 53 belong to Faculty of Computing and Information Technology, 14 belong to Faculty of Engineering and Technology, 30 belong to Faculty of Management and Administrative Sciences, 114 belong to Faculty of Sciences, 38 belong to Faculty of Social Sciences and 21 belong to Faculty of Design and Architecture.

Instrument of the Study

Data was collected through well-structured questionnaire. Questionnaire contained 2 sections. The first section contains information about age, gender, Faculty, Semester, Area, Birth order and no. of siblings. Second section of questionnaire contained questions associated to various factors on 5-point Likert scale.

Data Analysis Techniques

For the analysis of data, different statistical software are used such as Statistical packages for social sciences (SPSS), Statistica, AMOS. Statistical techniques used for the analysis are as following:

- Reliability Analysis
- Descriptive Statistics
- Bivariate Analysis
- Confirmatory Factor Analysis (CFA)
- Structure Equation Modeling (SEM).

Correlation results of Family Relationship with Other Factors

Family Relationship	Sig. value	Correlation variants
Social Attitude	.000	.276**
Self esteem	.000	.423**
Intellectual ability and achievement	.000	.359**
Satisfaction with life	.000	.443**
Mental Health	.001	.199**
Impact of Birth Order	.000	.230**

Family Relationship is correlated with Social Attitude with the value of .276, which means that there positive relationship exists because these factors tend to increase together such as good social attitude is correlated with good family relationship. Self-esteem is correlated with Family Relationship with the value of .423, which means that there is positive relationship exists because these factors tend to increase together such as good Family Relationship is correlated with higher self-esteem.

Intellectual ability is correlated with Family Relationship with the value of .359 which means that there is moderate positive relationship exists because these factors tend to increase together such as good Family Relationship is correlated with Intellectual ability.

Satisfaction with life is correlated with Family Relationship with the value of .443, which means that there is moderate positive relationship exists because these factors tend to increase together such as good Family Relationship is correlated with Satisfaction with life. Mental Health is also correlated with Family Relationship. Also the impact of birth order is related to Family Relationship. Hence all the factors are related to Family Relationship.

Measure of Goodness of Fit of Model for confirmatory Factor Analysis

Factor/Variable	$\chi2$	d.f	p-value	$\chi2$/ d.f	GFI	AGFI	RMSEA
Social Attitude	26.813	5.00	0.000	5.3	0.966	0.897	0.120
Family Relationship	113.89	14.0	0.000	8.1	0.891	0.782	0.165
Self Esteem	46.067	14	0.0027	3.09	0.976	0.943	0.071
Intellectual Ability and Achievement	23.188	9	0.005	2.5	0.976	0.943	0.071
Satisfaction with life	50.94	9	0.000	5.6	0.942	0.865	0.131
Mental Health	73.91	5	0.000	14.7	0.974	0.939	0.075
Impact of Birth Order	73.92	5.00	0.000	14.7	0.904	0.712	0.123
Recommended				≤ 3	$\geq .90$	$\geq .90$	≤ 0.08

RECOMMENDATIONS

This study cited some suggestions which would be beneficial for students for their better personality development.

1. Results of Bivariate analysis suggest that students should not take much stress as it affects their social attitude and interaction with people. Also their self-esteem is affected by their mental illness. And mental health has a great influence on academic achievement.
2. Result of SEM suggest that the personality traits are obviously different according to their birth positions.

REFERENCES

1. Al-Khayat, M.M. and Al-Adwan, F.E.Z. (2016). The Effect of Birth Order on Personality Traits and Academic Performance at Sample of families in Jordan. *European Journal of Social Sciences*, 52(2), 151-161.
2. Chandola, R. and Tiwari, S.C. (2016). A Study of Birth Order and Mental Health. *International Journal of Advanced Research and Review,* 1(10), 26-30.
3. Gupta, T. (2017). Birth Order and Personality. *The International Journal of Indian Psychology*, 5(1), 120-125.
4. Ha, T.S. and Tam, C.L. (2011). A Study of Birth Order, Academic Performance and Personality. *Journal of Social Science and Humanity,* 5(1), 18-24.
5. Julia, M.R., Egloff, B. and Stefan, C.S. (2015). Examining the effects of birth order on Personality. *International Journal of Psychological and Cognitive sciences*, 112(46), 14224-14229.
6. McGowan, H. and Beck, E. (2009). A Qualitative Investigation of Middle siblings. *Journal of student scholarship,* 11(1), 1-8.
7. Nhandi, D. (2017). Sibling’s Birth Order Interaction and Self Esteem Development. *International Journal of Education and Research*, 5(1), 51-68.
8. Soysal, F.S.O. (2016). A study on sibling relationships, life satisfaction and loneliness level of Adolescents. *Journal of Education and Training Studies*, 4(4), 58-67.
9. Tyagi, P. (2017). To assess the relation between parenting style of parents and birth order in personality of adolescents. *International Journal of Applied Home Science*, 4(7), 513-518.
10. Zyrianova, N.M., Chertkova, Y.D. and Pankratova, A.A. (2013). The influence of Birth Order and family size on the relationships between Cognitive Abilities and Personality traits. *Journal of Russian Psychological Society*, 86(1), 262-266.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 393-400

EFFECT OF MOTHER NUTRITIONAL KNOWLEDGE ON CHILD NUTRITION

Sana Abdul Ghafoor and **Erum Shahzadi**[§]
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: erum.shahzadi@uog.edu.pk

ABSTRACT

The objective of the study is to investigate the factors that effecting the child health and mother role in child nutrition. Child nutrition become major issue of developing countries. Family health is an essential matter. Mother plays a vital role for her family. She bears the responsibility of taking care of her children's health needs and caring. The purpose of this study to evaluate the factors that effects the child health and how mother health also effects the child health at early stage of life. So that people may get awareness about child health and gave complete attention to their children at first 3 year of child life. Sample of 350 mothers was taken from two Tehsil (Sialkot and Sambrial) using cluster sampling. Statistical techniques like descriptive statistics, confirmatory factor analysis, neural network chi- square test and binary logistic regression are used for the analysis of data. The results shows that education of mother, mother health, hygienic, medication, healthy food and medication somehow effects the child health and conclude that education of mother and hygienic has more effect on child nutrition and health. Mother health and medication also has significant effect on child health. The results also shows that age of mother, education of mother, number of children, mother health, hygienic, medication, healthy food and pregnancy complication are most significant risk factor of children health at first 3 year of child life.

1. INTRODUCTION

Family health is a basic issue. Mother assumes significant job concerning giving health care to her family. She is in charge of dealing with her kids' health needs and furthermore thinking about more seasoned grown-ups and there family. She additionally in charge of their nourishments, hygienic and health requirement. Numerous examinations uncovered that most significant components influencing family health are mother's instruction, age, financial status and frame of mind toward ensured estimation, treatment for illness and how to improve health.

Nutrition has become a main problem of health in many countries, whether in developed countries or in developing countries as well. The occurrence of nutritional problems to infants can actually be avoided if mother has a good knowledge degree on how to manage and serve food for child with balanced nutrition. Lack of knowledge of mothers on balanced nutrition can be a main factor of lack of ability to serve appropriate food according to nutritional needs for family members which certainly causes the nutritional status of children [Al-Shookri, et al. (2015)].

One of the factors that determine the level of health and wellbeing is nutrition. Nutrition is an important factor which plays a role in the human life cycle, especially infants and children who will be our future generation. As to serve daily food should contain balanced nutrition as needed to support optimal growth, to avoid deficiency diseases, to prevent poisoning, and also to prevent the beginning of diseases which might interfere with the survival children [Rao, S. et al. (2011)].

Good health is based on good food. Children that are well fed within first two years of life are healthier in their childhood then the other children who not fed in first two years of life. During six months of baby breast milk is enough for the baby. Milk contains all the nutrients that needed for the healthy growth and also keep child away from the infection Mother also take care of child hygienic. Boiled water is very good for the baby health it kills germs from the water. Before giving milk bottle to baby should sterilize it. Also advise their child to keep their self-clean because cleanliness also keep baby healthy [Ashworth (2002)].

2. MATERIAL AND METHODS

The cross sectional study is used. The study population consist of all those mother whose child age lies between 1-4 years. Data was selected from the well-defined questionnaire. The sample of 350 mothers is selected randomly. A cluster sampling technique was used for the selection of the study population. Sialkot district is randomly selected as a cluster. The cluster is further distributed into four (tehsil) (Sialkot, Sambrial, Pasrur and Bhopalwala) then using simple random sampling Sialkot Tehsil and Sambrial Tehsil is selected for the study. The minimum age of mother is 24 and maximum age of mother is 41 selected in the study.

Data Analysis Techniques

For data analysis, different statistical software are used such as Statistica and Statistical packages for the Social Sciences (SPSS). The statistical techniques that are used for data analysis are Descriptive Statistics, Confirmatory Factor Analysis (CFA), Neural Network, Logistic regression and Cross Tabulation. These analysis was used to explore the mother knowledge about child nutrition and to check that which factor mostly effects the child health.

Descriptive statistics is used for examine the data, to summarize data in dramatic way and also use to show our data in very simple way. Cross tabulation is used between two nominal variables mother that doing job causes pregnancy completion or not. Confirmatory factor analysis is a statistical methodology that is utilized to check the component making of set of found factors Confirmatory factor analysis is to confirm our factors that either our factors are confirm or not.

In neural network conclude all our independent variables (mother health, hygienic, medication, healthy food, pregnancy complication, no formal education, primary, middle, secondary and higher education of mother). And dependent variable in two categories malnutrition/child nutrition (No, Yes). No shows that child not have malnutrition and yes shows that child cause malnutrition. It also used to check the correct classification. In logistic regression it also used to check the correct classification and also used for the prediction. Neural network and logistic regression work same but we used both these technique that which analysis provide us good results.

3. RESULTS AND DISCUSSION

Descriptive Statistics

The table 1 shows descriptive statistics of scale variables. The data involves the age mother, total number of child and age of last child. According to the table we say that minimum age of mother is 24 and maximum age is 41. Average age of mother is 31.71 with standard deviation 3.656. The minimum child of respondent is 1 and maximum child of respondent is 6. Average children of respondent is 3.127 with 1.0041 standard deviation. The minimum age of child is 1 and maximum age of child is 4. The average age of respondent child is 1.79 with standard deviation 0.602.

All these 3 three variables helps us to see the nutritional knowledge of mother. The number of children can tell that how much mother have experience about the child nutrition and health.

Table 1
Descriptive Statistics

Variable	Minimum	Maximum	Mean	Std. Deviation
Age of Mother	24	41	31.71	3.656
Total Number of Child	1	6	3.127	1.0041
age of Last Child	1	4	1.79	.602

Cross Tabulation

In table 2 shows results of cross tabulation between two nominal scale variable. Mother doing job is divided into two categories No and yes also Pregnancy Complication is divided two categories Yes and No. The frequency of no pregnancy complication with not mother doing job is 132 and no pregnancy complication with doing job is 51. Similarly, the frequency for having pregnancy complication for not mother doing job is 132 and having pregnancy complication with mother that doing job is 44. The overall mother that is not doing job have no pregnancy complication 132 out of 350 and mother that are doing job and faced pregnancy complication is 44 out of 350. And using chi-square test it is used to check the association between the nominal variables. The p-value is less than 0.05 we say that there is association between the mother job and pregnancy complication.

Table 2
Cross Tabulation

<table>
<tr><th rowspan="2"></th><th rowspan="2" colspan="2">Pregnancy Complication</th><th colspan="2">Mother doing Job</th><th rowspan="2">Chi-Square</th><th rowspan="2">p-value</th></tr>
<tr><th>No</th><th>Yes</th></tr>
<tr><td rowspan="2">Association b/w M.J and P.C</td><td rowspan="2">Pregnancy Complication</td><td>No</td><td>132</td><td>51</td><td rowspan="2">31.566</td><td rowspan="2">0.049</td></tr>
<tr><td>Yes</td><td>123</td><td>44</td></tr>
</table>

Neural Network

Figure 1 present ROC curve (Receiver Operating Characteristic). ROC curve tells us the information about sensitivity and specificity. It also helps us to check the accuracy of the model. There are two curves with two different colors for each category. The blue curve shows for no category and green curve show for yes category. The mid line that pass from origin is called baseline. The curves more far from the baseline and near to 1 shows the more accuracy of model.

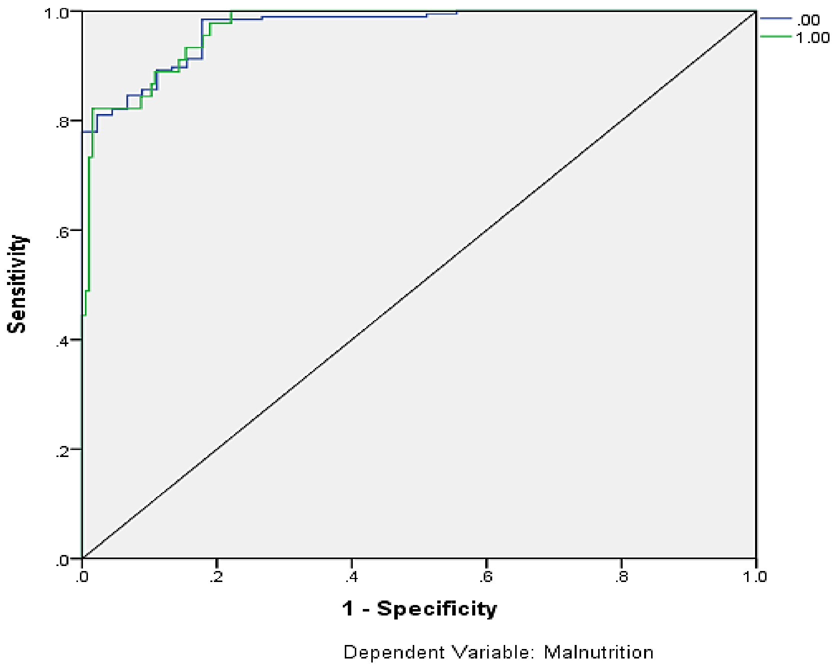


Figure 1:

In table 3 shows the area under the curve, it also shows the accuracy of the model. We say model is best fitted if the value of area under the curve close to 1 then it shows that 100% sensitivity and specificity. The value of area under curve is 0.97 for both yes and no it mean that model is best fitted.

Table 3
Area under the Curve

		Area
Malnutrition	**No**	.970
	Yes	.970

In Figure 2, shows the independent variable importance for the prediction of dependent variable. The child nutrition (malnutrition) have 100% normalized importance.

The model gives the higher importance to education, 2nd higher importance gives to hygienic 3rd, 4th, 5th and 6th higher importance gives to medication, mother health, pregnancy complication and healthy food. The percentage of having importance for education, hygienic, medication, mother health, pregnancy complication and healthy food is 100.0%, 95.4%, 91.1%, 90.4%, 88.9% and 87.4%.

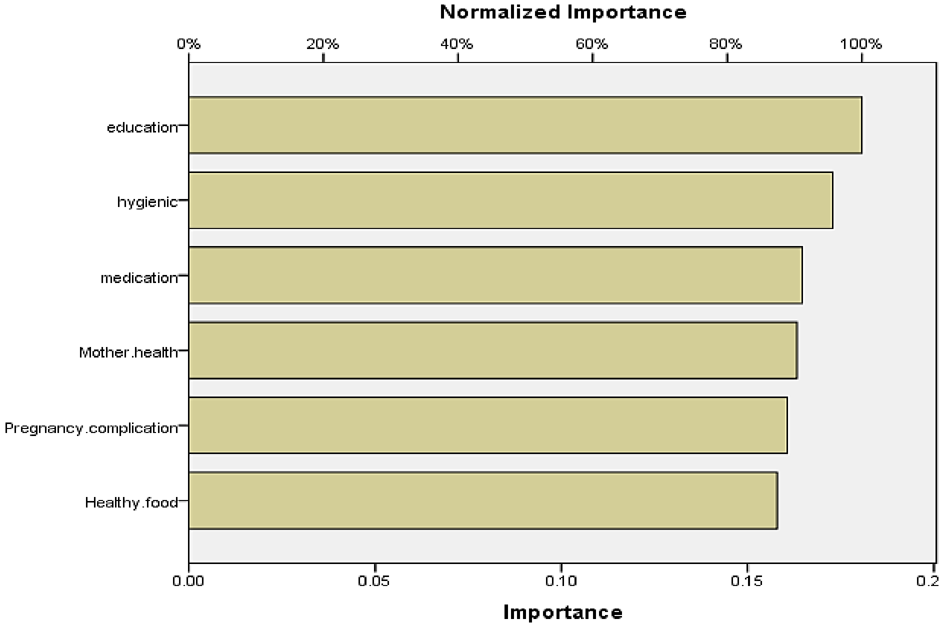


Figure 2:

Logistic Regression

In table 5 shows the classification table that explain the accuracy of model. The variables that are include for this analysis are mother age, total child, child age, education, mother health, hygienic, healthy food, medication and pregnancy complication. This table shows 80.7% decision is correctly specified by this model it means that model is better. There is 86.5% correct decision for no malnutrition and 71.6% for malnutrition.

Table 5
Classification Table

Observed		Predicted		
		Child Nutrition		Percentage Correct
		No Malnutrition	Malnutrition	
Child Nutrition	No Malnutrition	180	28	86.5
	Malnutrition	38	96	71.6
Overall Percentage				80.7

4. CONCLUSION

The results of the study shows that child nutrition is effected by the mother health, hygienic, medication and healthy food. Pregnancy complication not affect the health of children as compared to the other factor. Results also shows that mothers who doing job during pregnancy face more stress then other mothers. Education of mother also plays very important role for child nutrition. If mother is educated then there are chances that child nutrition is good as compared to uneducated mother.

The study also shows that malnutrition effect by food but here another factor like hygienic, medication, mother health also plays most important role that effects on child nutrition. Here another factors age of mother, education of mother, number of children, mother health, hygienic, medication, healthy food and pregnancy complication badly effect the malnutrition. Where number of children determine the knowledge of mother about the child nutrition.

Mass verbal exchange, media, non-governmental organization, coverage makers and numerous other agencies should work to accumulate in nutrition education. It will be beneficial to provide powerful and continual nutritional education and energetic participation should enhanced a higher outcomes.

Nurses must additionally be worried in improving mother knowledge on this advice in order that there will well vitamins and healthful child. And also suggest mother to feed their baby should 6 months and extra properly to feed baby for 2 years.

REFERENCES

1. Alderman, H., Hoddinott, J. and Kinsey, B. (2006). Long term consequences of early childhood malnutrition.*Oxford Economic Papers*, 58(3), 450-474.
2. Al-Shookri, A., Al-Shukaily, L., Hassan, F., Al-Sheraji, S. and Al-Tobi, S.(2011). Effect of Mothers Nutritional Knowledge and Attitudes on Omani Children’s Dietary Intake. *Oman Medical Journal*, 26(4), 253-257.
3. Amsalu, S. and Tigabu, Z. (2008). Risk factors for severe acute malnutrition in children under the age of five, a case-control study. *Ethiopian Journal of Health Development*, 22(1), 21-25.
4. Ashworth, A. (2002). Feeding Young Children. Foot Steps, *KNST Space Journal,* 12(8).
5. Beaver, M., Brewster, J., Jones, P., Keen, A., Neaum, S. and Tallack, J. (2002). Babies and Young Children. *Stanley Thrones limited. Cheltenham Gr. Britain,* 182-183.
6. Berihu, A., Abera, G.G.B., Berhe, H. and Kidanu, K. (2013). Mother's Knowledge on Nutritional Requirement of Infant and Young Child Feeding in Mekelle, Ethiopia, Cross Sectional Study. *Global Journals Inc. (US)*, 13(6), 12-35.
7. Siagian, C.M. and Halisitijayani, M. (2015). Mothers Knowledge on Balanced Nutrition to Nutritional Status of Children in Puskesmas (Public Health Center) In The District of Pancoran, Southern Jakarta 2014. *International Journal of Current Microbiology and Applied Sciences (IJCMAS)*, 4(7), 815-826.

8. Elsabagh, H.M., Atlam, S.A. and Shehab, N.S. (2016). Knowledge, Attitude and Practice Regarding Personal Hygiene among Preschool Children in Tanta city, Gharbia Governorate, Egypt. *Int J Med Res Prof*, 2(2), 255-61.
9. Esteves, T.M.B., Daumas, R.P., Oliveira, M.I.C.D., Andrade, C.A.D.F.D. and Leite, I.C. (2014). Factors associated to breastfeeding in the first hour of life: systematic review. *Rev Saúde Pública*, 48(4), 697-708.
10. Ezzati, M., Lopez, A.D., Rodgers, A., Vander, Hoorn S. and Murray, C.J.L. (2003). The Comparative Risk Assessment Collaborating Group, Selected Major Risk Factors and Global and Regional Burden of Disease. *Lancet*, 360, 1347-60.
11. Bongaarts, J. (2007). Food and Agriculture Organization of the United Nations: The state of food and agriculture: Agricultural trade and poverty: can trade work for the poor? *Population and Development Review*, 33(1), 197-198.
12. Greenaway, E.S., Leon, J. and Baker, D. P.(2012). Understanding The Association Between Maternal Education And Use of Health Servises in Ghana, Exploring the Role of Health Knowlede. *J. Biosoc. Sci.*, 44(6), 733-747.
13. Hobcraft, J., McDonald, J.W. and Rutstein, S.O. (1984). Socioeconomic Factors in Infant and Child Mortality: A Cross-national Comparison. *Population Studies*, 38(2), 193-223.
14. Huynh, G., Huynh, Q.H.N., Nguyen, N.H.T., Do, Q.T. and Van Tran, K. (2019). Malnutrition among 6–59-Month-Old Children at District 2 Hospital, Ho Chi Minh City, Vietnam: Prevalence and Associated Factors. *BioMed Research International*, Article ID 6921312, 8.
15. Khodaee G.H., Moghadam, Z.E. and Saeidi, M. (2015). Healthy Diet in Children., Facts and Keys. *Int. J. Pediatr*, 3(24), 6-2.
16. Levinson, F. and Bassett, L. (2007). Malnutrition is Still a Major Contributor to Child Deaths. *Population Reference Bureau*. https://u.demog.berkeley.edu/~jrw/Biblio/Eprints/PRB/files/Nutrition2007.pdf
17. Lim Z.X., Wong J.L., Lim P.Y. and Soon L.K. (2018). knowledge of nutrition during pregnancy and associated factors among antenatal mothers. *International Journal of Public Health and Clinical Sciences*, 5(1), 2289-7577.
18. Mahgoub, A.T., Elkhalifa, M.Y., Medani, K.A. and Abdalla, W.M. (2014). Mothers Health Awareness and Nutritional Status of Children in Khartoum State-Sudan. *Medical Journal of Islamic World Academy of Sciences*, 22(2), 61-68.
19. Pandey, M.K. (2012). *Maternal Health and Child Mortality in Rural India*. ASARC WP 2009-12.
20. Mariara, K.J., Ndenge, G.K. and Mwabu, D.K. (2009). Determinants of Children's Nutritional Status in Kenya: Evidence from Demographic and Health Surveys. *Journal of African Economies* 18(3), 363-87.
21. Balogun, M.R., Okpalugo, O.A., Ogunyemi, A.O. and Sekoni, A.O. (2017). Knowledge, attitude, and practice of breastfeeding: A comparative study of mothers in urban and rural communities of Lagos, Southwest Nigeria. *Nigerian Medical Journal: Journal of the Nigeria Medical Association*, 58(4), 123-130.
22. Ozdogan, Y., Ucar, A., Akan, L.S., Yılmaz, M.V., Surucuoglu, M.S., Cakıroglu F.P. and Ozcelik, A.O. (2012). Nutritional knowledge of mothers with children aged between 0-24 months. *Journal of Food, Agriculture & Environment*, 10(1), 173-175.

23. Pelletier, D.L., Frongillo, E.A.J., Schroeder, D.G. and Habicht J.P. (1994). A Methodology for Estimating the Contribution of Malnutrition to Child Mortality in Developing Countries. *Journal of Nutrition*, 124(10), 2106S-2122S.
24. Pyle, S.A., Sharkey, J., Yetter, G., Felix, E., Furlong, M.J. and Poston, W.C. (2006). Fighting an epidemic: The role of schools in reducing childhood obesity. *Psychology in the Schools*, 43(3), 361-376.
25. Rao, S., Swathi, P.M., Unnikrishnan, B. and Hegde, A. (2011). Study of complementary feeding practices among mothers of children aged six months to two years-A study from coastal south India. *The Australasian Medical Journal*, 4(5), 252-257.
26. Saaka, M. (2014). Relationship between Mothers' Nutritional Knowledge in Childcare Practices and the Growth of Children Living in Impoverished Rural Communities. *Health Populnutr*, 32(2), 237-248.
27. Siagian, C.M. and Halisitijayani, M. (2015). Mothers Knowledge on Balanced Nutrition to Nutritional Status of Children in Puskesmas (Public Health Center) in the District of Pancoran, Southern Jakarta. *Int. J. Curr. Microbiol. App. Sci.*, 4(7), 815-826.
28. Taleb S., Kaibi M. and Deghboudj N. (2011). Assessment of Nutritional Status of Pregnant Women Attending the City Tebessa PMI (Algeria). *National Journal of Physiology, Pharmacy & Pharmacology*, 1(2), 97-105.
29. UNICEF (2006). *Progress for children, a report card on Nutrition*. UNICEF.
30. Walvekar P.R., Naik V.A., Wantamutte A.S. and Mallapur M.D. (2006). Impact of Child to Child Programme on Knowledge, Attitude Practice Regarding Diarrhoea among Rural School Children. *Indian Journal of Community Medicine*, 31(2), 56-9.
31. World Health Organization (2000). Complementary feeding: Family foods for breastfed Children. *WHO, Geneva*, 43(3), 8-16.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 401-408

CAUSES OF CHILD LABOR IN DISTRICT SIALKOT

Erum Shahzadi[§] and **Asifa Afzal**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: erum.shahzadi@uog.edu.pk

ABSTRACT

Child labor is an important issue in any developing country. It's not only harmful for the children also it damages the moral fabrics of the society. Child labor is problem that is increase day by day. The children who are involve in work in early age lose their mental as well as physical abilities. Due to work children not go to any institute for getting education. This study focus on the factors due to children work in early age like poverty, self-will, parental will, lack of education, over population etc. The main purpose of this study was to identify the causes of child labor at the district Sialkot. A cross sectional study was conducted at the district Sialkot. A well designed questionnaire was used as a tool of data collection. A sample size of 300 children was selected. Purposive sampling technique was used for this research work. The respondents of this study are the children who are involved in any type of work in early age at district Sialkot. Statistical techniques like correlation, cross tabulation, confirmatory factor analysis and structural equation modeling (SEM) was used for the analysis of the data. The findings of this study indicate that the children who belong to rural areas are more involve in work as compare to urban areas. Results indicate that there is significant relationship of child labor with poverty, self-will, parental will, education, overpopulation and family environment. Child labor associated with family status. The relationship between child labor and overpopulation is strong then all other factors.

1. INTRODUCTION

Childhood is the most beautiful period in social life. It is the stage when child is away from any pressures and is the sweetheart of all family members. And ready to learn and discover the new things. But it is only one side of the story. The other side of story is full of pressure and problems. It is perceived that poor families have large number of children, so it is very difficult for them to survive and take over their children on the salary of only one family member. So they send their children in different factories, workshops, even sell different things in streets and make their children their source of money. Many families are incapable to afford the basic needs of their kids, and they send their children in work rather than school. Some families are only depending on the child income [Agarwal, (2017)].

The term child labor states the child involvement in any activity that dispossessed him of his right to learning and childhood, and its harmful for the child emotionally, physically, socially and ethically. Child labor is happening just because of poverty, parents less education, large family size, lack of alertness about rules and health care [Awan, (2011)].

According to UNICEF information it is estimated that 250 million children between 5-14 age are involve in child labor all over the world and this figure is constantly increasing. According to the ILO (International Labor Organization) report 165 million youngsters aged 5-14 years are doing work as a labor. A large ratio of children works for long hours and in unsafe environments [Agarwal, (2017)].

According to ILO (International labor organization) resolution, child work is any labor activity in all teenagers that is done lower than 18 in the labor market or industrial. All youngsters that are involve in work activity damage their schooling activities and also their own carrier. Child work is mostly resolute in Asia and Africa, more than 90 percent child employment in both countries. In Asia the ratio of child worker is more than anywhere else, a large ratio of African children is involving in the labor force. In India 44 million teen-agers work as a labor, and it giving the biggest labor force in the world. In Pakistan 10 percent workers having age between 10 to 14 involve in child labor. In Nigeria 12 million child are include in the labor work. In Brazil 7 million kids effort as a labor. In South America child labor is common [Zaidi, (2013)].

Due to child labor not only children and their families affected, a range of societies and organizations are affected by the child labor. Different firms, trades, labor market, state economy, and transactional trade also affected by the child labor. Many employers and manufacturing companies take use the children and take benefits from them and in return paid less. In the results of child labor many unskilled children are hurt by the child labor because they have no idea about the work, so it is considered that child labor increases the rate of untrained workforce. Due to unskilled workers the nation economy is affected. So, there is need to eradicate the child work because children and economy both are in danger [Richard, (2000)].

Kashif & Hassan (2013) conducted a study to examine the child labor and its solution. Sample size was selected through stratified random sampling from different cities of Pakistan. The questionnaire was based on likert scale and data is collected through email and by hand. The results showed that the child labor have negative effect on the society and it is illegal fact. The results also showed that the poverty, overpopulation, unemployment and parental no awareness are the main reasons of child labor and government play important role to eliminate this worse problem.

Zaidi *et al.* (2013) carried out a study to examine the effect the child labor in Pakistan. Sample size of 700 kids was selected through cross sectional study. The results indicated that majority of the children enforce to work because of shortage and having large dependent ratio of family. The results also indicates that they work more and receive less money in return.

Chanda (2014) conducted a study to determine the effect of child domestic labor on children education. The participant for this study was selected through purposive snowball sampling. The results indicates that the domestic work play negative effect both boys and girl’s education. The children are drop out of schools due to the domestic work.

Ubajaka *et al.* (2010) showed a study to explore the parent’s perception on the sound effects of child labor. A simple random sampling and modified cluster sampling technique was used for the sample selection. The data was collected through the well-defined questionnaire. Results indicates that some parents consider work is good for the children and some parents consider childhood work results harassment, teenage pregnancy and sexually diseases.

Parikh & Sodoulet (2005) carried out a study to investigate the effect of parent's profession on child labor and school presence in Brazil. They conducted a cross sectional study including sample of 28,891 includes 14,613 boys and 14,206 girls from the overall Brazil. Results significantly indicates that the families having large independent ratio prefer their offspring to work. They concluded that the parents who are not employed their children are more likely to work as compare to those parents who are employed.

2. MATERIAL AND METHODS

The sample of children was selected by purposive sampling technique. In this study the target population was children are selected from the industries, workshops and street sellers from the district Sialkot. The study was cross sectional study; information is composed from the children who are involve in child labor at district Sialkot. The sample of 300 children was selected, from which 30 percent involve in industrial work, 39.0 percent involve in workshop and 31.0 percent involve another work. The information was collected using well-defined questionnaire. The questionnaire consists of two sections. The first part consists of demographic material like age, family size, household income, No. of earners in family, region, family status, father occupation and type of work. Second segment of questionnaire consist questions related to various on 5-point Likert scale. Only one factor consists of binary scale.

Data Analysis Technique

Different statistical software was used for the analysis of data such as, Statistica, Amos, Statistical packages for social sciences (SPSS). Statistical technique used for the analysis of data are Descriptive Statistics, Correlation Analysis, Cross Tabulation, Confirmatory Factor Analysis (CFA) and Structural Equation Modeling (SEM).

3. RESULTS AND DISCUSSION

Descriptive Statistics is a term in which discuss the summaries about the data. Table contain Descriptive statistics (minimum, maximum, mean and standard deviation) of the altered measurable variables of respondents of this study like Age, Family size, Household income and No. of earners in family. It specifies that the lowest age of children in this study is 9 and extreme age is 16. The middling age is approximately 13 with 2.010 standard deviation. The minimum size of family is 4 and maximum is 16. The average of family size is close to 9 with the standard deviation 2.355. The minimum household income is 10000 and maximum household income is 38000. The average of household income is 25350 with the standard deviation 6368. 361. The minimum and maximum No. of earners in family is 1 and 4 respectively. The mediocre of No. of earners is round about 2 with 0.798 standard deviation.

Descriptive Statistics

Variables	Minimum	Maximum	Mean	St. Deviation
Age	9	16	13.26	2.010
Family size	4	16	8.32	2.355
Household income	10000	38000	25350.00	6368.361
No. of earners in family	1	4	2.28	0.798

Child labor is connected with family environment with value -.147 which implies that there is negative correlation between child labor and family environment. Means if the environment of family is good than the rate of child work decrease. Child labor is correlated with the overpopulation with .383 value, which indicate that there is positive affiliation exist between them because if population increase the ratio of child labor also increase. There is positive correlation between the poverty and child labor with value .172, it means that there is helpful affiliation between the poverty and child labor, if poverty increase than the child labor also increases. Also, there is positive relationship between remaining factors education, self-will and parental will with values .315, .312 and .204.

Correlation Analysis of Child labor with all other Factors

Child Labor	Child Labor	Family Environment	Poverty	Self will	Parental will	Education	Over Population
Child Labor	1	-.147*	.172**	.312**	.204**	.315**	.383**
		.011	.003	.000	.000	.000	.000

Table show the goodness of fit measures of Confirmatory Factor Analysis. Chi-square is the capacity of the contrasts between the watched co-variances and the co-variances suggested by the model. The choice standard which may be connected when the chi-square will irrelevant and test statistic/d.f <=3 then the model fits the sufficiently. The estimation of GFI, AGFI and CFI near 0.9 demonstrate that model is solid match. RMSEA (Root mean square blunder estimation) demonstrate that when worth is 0.08 than model is great fitted. In factor poverty the estimation of chi-square/d.f is 0.0636 which is under 3 it implies that model is great fitted. The estimation of AGI and AGFI is 0.960 and 0.907 which implies that model is great fitted. The estimation of RMSEA is 0.064 which is under 0.08 it implies that model is great fitted. Generally these indices show that model is great fitted. In Self-will factor chi-square/d.f is 0.0688 which is under 3 it implies that model is great fitted. The estimation of GFI and AGFI is 0.961 and 0.884 individually which implies that model is great fitted. The estimation of RMSEA is 0.069 which demonstrates that model is great fitted. In general indices demonstrate that model is great. In Parental will factor chi-square/d.f is 0.0701 which demonstrates that model is great fitted. The estimation of GFI is 0.954 and AGFI is 0.893 it implies that model is great fitted. The estimation of RMSEA is 0.070 which is under 0.08 it implies that model is great. Generally these indices show that model is great fitted. In the factor Education the chi-square/d.f is 0.064 which is under 3 it implies that model is very much fitted. The estimation of GFI and AGFI is 0.971 and 0.914 it implies that model is excellent fitted. The estimation of RMSEA is 0.064 it implies that model is great fitted. By and large indices demonstrate that model is generally excellent fitted. In the factor overpopulation the estimation of chi-square/d.f is 0.123 methods model is great fitted. The estimation of GFI and AGFI is 0.896 and 0.787 which means model is great. The estimation of RMSEA is 0.023 which is under 0.08 it implies that model is great fitted. In general indices demonstrate that model is fit sufficiently. In the factor child labor the estimation of chi-square/d.f is 0.075 which is under 3 it methods model is great fitted. The estimation of GFI is 0.935 and AGFI is 0.882 which implies that model is great fitted. The estimation of RMSEA is 0.075 which is under 0.08 it implies that model is great fitted. In general all these indices demonstrate that model is great fitted.

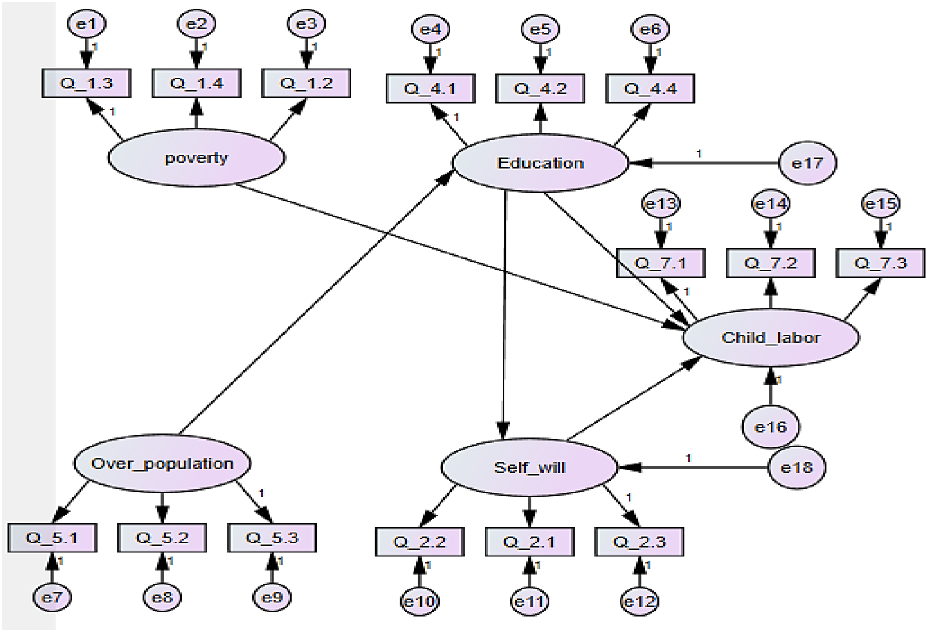
Measure of Goodness of Fit of Model for Confirmatory Analysis

Factor/ Variable	$\chi 2$	d.f	p-value	$\chi 2$/ d.f	GFI	AGFI	RMSEA
Poverty	37.229	9	0.000024	0.0636	0.960	0.907	0.064
Self will	32.9216	5	0.000004	0.0688	0.961	0.884	0.069
Parental will	44.5184	9	0.000001	0.0701	0.954	0.893	0.070
Education	23.3128	5	0.000294	0.064	0.971	0.914	0.064
Over population	85.3458	5	0.000000	0.123	0.896	0.787	0.023
Child labor	87.2623	20	0.000000	0.075	0.935	0.882	0.075
Recommended				≤ 3	$\geq .90$	$\geq .90$	≤ 0.08

The table demonstrate the goodness of fit measures. The end will be taken when the chi-square will insignificant and test statistic/d.f is not exactly or equivalent to 3 at that point model is fit adequately. The p-value of chi-square is 0.000 which is under 0.05 it implies that value is significant it happens in view of enormous sample size. For the most part we get chi-square value is significant so it's anything but a major problem. The value of test statistic/d.f is 1.8030 which is under 3 it implies that model is great fitted. In the event that the value of GFI, AGFI, NFI and CFI near 0.9 it shows that model is great fixed. The value of RMSEA (root mean square estimated approximation) is under 0.08 than model is great fitted. Here the value of GFI is .937 which is bigger than 0.9 it implies that model is generally excellent fitted. The value of AGFI is .911 it implies that model is excellent. The value of RMSEA is .052 which is under 0.08 which mean model is great fitted. In general, we state that model is fitted adequately.

Measure of Goodness of Fit Model

	$\chi 2$	d.f	p-value	$\chi 2$ /d.f	GFI	AGFI	RMSEA
SEM	153.259	85	.000	1.8030	.937	.911	.052
Recommended				≤ 3	$\geq .90$	$\geq .90$	≤ 0.08



4. CONCLUSION

This chapter of study covers debate about the results. Results indicate that the children involve in child labor is more in rural areas as compare to urban areas. Remaining all the factors perform significant role in the study. Also result indicate that there is positive association of child labor with all the factors. There is strong connection among the child labor and overpopulation as compare to poverty, self-will, parental will and education. It is concluded that there is significant association between child labor and family status. Results also conclude that all the factors are related to each other. And model is good fitted. According to the results it shows that education and self-will play significant role in measure the child- labor. It also shows that poverty is directly related to child labor and they play significant role in measure the child-labor. In many previous studies it is reported that poverty, overpopulation, unemployment and lack of parent's education is the main cause of child labor.

Poverty is the main factor of child labor so government needs to increase the life style of common man and fulfill the basic needs. It is must be insured that children spend their childhood in schools. It is the duty of government to give the awareness about overpopulation to people because it is the main reason of child labor, parents should work on family planning according to their resources. It is recommended that the state take steps about this illegal work and punish them who are involve to damage the future of children. As well as state provide the opportunities of employment for parents then the rate of child labor decrease.

REFERENCES

1. Agarwal, S. (2017). Impact of Child Labor on the Nutritional Level and Developmental 9-12 Years. *Journal of Textile Science and Engineering*, 7(3), 1-4.
2. Ahmad, A. (2017). Poverty Education and Child Labor in Aligarh city India. *Journal of Studies on Home and Community Science* 6(3), 165-172.
3. Anker, R. (2000). *Conceptual and research frameworks for the economics of child labour and its elimination*. ILO-IPEC.
4. Awan, M.S., Waqas, M. and Aslam, M.A. (2011). Why do Parents Make their Children Work? Evidence from Multiple Indicator Cluster Survey. *International Journal of Academic Research,* 2(3), 545-549.
5. Bukhari, A. and Fasih, T. (1998). Households' Non-leisure Time Allocation for Children and Determinants of Child Labour in Punjab, Pakistan. *Journal of Pakistan Development Review,* 37(4), 899-918.
6. Chanda, P. (2014). Impact of Child Domestic Labor on Children's Education. A case study of Lusaka city in Zambia. *European Scientific Journal*, 1857-7881.
7. Haider, S.Z. and Qureshi, A. (2016). Are All Children Equal? Causative Factors of Child Labour in Selected Districts of South Punjab, Pakistan. *Journal of New Approaches in Educational Research,* 5(1), 3-10.
8. Johansson, J. (2009). Causes of child labor – A Case Study in Babtai Town, Tanazania. *Journal of Global Development,* 1-43.
9. Kashif, M. and Hussain, M. (2013). A Quantitative Study on Child Labor in Pakistan and its Dynamic Solutions. *Journal of Pakistan Studies,* 5(3), 1-13.
10. Kazmi, S.M.A. (2015). How Does Socio-Economic Factors Force Children into Child Labour? A Case Study of Sahiwal district, Punjab, Pakistan. *Journal of Sustainable Development Policy Institute,* 1-15.
11. Khan, J.M., Sadozai, N.K., Khan, K., Khan, M.F., Khattak, R.M. and Ali, H. (2018). Determinants and Working Conditions of Child Labor: A Case Study of Children Working at Automobiles Workshop at Khyber Pakhtunkhwa (KP), Pakistan. *Journal of Arts and Social Sciences*, 9(2), 2151-6200.
12. Kulsoom, R. (2009). Child labor at District Level: A Case Study of Rawalpindi. *Journal of Munich Personal Repec Archive,* 1-23.
13. Kurosaki, T., ITO, S., Fuwa, N., Kubo, K. and Sawada, Y. (2006). Child labor and School Enrolments in Rural India: whose Education matter. *Journal of Environmental Law*, 17(2), 440-64.
14. Latif, A., Ali, S., Awan, A. and Kataria, J.R. (2016). Socio-Economic and Political Determinants of Child Labor at Brick kilns: A case study of district Jhang. *A Research Journal of South Asian Studies*, 31(1), 161-174.
15. Mazhar, Z.A. (2008). *Child labor in Pakistan: including the excluded Subsidized Education and Subsidy: Tools to Combat the Worst forms of Child Labor*. Master Dissertation, Institute of Social Studies, Graduate School of Development Studies 1-64.
16. Moses, C. (2015). Culture and Child Labor Practices in Ebonyi State Nigeria. *Multi-Disciplinary Journal of Research and Development Perspective,* 4(2), 29-42.

17. Nafees, A.A., Khan, K.S., Fatmi, Z. and Aslam, M. (2012). Situation Analysis of Child Labour in Karachi, Pakistan: a qualitative study, *Journal of Pakistan Medical Association*, 62(10), 1075-1082.
18. Nyamubi, G.J. (2015). The Impact of Child Labour on Primary School Children's Access to and Participation in Basic Education in Tanzania. *International Journal of Learning, Teaching and Educational Research*, 13(2), 26-36.
19. Parikh, A. and Sadoulet, E. (2005). *The Effect of Parents' Occupation on Child Labor and School Attendance in Brazil*. https://ageconsearch.umn.edu/record/25045.
20. Ringera, M.J. (2017). Influence of Child Labor on Access to Education of Children with Disabilities in Samburu Country. Doctoral Dissertation, Kenyatta University.
21. Sughis, M., Nawrot, S.T., Haufroid, V. and Nemery, B. (2012). Adverse Health Effects of Child Labor: high Exposure to Chromium and Oxidative DNA damage in Children Manufacturing Surgical instruments. *Journal of Environmental Health Perspectives* 120(10), 1469-1475.
22. Ubajaka, C., Duru, C.B., Nnebue, C.C., Okwaraoha, O.B. and Ifeadike, G.O. (2010). Parents' Perception on the Effects of Child Labour in a community in Anambra State, Nigeria. *Journal of Afridimedic*, 1(1), 15-19.
23. Zaidi, N., Javed, N. and Khan, M.S. (2013). Child labor in Pakistan - A Tip of An Iceberg. *Journal of Postgraduate Medical Institute*, 27(3), 285-289.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 409-422

PREVALENCE OF MIGRAINE AND ITS EFFECTS ON QUALITY OF LIFE

Sana Muneer and **Erum Shahzadi**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: muneersana07@gmail.com
erum.shahzadi@uog.edu.pk

ABSTRACT

Headache is one of the common health problems that disturb human activities. It can vary based on its duration, frequency, causes, and severity. One of the severe types of headache is migraine. Migraine is not just a headache, instead, it is much more than a headache. Migraine is a disorder of the brain that can occur in many stages. Migraine headache is unilateral or sometimes can be bilateral, and may occur with other symptoms such as nausea, vomiting, speech difficulty, and sensitivity to light, etc. These symptoms not only vary from person to person but also can differ in an individual from one attack to another and affects the quality of life. As by knowing disease's triggers one can get the idea of how to cope up with it. So, it seems to be beneficial to detect triggers and symptoms for the treatment of migraine attacks.

The aim of the study "Prevalence of migraine and its effects on quality of life" was to assess the common symptoms, trigger and coping strategy of migraine and to analyze the effects of migraine on migraineur's quality of life. To meet these objectives responses from the survey questionnaire have been analyzed by applying odds ratio and ordinal logistic regression model.

1. INTRODUCTION

Headaches are a common health problem. Headache can be mild to severe and if severe can affect the quality of life. Headaches can vary on the basis of duration, causes, and severity. International Headache Society (IHS) classified headache disorder into four categories, Migraine, Tension-Type Headache, Cluster Headache and miscellaneous (Ridenour, 1998). According to this study (42%) had a migraine (31%) suffer from tension-type headache, (0.9%) had a cluster headache and (26%) had a combination headache (Nasir & Khalid, n.d.).

The most common type of headache is a tension headache. These forms of headaches usually occur due to the Contraction/ tightening of muscles in the shoulders, neck, scalp, face, and jaw, and includes symptoms such as Dull, steady pain, like a tight band. It can occur in a chronic or episodic form. In tension-headache pain-intensity is mild as compared to migraine. Tension headaches are trigger by stress, fatigue or muscle strain. This type of headache may occur due to activities that put the head and neck in a stressed such as prolonged computer use, reading, gum chewing, etc.

Cluster headache is a disorder frequently seen in men and this form of headache may last from 15 minutes to 3 hours. The pain usually occurs on one side of the head or around the eye. Occur with the Symptom such as tearing, sweating and nasal congestion. It is more common in males than in females (5:1). (Digre, Baggaley, Brennan, & Jeffries, 2011).

Another common type of headache is **Migraine**. Moreover, Migraine is not just a headache, instead, it is much more than a headache. Migraine is a disorder of the brain that can occur in many stages. Migraine with a prevalence of 10-12%, is ranking as 19th disease occurring in Pakistan. Migraine is more common in females than in males (Bhurgri, Buksh, & Qureshi, 2009; Zahid et al., 2014). If the migraineur also suffers from other disorders, then treating that disorder may also help to reduce the migraine attack and its severity (Shehbaz, Ali, Akhtar, & Aziz, 2007; Wencel, 2015).\

1.1 Migraine Type

Migraine can be classified in terms of frequency (numbers of days) into two types: (1) Episodic migraine and (2) Chronic migraine. Episodic migraine occurs on less than 15 days per month whereas chronic migraine has more than 15 days of headache per month. During chronic migraine, patients also suffer from other migraine symptoms for at least 8 days (Wencel, 2015). Furthermore, it can also be classified in terms of the appearance of the disorder named (1) Migraine with aura and (2) Migraine without aura. Migraine pain will always present on one side only (although the side can change between attacks).

1.2 Migraine Stages

As the symptoms are not the same in each patient even these symptoms may change in a migraineur from one attack to another attack, these symptoms usually included five stages (Nasir & Khalid, n.d.; Ridenour, 1998). When migraine attack occurs in migraineur it may include all of these phases or may in some patient only one or two phase occur. These phases may vary from person to person or also can vary from one attack to another. Migraine attack includes five phases: (Prodrome, Aura, Headache Phase, Resolution and Postdrome phase). The prodrome phase and the aura phase of migraine are the first two possible phases of the migraine.

1.3 Causes of Migraine

One theory considers migraine to be originated by problems with blood vessels. This theory is now largely discredited. In recent studies phenomenon known as cortical spreading depression is considered to be the cause of migraine disorder. In migraineurs, during a migraine attack, abnormal electrical activity may occur this activity is called "spreading depression ".Ion channels are like chemical gate that are responsible to control the flow of elements like potassium and sodium in and out of nerve cells, due to the abnormalities of these gates, a migraine headache may occur. Migraine triggers, push them over the edges, which cases electrical disturbance due to which migraine attack occur. unusual visual phenomenon or symptoms of migraine like wavy light, blind spot or complete loss of vision is a result of spreading depression in the visual area of brain. Similarly abnormal cortical brain activity over other region of cortex causes migraine symptoms like numbness, paralysis and speech difficulty (Teixido & Carey, 2014).

1.4 Triggers

"Triggers" are specific factors that may increase the risk of having a migraine attack. Triggers do not cause migraine. Instead, they are thought to activate processes that cause migraine in people who are prone to the condition.

The migraine sufferer has inherited a sensitive nervous system that under certain circumstances can lead to migraine. Migraine may trigger by number of factors/triggers that can vary in another attack. Basically, Triggers may effects “migraine generator” that initiates migraine. There are number of factors that can trigger migraine attacks certain foods, certain odors, skipping a meal, overuse of medications, bright lights/sunlight/ fluorescent lights/ loud noise, Stress /depression. For migraine patients, knowing the triggers (such as certain foods, stress, sounds, sleeping pattern, or bright lights) is helpful to overcome future migraine attacks.

1.5 Significance of the Study

Headache disturbs man’s activity, one of which is migraine headache that also influences daily life activities when a migraine’s patient face it. Migraine occurs due to the number of reasons and has multiple symptoms that are necessary to identify to cope with it. By knowing disease‘s triggers one can get the idea of how to cope with it. Detection of these triggers and symptoms are helpful for the treatment of migraine attack. So, this study is based on the following objectives:

- To determine the symptoms, trigger, coping strategies of migraine
- To examine the intensity of migraine in the presence of any other disorder
- To evaluate the effect of migraine on "quality of life" of the migraineur.

2. MATERIAL AND METHODS

To conduct a study on the topic “prevalence of migraine and its effect on quality of life”. Data of migraine patients were collected from the hospital of district Gujrat (Gujrat and Lalamusa), data was also collected from the student of University of Gujrat that suffer from migraine headaches. As, after a discussion with doctors at the different hospital it was recognized that migraine was common among people, but several patients not visited the hospital for their treatment. So, on the basis of this point data was also collected from the respondent that was accessible but not visit any hospital for their treatment.

2.1 Sampling Technique

For data collection, cross-sectional study design and purposive sampling were used. Data were collected from patient visited Gujrat Hospital and Aziz Bhatti Hospital of city Gujrat, from Abdullah Hospital of city Lalamusa within the duration of 2 months, data were also collected from the student of the University of Gujrat that suffer from migraine disorder. In collecting data from hospitals for a better understanding of the question to migraine suffer, the question was asked in URDU and PUNJABI verbally, so that they can respond accurately.

2.2 Sample Size

A sample of 120 migraine patients was collected from the target population, consist of both males and females, belonging to different age groups, marital status (single or married) and of different occupations and education levels.

2.3 Research Instrument

To conduct a study, a questionnaire was used for data collection, that consists of demographic information such as age, gender, marital status, family income, occupation, and education level. Another part of the questionnaire was based to get information about a number of regular headaches, migraine attacks, and its frequency, symptoms, triggers and coping strategies. Questions about medication effect, weather effect was also asked. The last section of the questionnaire was based to know the effect of migraine on Quality of Life.

2.4 Data Analysis Technique

To analyzing data, different statistical techniques were used that includes:

1. Descriptive Statistics
2. Odds Ratio
3. Ordinal Logistic model

2.4.1 Descriptive Statistics

Descriptive statistics is a way to present, summarize and to interpret basics features of data, or we can say that it is used to describe the data in numbers, which is a more presentable way. Descriptive statistics is a part of the statistics that provide a summary of data in a more realistic form. So, to describe common symptoms, triggers and coping strategies frequency table was formed. Maximum, minimum, mean and standard deviation value was collected of those variable variables that are in scale form.

2.4.2 Odds Ratio

The odds ratio is a statistical technique that is used to measure the strength of association between two categorical variables. Odds ratio can be defined on the basis of odds, basically, odds are the ratio of the probability of success divided by the probability of failure.

2.4.3 Ordinal Logistic Regression Model

When the response variable/dependent variable has more than two categories, and these categories have some order such as category 1 < category 2 < …. category j. response variable depends on predictor variable that can be of any type continuous or categorical, then to model the response variable, ordinal logistic is used.

Category 1 is usually considered to be low level category whereas last category is considered to be high level category. A number of logit function can be used to model one of which is through cumulative probability function of y.

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

Descriptive statistics are used to summarize data and to present it in a more representable form. Most of the variable involved in the questionnaire is in categorical form. So, frequency and percentage are used for qualitative variables. Table A-1 explain demographic information, the total sample consists of 21.7% male, and 78.3% female. Almost 23.3% of the respondent belonging to rural areas and 20% of semi-urban areas.

Table A-2 display percentages about migraine‘s suffers health, headache, effect of medication on migraine headache, migraine effect on life and about migraine occurrence. Table A-3 show the percentages about where the migraine pain located, pain type and other disorder.

Figure (1, 2) illustrates the percentage of Symptoms of Migraine and Aura Symptoms, respectively. While the percentage of migraine triggers and coping strategies has been depicted in Figure-03 and 04 respectively.

Findings suggest that the most common trigger of migraine was found to be loud sound (58.3%), next one is stress (43.3%), about thirty-six percent of the respondents report that their migraine pain is initiated due to too little sleep.

Moreover, About sixty-eight percent of the respondent cope up migraine by taking rest or sleep (67.5%), next one is by using pain killers (53.3%), about thirty-six percent of the respondent cope up to migraine in quiet and darkness(35.8%), about thirty-one percent of the respondent massage on their head(30.8%), by tying something around head (13.3%), by using hot pack on head (8.3%), cold pack(4.2%), warm shower (3.3%) and pacing back-and-forth (1.7%).

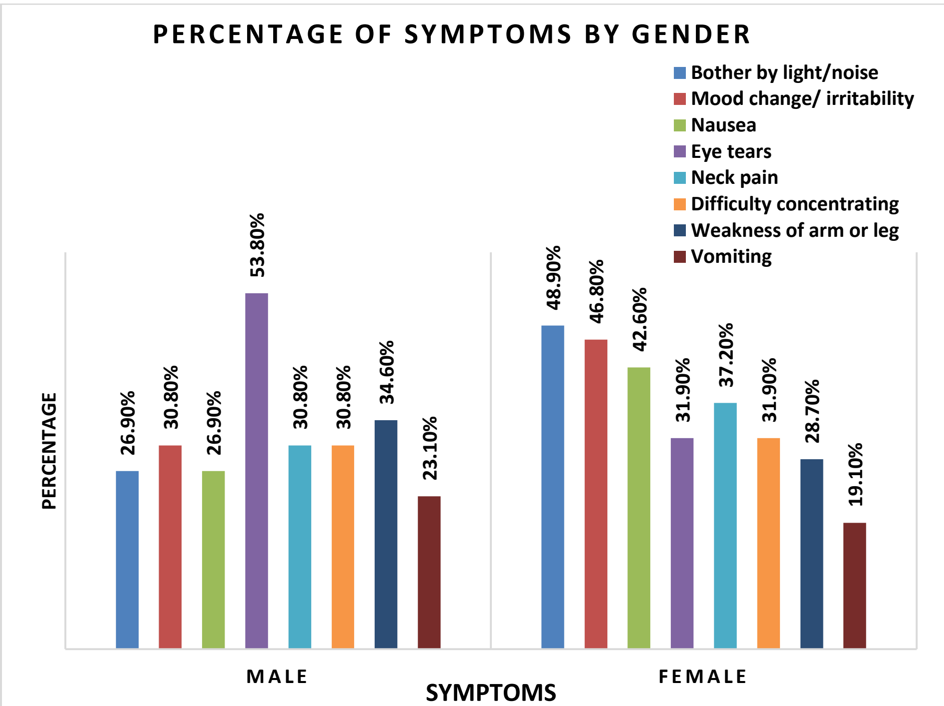


Figure 1: Percentage of Symptoms of Migraine

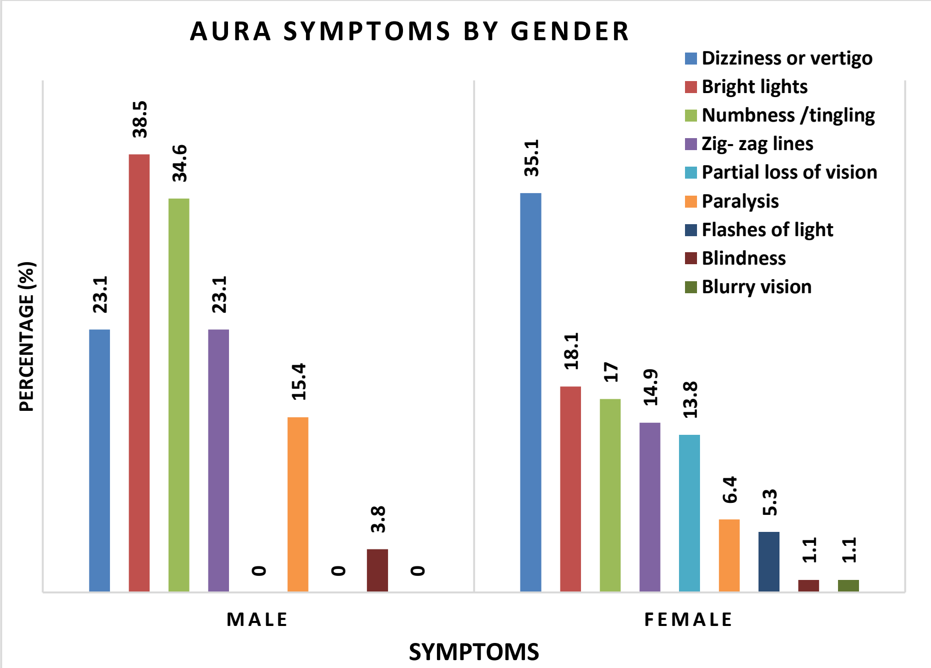


Figure 2: Percentage of Aura Symptom of Migraine

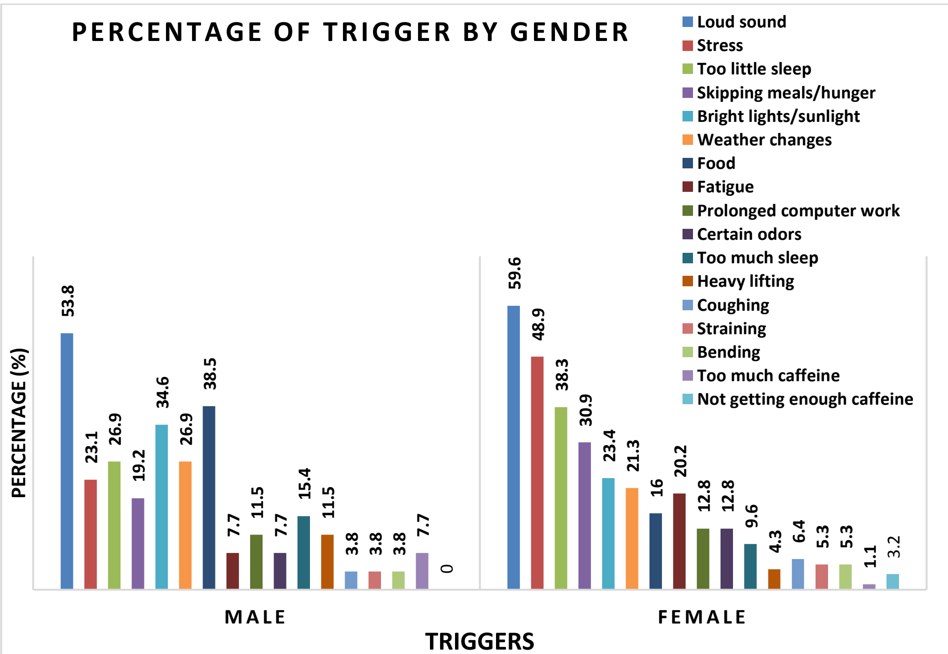


Figure 3: Percentage of Trigger of Migraine

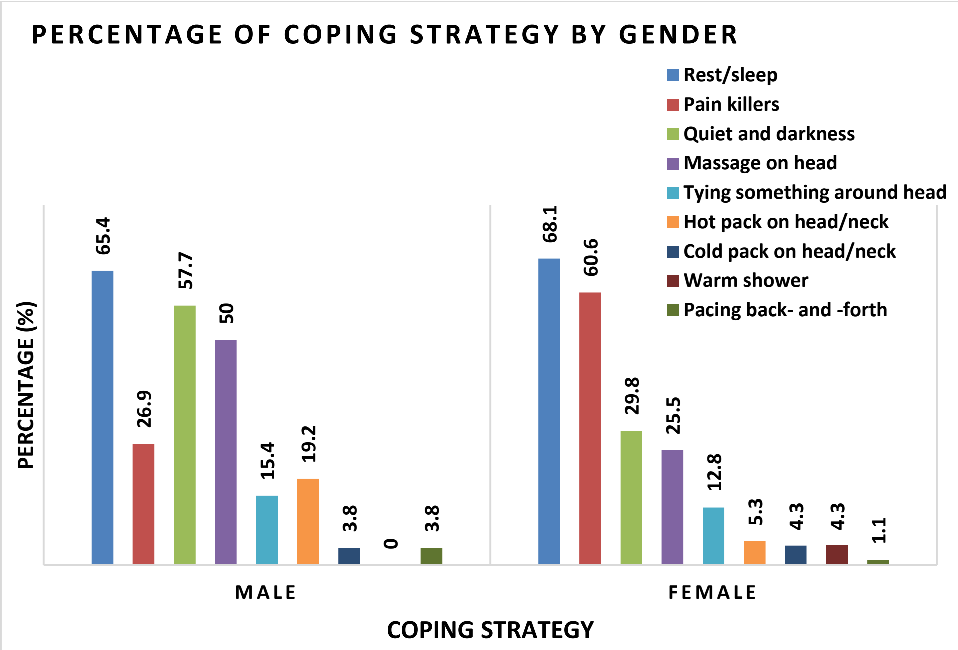


Figure 4: Percentage of Coping Strategy of Migraine

3.2 Odds Ratio

Table A-4 show that, the odd ratio relative risk of a dependent variable "migraine effect on quality of life " with its independent variable such as regular headache, migraine headache timing, attack, depression, prevent work, prevent school, migraine is inherited or not. The estimated odd to have high effect of migraine on quality of life are 2.2857 time as large as in those migraine suffer with depression than those without depression also the estimated odd of low effect are 0.4375 time as large as in migraine with depression than migraineur without depression, whereas the estimated odd of high effect on quality of life are 0.4375 times as large as in migraine suffer without depression than those with depression also the estimated odd of low effect of migraine on quality of life are 2.2857 time as large as in migraineur without depression than those with depression. The high effect of migraine on quality of life is 1.6 time as likely as in those migraineur who are also suffer with depression than those without depression, whereas the relative risk value report that the low effect of migraine is estimated to be 0.7 time as likely for those with depression than without depression.

Estimated odd of high effect of migraine on quality of life is 3.1671 time as large as in those respondents who are suffering more than 24 hours with migraine headache than up to 24 hours, the estimated odd of a High effect are 0.3157 time as large as in patients that have migraine pain up to 24 hours than those who suffer more than 24 hours, the estimated odd of low effect is 3.1671 times as large as in migraine suffer from migraine headache up to 24 hours than more than 24 hours, the estimated odd of low effect is 0.3157 time as large as In those who have migraine headache more than 24 hours than those up to 24 hours. The probability to have a high effect on quality of life are 1.82 time as large for those migraine suffers who also have migraine headache more than 24 hours than those up to 24 hours.

The estimated odd of high effect is 2.169 times as large as in respondents that prevent school due to their migraine than those who not prevent school, also the estimated odd of low effect is 0.4610 times as large as in respondents that prevent school than those who not prevent school. The estimated odd of high effect is 0.4610 times as large as in the respondent that does not prevent school due to migraine headache than those who prevent school.

Estimated odds to have a high effect are 2% time larger in migraine patients that do not have migraine genetically than those who have a positive history of migraine. The estimated odd to have a high effect are 0.9726 time as large as in migraineur that has migraine genetically than those who not have positive history of migraine. The probability to have low effect are 1.01 time as large in those migraine suffers who have positive history of migraine than those who not have migraine genetically.

3.3 Ordinal Logistic Model

Table A-5 show the results of ordinal logistic model, migraine effect on quality of life is used as dependent variable having three categories (low effect, medium effect, and high effect). Whereas the independent variable is number of migraine attack, health rate in last month, prevent work due to migraine, prevent school due to migraine and regular headache and migraine timing. To check the significance of the coefficient of variable wald statistics is used and the variables said to be significant if its p-value is smaller whereas t-value is larger. On the basis of this statement all the variables are significant except regular headache and migraine timing, as this variable is theoretically important. So, we did not exclude this from our model. Model is repeated number of times and that model is preferable that provides good prediction with a minimum value of AIC and residual deviance, the value of AIC and residual deviance is (177.918, 149.918 respectively).

Value of the coefficient of variable (attack 3-5 = 0.2129), (more than 5 attacks = -1.3478), (fair health rate = 2.1786), (good health rate = 2.0190), (excellent health rate = 1.7019), (prevent work = -1.5986), (prevent school = -2.4554), (regular headache more than 10 = -0.0087), (MT3_4 hours = 0.4036), (MT 5_24 hours = 0.4976), (MT several days = -0.8126) and (MT 1 week or longer = 0.5909) and $\beta 1o = -2.500$, $\beta 2o = 2.4247$ the value of the odds ratio is calculated by using the formula $\exp(c\beta r)$, The estimated odd to have low effect of migraine $P(Y=1)$ on quality of life vs. moderate and high effect $(P \geq 1)$ is 5.484 times as large as if a person have excellent health rate than those whose health rate are not excellent holding the other variable constant. The estimated odd to have low effect of migraine $P(Y=1)$ in quality of life vs. moderate and high effect $(P \geq 1)$ is 0.086 times as large as if a person prevent school due to migraine than those who not prevent school due to migraine. Its means that if a person is not able to go to school or their workplace due to migraine then they have large effect of migraine on their quality of life holding the other variable constant.

Table A-7 reveals the classification of the model based on 120 samples, 6 out of 10 cases classified to have a low effect of migraine on quality of life (60%). 80 out of 103 cases classified correctly to have the medium or moderate effect of migraine on their quality of life (77.67%). 6 out of 7 cases classified correctly to have the high effect of migraine on their quality of life. Overall, 92 out of 120 cases classified correctly (76.67%).

4. CONCLUSION, LIMITATION, RECOMMENDATION

4.1 Conclusion

This result shows that migraine usually occurs between the ages to 20-30, migraine is more occur in female than male. Number of study report that migraine is inherited in nature, in our study 39.2% of the respondent have positive family history of migraine. This study also evaluate that medication help to get rid of migraine. As according to this study, by using medication migraine suffer overcome their migraine headache in minutes and hours 35% and 57.5% respectively. On the contrary, without medication in the majority, migraine is prevented in hours or days 47.5% and 48.3% respectively. Migraine also disturb life as 51.7% of the respondent prevent their work due to migraine. Migraine is also affected by weather as 65.8% of the respondent report that their migraine is usually more occur in summer. in the majority of the respondent migraine pain is located behind the eye (50.8%).one of the common symptoms is sensitivity to light and sound (44.2%), among aura symptoms the most common one is Dizziness or vertigo (32.5%). A common trigger is loud sound whereas the most common coping strategy is sleep/rest. Results of the odds ratio show that in females risk of the major attack is more as compared to male (OR=1.6173). Odds ratio also interprets that effect of migraine on quality of life is also associated with depression, migraine timing and prevent school as the value of odd ratio high. A number of migraine attacks per month are also associated with the common triggers of this study. In other models to predict the effect of migraine on quality of life prevent school is a highly significant variable, and this model classifies 76.67% cases correctly.

4.2 Limitation

There is also a need to point out the limitation of the study, first of all, the data was collected only from the two cities. The second one is a smaller sample size as several patients not visit hospitals for their treatment and they do not know the prevalence of migraine. Other limitation of the study is that to check out the association of migraine with other psychiatric disorders, due to limitation of time valid scale to measure psychiatric disorder is not used. Association is only measured through odds ratio. So, further study can be conducted on larger sample size and by using a scale to measure psychiatric disorders such as depression and anxiety and its association with a migraine headache for proper treatment of migraine.

4.3 Recommendation

1. First of all, the descriptive result shows that one of the common triggers is a loud sound. So, patients should avoid sitting in loud sound places. Migraine is also triggered by stress. So, try to keep calm. The sleeping disturbance is one of the causes of migraine, therefore take proper sleep.
2. Result of odds ratio also evaluates that migraine attack is trigger by a loud sound and bright lights. So, use sunglasses to avoid migraine attacks up to some extent.

3. The result of the ordinal logistic model also reveals that by increasing stress, the risk of the number of attacks also increases. Bright lights, loud sounds, skipping meals also have a significant effect on numbers of migraine attacks. So, one can cope up with migraine by eating healthy food, reducing stress, avoiding loud sound and exposure to bright light.

REFERENCES

1. Bhurgri, G.R., Buksh, H. and Qureshi, M.A. (2009). Frequency of depression in migraine patients. *Gomal Journal of Medical Sciences*, 7(2), 106-108.
2. Digre, K.B., Baggaley, S.K., Brennan, K.C. and Jeffries, K. (2011). *Headache: A Patient's Guide*. University of Utah Clinical.
3. Nasir, R. and Khalid, H. (n.d.). *Migraine a Mystery*. (S. Ahmed, S. Noushad, & S.M. Tahir, Eds.). International Journal of Endorsing Health Science Research (IJEHSR) and Advance Educational Institute Research Centre (AEIRC).
4. Ridenour, K. (1998). *Diagnostic Approach for Migraine Headache, with a Comparison of Subcutaneous and Intranasal Treatment Option*. Master Dissertation. Intercollegiate Center for Nursing Education. Washington State University.
5. Shehbaz, N., Ali, S., Akhtar, S.W. and Aziz, H. (2007). Migraine: Comoribidity with Depression. *Pakistan Journal of Medical Sciences*, 23(1), 95-99.
6. Teixido, M. and Carey, J. (2014). Migraine–More than a Headache. *Otolaryngol Head Neck Surg*, 14, 1-14.
7. Wencel, L.W. (2015). *Psychiatric Symptomatology in Migraine Sufferers and the Relationship with Headache- Related Self-Efficacy A*. (Graduation, University of Mississippi Oxford).
8. Zahid, M., Sthanadar, A.A., Kaleem, M., Latif, M., Sthanadar, I.A., Ali, P.A., Sthanadar, I.A., Ismail, M., Imtiaz, N. and Shah, M. (2014). Prevalence and perceptions about Migraine among students and patients in Khyber Pakhtunkhwa province, Pakistan. *Advances in Bioscience and Biotechnology*, 5, 508-516

APPENDIX-A

Table A1
Percentage of Demographic Variables

Gender	**Male**		**Female**		
	21.7		78.3		
Area	**Rural**	**Urban**	**Semi-urban**		
	23.3	56.7	20		
Marital Status	**Single**	**Married**	**Divorced**		
	25.8	73.3	.8		
Occupation	**Student**	**Unemployed or housewife**	**Others**		
	72.5	15.8	11.7		
Age	**10-20**	**21-30**	**31-40**	**41-51**	
	30.8	55.8	9.2	4.2	
Education	**Under middle**	**Middle**	**Metric**	**Inter**	**others**
	5	7.5	10.8	10.8	65.8

Table A2
Percentage of Other Variables

Inherited (positive family history of migraine)	**Yes**				
	39.2				
Frequent on	**Weekdays**	**Weekends**	**Vary**		
	75.8	16.7	7.5		
Headache begin	**Gradually**	**Suddenly**	**Vary**		
	28.3	56.7	15		
Headache last with medication	**Minutes**	**Hours**	**Days**		
	35.0	57.5	7.5		
Without medication	**Minutes**	**Hours**	**Days**		
	4.2	47.5	48.3		
Prevent activities	**School**	**Work**	**Household**		
	22.5	51.7	40.8		
Health rate in last month	**Excellent**	**Good**	**Fair**	**Poor**	
	10.8	33.3	35	20.8	
Migraine affect quality of life	**Not at all**	**Very little**	**Moderately**	**Extremely**	
	5	20	45.8	29.2	
Begin in Day timing	**Morning**	**Afternoon**	**Evening**	**Night**	**Vary**
	19.2	33.3	25.8	6.7	15
Affected by weather	**Spring**	**Summer**	**Fall**	**Winter**	**none**
	2.5	65.8	5.8	11.7	14.2
Doctor	**Neurologist**	**Physician**	**Walk in Clinic**	**Emergency department**	**none**
	30.8	19.2	18.3	2.5	29.2

Table A3
Percentage about Pain Type, Pain Location and other Disorder

		Gender		Total
		Male	Female	
Pain Location	Behind eye(either left or right)	61.5	47.9	50.8
	Temple(either left or right)	61.5	45.7	49.2
	Above eyebrow(either left or right)	50	22.3	28.3
	Back of head(either left or right)	23.1	29.8	28.3
	Forehead	26.9	24.5	25
	Neck	26.9	16	18.3
Pain Type	Pressure	46.2	55.3	53.8
	Stabbing	50	23.4	29.2
	Throbbing	30.8	20.2	22.5
	Dull ache	15.4	7.4	9.2
	Burning	3.8	3.2	3.3
Other Disorder	Depression	34.6	38.3	37.5
	Neck pain	11.5	34.0	29.2
	High blood pressure	19.2	11.7	13.3
	Gastric ulcer	7.7	12.8	11.7
	Skin problem	7.7	10.6	10.0
	Stroke allergies	7.7	6.4	6.7
	Heart disease	.0	4.3	3.3
	Asthma	.0	2.1	1.7
	Coronary disease	.0	1.1	.8

Table A4
Contingency Table, Odds Ratio and Relative Risk of Migraine Effect on Quality of Life with other Variables

		Migraine Effects on Quality of Life		Odds Ratio (OR)	1/OR	Relative Risk(S)	R.R (F)
		High effect	Low effect				
Depression	Yes	24	21	2.2857	0.4375	1.6	0.7
	No	25	50				
Migraine Headache Timing	More than 24	18	11	3.1671	0.3157	1.82	0.58
	Upto 24 hours	31	60				
Regular Headache	More than 10	17	26	0.9194	1.0875	0.95	1.03
	1-10	32	45				
Attack	Major	30	32	1.9243	0.5196	1.48	0.77
	Minor	19	39				
Prevent School	Yes	15	12	2.169	0.4610	1.52	0.7
	No	34	59				
Prevent Work	Yes	26	36	1.099	0.909	1.06	0.96
	No	23	35				
Inherited	Yes	19	28	0.9726	1.0281	0.98	1.01
	No	30	43				

Table A5
Parameter Estimation of Ordinal Logistic Model for Migraine Effects on Quality of Life

	value(η)	**β= -η**	**Std. Error**	**up**	**low**	**Exp (cβ)**	**t value**
Attack 2 3-5 Attack	-0.2129	0.2129	0.5229	1.238	-0.812	1.237	0.40717
Attack 2 more than 5 Attack	1.34780	-1.3478	0.6678	-0.039	-2.657	0.260	2.01851
PW Yes	1.5986	-1.5986	0.4799	-0.658	-2.539	0.202	3.33104
PS Yes	2.4554	-2.4554	0.6205	-1.239	-3.672	0.086	3.95694
Health Rate Fair	-2.1786	2.1786	0.6659	3.484	0.873	8.834	3.27161
Health Rate Good	-2.0190	2.0190	0.6828	3.357	0.681	7.531	2.95704
Health Rate Excellent	-1.7019	1.7019	0.8785	3.424	-0.020	5.484	1.93719
Reg. Headache more than 10	0.0087	-0.0087	0.5201	1.011	-1.028	0.991	0.01668
MT 3-4 hours	-0.4036	0.4036	0.7046	1.785	-0.977	1.497	0.57284
MT 5-24 hours	-0.4976	0.4976	0.7213	1.911	-0.916	1.645	0.68975
MT Several days	0.8126	-0.8126	0.8426	0.839	-2.464	0.444	0.96437
MT 1 Week or Longer	-0.5909	0.5909	0.9912	2.534	-1.353	1.806	0.59600
Intercepts:							
Low-Medium	-2.5001		0.8863				-2.8209
Medium-High	2.4247		0.8765				2.7664

Residual Deviance: **149.918** AIC: **177.918**

Table A6
Likelihood Ratio Test of Ordinal Logistic Model for Migraine Effects on Quality of Life

LR	**Chisq**	**Df**	**Pr(>Chisq)**
Attack 2	5.6995	2	0.0578582
PW	12.1851	1	0.0004817 ***
PS	18.9120	1	1.369e-05 ***
Health Rate	13.1591	3	0.0043049 **
Reg. Headache	0.0003	1	0.9855543
MT	4.6425	4	0.3259820

$$logit\big(P(Y \le j)\big) = \beta_{jO} + 0.219(attack2\ 3_{5attack})$$
$$-1.3478(attack\ 2\ more\ than\ 5\ attack\ \ ) - 1.5986(\text{PWyes})$$
$$-2.4554(\text{PSyes}) + 2.1786(\text{health. RATEfair })$$
$$+2.0190\ (\text{health. RATEgood}) +\ 1.7019(\text{ health. RATEexcellent})$$
$$-\ 0.0087\ (\text{reg. headache more than }10) + 0.4036(\text{MT }3_{\text{4hours}})$$
$$+0.4976(\text{MT }5_{\text{24hours}}) - 0.8126\ (\text{ MTseveral days})$$
$$+0.5909(\text{ MT 1week or longer })$$

where

$$\beta_{1o} = \text{low} \mid \text{medium} = -2.5001$$
$$\beta_{20} = \ \text{medium} \mid \text{high} =\ 2.4247$$

Table A7
Prediction of Migraine Effects on Quality of Life

	Low	**Medium**	**High**	
Low	6	13	0	31.58%
Medium	3	80	1	95.24%
High	1	10	6	35.29%
	60%	77.67%	85.71%	76.67%

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 423-434

EFFECTS OF SOCIAL MESSAGING ADDICTION ON STUDENTS' AT UOG

Mehwish Arshad and **Zaheer Abbas**[§]
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: zaheerabbas@uog.edu.pk

ABSTRACT

The new technologies are increasing day by day. People spend most of their time on smart phones doing social messaging. These technologies are not only good but also harmful for us. Now-a-days people prefer social messaging on smart phone rather than talk about their passions, ideas and information directly. These days, mobile phone is typically a way in which people relate to one another socially through Social messaging. This research aims to examine the five independent factors: Health, Family Relations, Academic Performance, Wastage of time, Behavior. The aim of this study is to examine the effect of Social messaging on health, education, family relation and social relation with society of students and to recognize the significant factors which can differentiate respondents among the low and high Social messaging addiction. For this research a sample of 150 students of university of Gujrat was selected using Simple Random Sampling technique and questionnaire was designed to collect data from students. Reliability of the data was checked before further analysis. Confirmatory Factor Analysis was used to conform the factors. The t-test for two independent samples, Neural network and Anova was used for data analysis purpose.

KEYWORDS

Mobile Phones, Social Messages.

1. INTRODUCTION

Since the last few decades, the continuous evolving technologies have impacted the human life in many ways. According to Ochongor et al. (2012) due to advanced technologies, now-a-days one can communicate in any part of the world within no time.

Vrinda (2010) among all the new inventions, Mobile Phone is the most commonly used device in these days. Mobile companies introduced new features to keep the customers attached.

Evolution in mobile phone features, especially in social messaging have influenced everyone in numerous ways e.g. effects on health, education, family relations, and social relations etc. all over the universe. Young people mostly use social messaging to communicate with each other. People are getting used to social messaging every day. Due to the addiction of using social messaging, people not perform their work properly and it affects their abilities to work accurately. Health and social well-being of people are secretly worsening due to social messaging.

According to Attaa (2010) telecommunication is the fastest growing business and Pakistan is no exception to it. The figures given by Pakistan telecommunication shows that (0.72%) users in 2010 increase to 3.8% users in 2017 which shows significant upward trend. Mobile Phones were not the same as we are using today. Initially, it was like a two-way radios in 1948s. At that time, they were mostly used for communication purpose while transportation and emergencies. The limitations present in these two-way radios were improved with the invention of Landline by AT&T. Which was connected to a base station but still it has deficiencies.

The mobile phones introduced initially were very Ericsson Company. The weight was in kg's which decreased into grams with the passage of time. Second generation phones were considerably of less weight due to the invention of digital circuits.

Power and Horstmanshof (2004) also Taylor and Harper (2001) suggested that with the use of these devices, people get closer and make new relations. In case of urgency, mobile phones provide a sense of security. Tjong et al. (2003) described that mobile phones is a major source of social belongingness. Mobile phones connect the members of a family (Mathew) Similarly, Chen and Katz (2009) suggest that "mobile phone is also useful for persons to distribute their experiences with their family members and providing a passionate and devotional guidance from their families".

2. SIGNIFICANCE AND OBJECTIVE OF STUDY

Despite the increased use of social messaging among university students researchers have pay little or no attention to understanding the effect of social messaging on the students' development. Social messaging use has become a popular and fundamental aspect of the social lives of many students today and hence investigation of its outcomes deserves attention. The recent study is very important particularly for parents, teachers, students, NGOs and policy makers as it provides a very good analysis of bad effects caused by social messaging addiction. By understanding the consequences of recent study parents can guide their children in a better way. Moreover, they will know that it is very important to keep an eye on their children. Teachers will be able to understand that how badly social messaging affects their students' studies and how they can train them to achieve better grades by avoiding this addiction. Students will be themselves able to know that how badly social messaging addiction affects them in different ways, thus they can minimize social messaging usage.

The main objective of the study was to investigate the impact of the social messaging on students' life. The specific objectives of the study were:

1. To examine the types of social messaging platforms commonly used by students.
2. To find out the amount of time students spent on social messaging.
3. To investigate the purposes for which students used social messaging.
4. To examine the perceived impact of social messaging use on students' life.

Factors

Social messaging addiction is determined by different factors:

1. Health
2. Family Relations
3. Academic Performance
4. Wastage of time
5. Behavior

3. METHODOLOGY

This includes the information about target and sampled population, research design, data collection plan, sampling technique, sample size and statistical techniques for the analysis. In this study population was the students which are studying at the present time in University of Gujrat; Hafiz Hayat Campus. In this research the sample of students is selected with simple Random Sampling. Therefore, every unit has an equal probability of being chosen as sample, a simple random sample is meant to be an unbiased representation of population. For this research a sample of 150 students is selected from University of Gujrat by using simple random sampling. In this study, cross sectional study design has been used because only single time information is collected from respondents. For this research survey method for data collection was used. The data was collected by using questionnaire from selected sample. The questionnaire consisted of two portions, first included demographic information and second part contained the information related to Social Messaging addiction and their effects on student's health, family relations, academic performance, wastage of time, behavior. To take respondents opinion five point likert scales were used. The reliability of questionnaire was tested through Cronbach's Alpha. Confirmatory Factor Analysis was used to conform the factors. The t-test for two independent samples, Neural network and Anova was used for data analysis purpose and to investigate the objectives of the study.

4. RESULTS AND DISCUSSION

Reliability of the questionnaire is a key factor for statistical results. Table 1 shows the results of reliability statistic. The Cronbach's Alpha test is used to test the reliability of data. In this study the calculated value of Cronbach's Alpha is 0.785 which is greater than 0.70 shows that the data is acceptable.

Descriptive statistics are considered primarily to explain the basic features of the data. It provides simple summaries about the sample measures. The Table 2 shows the descriptive statistics of nominal scale variables. It shows that 44 males and 106 females. 36% are from joint family system and 64% are from nuclear family system. 46% are rural and 53% are urban. There are 78% from BS Students and 21% were from MS Students. 7% are Fb users, 79% users are of whatsapp, 10% are of Instagram users, 0.7% are twitter users and 2% are Tiktok users.

Table 3 shows the descriptive statistics of quantitative variables. It shows that the mean age is 20.28 years, with standard deviation 2.688 years. The minimum age is 17 years and maximum age is 35 years. The range of age is found to be 18 years. The mean

GPA/CGPA is 1.88, with standard deviation 0.794. The minimum GPA/CGPA is 1.5 and maximum GPA/CGPA is 4. The range of GPA/CGPA is found to be 2. It shows that the mean of Number of messages in a day is 605.56, with standard deviation 1702.544. The minimum Number of messages in a day is 10 and maximum Number of messages in a day is 20000. The range of Number of messages in a day is found to be 19990.

The Kolmogorov Smirnov Test (one sample K-S test) used to test whether the data follow specific distribution. In our study, this test is used to test whether the data follow normal distribution. Table 4 shows normality test of the scale variables "Score of Health, Score of Family Relation, Score of Academic Performance, Score of Time, Score of Behavior". We applied Kolmogorov-smirnov test. The above table reveals the results of the normality test. From the table it can be seen that the p-value of the test statistics for "Score of Family Relation, Score of Academic Performance, Score of Behavior" is respectively (.001, .000, .001) that is less than $\alpha = 0.05$ and the p-value of the test statistics for "Score of Health, Score of Time" is respectively (.060, .200) that is greater than $\alpha = 0.05$. The greater p-value to non-rejection of null hypothesis and the less p-value to reject null hypothesis that the test variable is not normal.

The independent sample t-Test test the significance of the difference between two sample means. Table 5 shows the nonparametric test of independent sample t-test. Previous table shows that the dependent variable Academic Performance score not follows the normal distribution then to move the non-parametric test Mann-Whitney U test used. The statistical significance (p-value) is 0.509 which is greater than 0.05 then we do not reject the null hypothesis. It indicates that the medians of the two groups are not significantly different.

The one-way analysis of variance (ANOVA) is used to determine whether there are any statistically significant differences between the means of two or more independent (unrelated) groups (although you tend to only see it used when there are a minimum of three, rather than two groups). Table 6 shows the previous table shows that the dependent variable Score of Academic Performance not follows the normal distribution then to move the non-parametric test Kruskal Wallis test used. The statistics significance (p-value) is 0.438 which is greater than 0.05 then we do not reject the null hypothesis.

Table 7 shows the CFA analysis. We have to use the CFA to confirm the construct of the study. In CFA models there is no need to independent or dependent variables. Here we have 5 factors in which there are lots of observed variables. The model estimate of confirmatory factor analysis shows that all the variables in all factors are significant, CFA the null hypothesis is that all variables are not confirmed for the factors. The level of significance is less than 0.05, so we reject the null hypothesis and determine that all variables are confirmed for the factors. The variable which has greater parameter estimate value, its contribution is more in factor confirmation. This is our objective to confirm all the factors. According to results of CFA all the underlying variables are confirmed. So, we can say that these confirmed factors are playing important role in determining the effect of addiction of social messaging on student's life.

Table 8 shows the Goodness of Fit measures. Chi-square ($\chi 2$) is the fundamental Goodness of fit measure used in CFA. Recommended value of ($\chi 2$ / d.f) is less than 3. So it also supports our estimated model. Goodness of Fit Index (GFI) is absolute goodness of fit measure. GFI value is .727 and AGFI value is .579, which supports our estimated model. Another measure to assess the goodness of fit of estimated model is Root Mean Square Error of Approximation (RMSEA). Recommended range of RMSEA is less than .08. Typically values are below 0.10 for most acceptable model. In this model, the value of RMSEA is .072 that is less than .08. So RMSEA is supported to fitted model. Almost all goodness of fit measures meets the recommendation level for this factor. So goodness of fit measures supports our estimated model.

We want to predict Addiction of Social messaging by multilayer Perceptron. Here we have dependent variable Gender with two categories. Table 9 classification matrix shows that 13 from 31 cases of the category “male” in the training sample are correctly specified, making 58.1% are correctly classified. There are 57 from 15 cases of the category “female” in the training sample are correctly specified, making 79.2% are correctly classified. Overall 72.8% cases are correctly specified in training sample. In holdout sample 7 from 13 cases of the category “male”are making 46.2% are correctly classified. 28 from 34 cases of the category “female”are making 82.4% are correctly classified. Overall 72.3% are correctly classifiec in testing sample.

Table 10 shows that Area under the curve is another way to check the fitness of the model. Our data under the curve is greater than 0.7 for dependent variable which means that our data is fit in the model. The recommendation criteria’s of area under the curve is close to 1. If the value of area is 0.5 than it means that data is not good for the model. Here it is 0.799 which means that our data is fit for the model.

Graph-1 shows about the sensitivity and specificity is the number of male cases correctly specified and specificity is the number of female cases correctly specified. The green line shows the females and the blue line shows the male cases.

Table 11 shows that Score of Time has 100% most important variable. Score of Behavior, Score of Academic Performance has 89.8% and 89.0% are important variables. Score of Health is 82.9% important variable. Score of Family Relation has 33.7% has important variables.

Graph-2 shows that all independent variables are important to prediction of the model in the bases of dependent variables, because the difference between the independent variables is not wider. Also, the most important factor is Score of Time.

Table 12 shows the parameter estimates. The input layer has 5 independent variables. Output layer uses softmax activation function. Hidden layer is used to assign the weights to input layers to predict the output layer. There are one hidden layer and two output layers.

5. CONCLUSION

People are going to be more addicted of using mobile phone usually social messaging apps. Social messaging has affected a common man in a number of ways such as effect on health, education, family relationships, time and behavior. By wasting more time on doing social messages, have decrease the overall capabilities of students to perform their work efficiently. Social messaging negatively affects physical and psychological health in several ways such as anxiety related to not receiving messages, feeling that nobody loves me when no message is received, sensation of hearing the message tone etc. social messages affects student's academic performance either negatively or positively. Positive in a sense that some use it as important academic messages both at university and at home. Negative in a sense as much of them addicted to social messaging and most of their time is wasted.

REFERENCES

1. Attaa, A. (2010). *Pakistan registers 98 million mobile phone users*. Retrieved from https://propakistani.pk/2010/06/24/pakistan-registers-98-million-mobile-phone-users/
2. Hord, J. (2006). How SMS works. Retrieved from https://computer.howstuffworks.com/e-mail-messaging/sms.htm
3. Isfahani, N.N. (2012). Study and Comparison the Family Climate of SMS Addicted and Normal Adolescent Girls. *J. Life Sci. Biomed*., 3(1), 52-55.
4. Mathews, R. (2004). The psychosocial aspects of mobile phone use amongst adolescents. *InPsych*, 26(6), 16-19.
5. Quain, (2012). *Negative Effects of Text Messaging*.
6. Vrinda (2010). *Texting SMS Affects (Mental) Health*.
7. http://www.avtext.com/the-history-of-sms-messaging.php

APPENDIX

Table 1: Reliability Analysis of all factors

Reliability Statistics	
Cronbach's Alpha	**N of Items**
0.785	23

Table 2
Descriptive Statistics of Nominal Scale Variables

Variables	Valid percent	Frequency
Male	29.3	44
Female	70.7	106
Joint	36.0	54
Nuclear	64.0	96
Rural	46.7	70
Urban	53.3	80
MSc	21.3	32
BS	78.7	118
Facebook	7.3	11
Whatsapp	79.3	119
Instagram	10.0	15
Twitter	0.7	1
Tiktok	2.7	4

Table 3
Descriptive Statistics of Quantitative Variables

Variables	Minimum	Maximum	Mean	S.D	Range
Age	17	35	20.28	2.688	18
GPA/CGPA	1	4	2.29	0.992	3
No-of-msgs	10	20000	605.56	1702.544	19990

Table 4
Normality Test of Data

Tests of Normality			
	Kolmogorov-Smirnov[a]		
	Statistic	**Df**	**Sig.**
Sc_Health	.071	150	.060
Sc_Fam_Rel	.100	150	.001
Sc_Acad_Per	.139	150	.000
Sc_Time	.054	150	.200
Sc_Beh	.102	150	.001

Table 5
Non-Parametric Test

Null Hypothesis	Test	Sig.	Decision
The distribution of Sc_Acad_Per is the same across categories of Gender	Independent Samples Mann-Whitney U Test	0.509	Retain the null hypothesis

Table 6
One Way ANOVA

Test Statistics	
	Sc_Acad_Per
Chi-Square	3.769
Df	4
Asymp. Sig.	.438
a. Kruskal Wallis Test	
b. Grouping Variable: which social app do you use the most?	

Table 7
Model Estimates of Confirmatory Factor Analysis of all Factors

	Estimate	S.E.	C.R.	P-value
Q6 Acad Per ← Academic Performance	1.000			
Q5 Acad Per ← Academic Performance	1.223	.256	4.774	***
Q4 Acad Per ← Academic Performance	1.234	.250	4.941	***
Q3 Acad Per ← Academic Performance	.896	.215	4.159	***
Q2 Acad Per ← Academic Performance	.705	.203	3.483	***
Q1 Acad Per ← Academic Performance	.904	.222	4.079	***
Q1 Fam Rel ← Family Relation	1.000			
Q2 Fam Rel ← Family Relation	1.282	.267	4.799	***
Q3 Fam Rel ← Family Relation	1.212	.266	4.563	***
Q4 Fam Rel ← Family Relation	1.088	.237	4.591	***
Q5 Health ← Health	1.000			
Q4 Health ← Health	1.109	.276	4.017	***
Q3 Health ← Health	1.099	.281	3.906	***
Q2 Health ← Health	.862	.224	3.857	***
Q1 Health ← Health	.690	.194	3.547	***
Q1 Time ← Time	1.000			
Q2 Time ← Time	1.719	.474	3.627	***
Q3 Time ← Time	1.628	.452	3.603	***
Q4 Time ← Time	1.719	.481	3.572	***
Q5 Time ← Time	.932	.334	2.793	.005
Q3 Behavior ← Behavior	1.000			
Q2 Behavior ← Behavior	.875	.207	4.231	***
Q1 Behavior ← Behavior	.838	.207	4.054	***

Table 8
Measures of Goodness of Fit of CFA Model

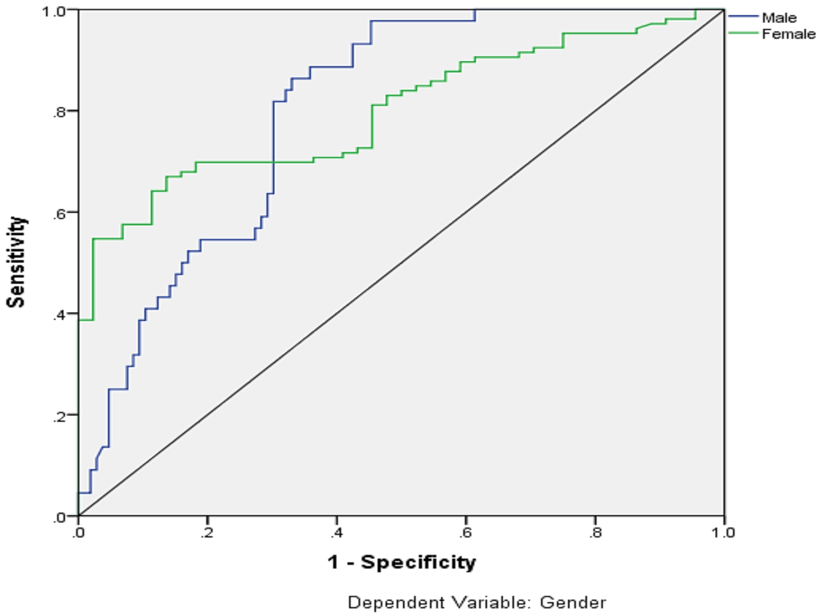
Minimum was achieved	Chi-square =		389.885
	Degrees of freedom =		220
	χ^2/d.f =		1.772
	Probability level =		.000
Model	**GFI**	**AGFI**	**RMSEA**
Default model	.727	.579	.072

Table 9
Classification

Sample	**Observed**	**Predicted**		
		Male	**Female**	**Percent Correct**
Training	**Male**	18	13	58.1%
	Female	15	57	79.2%
	Overall Percent	32.0%	68.0%	72.8%
Testing	**Male**	6	7	46.2%
	Female	6	28	82.4%
	Overall Percent	25.5%	74.5%	72.3%
Dependent Variable: Gender				

Table 10
Area under the Curve

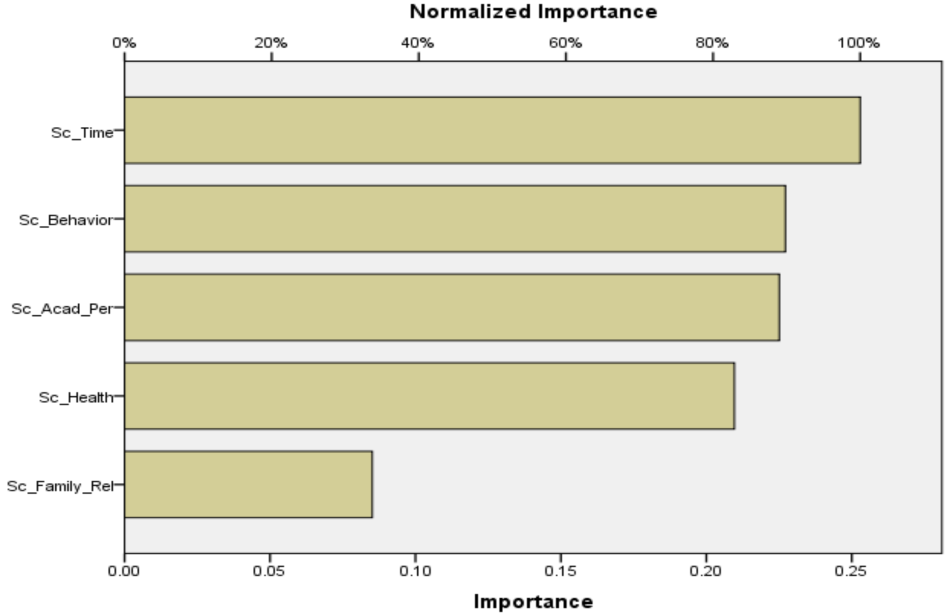
		Area
Gender	**Male**	.799
	Female	.799



Graph-1:

Table 11
Independent Variable Importance

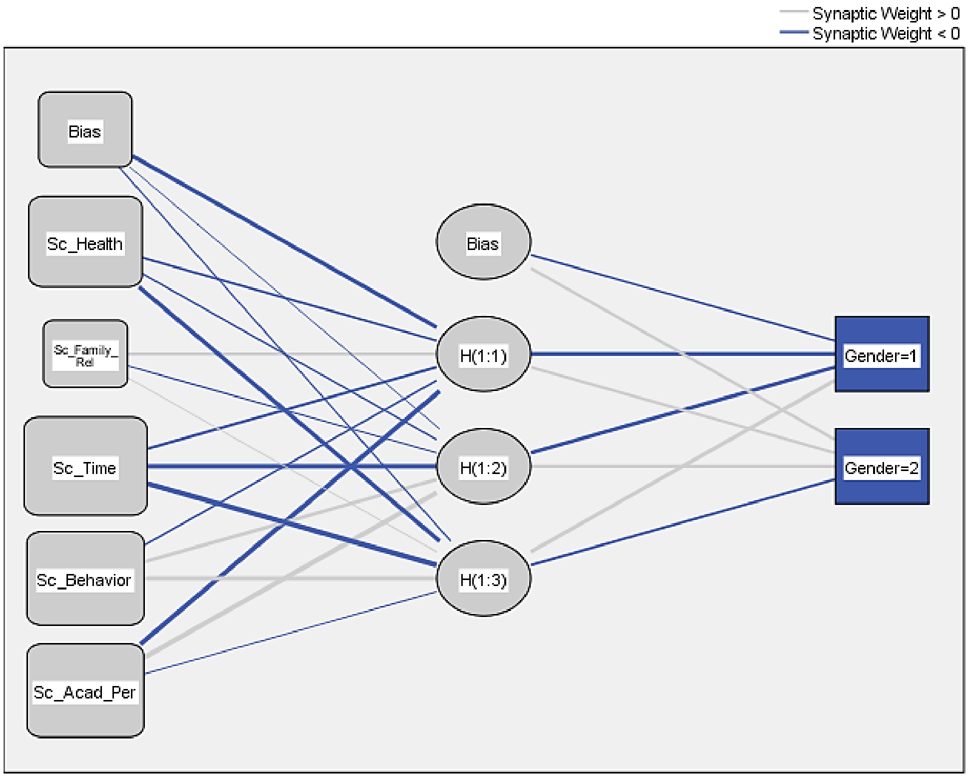
	Importance	Normalized Importance
Sc_Health	.210	82.9%
Sc_Family_Rel	.085	33.7%
Sc_Time	.253	100.0%
Sc_Behavior	.227	89.8%
Sc_Acad_Per	.225	89.0%



Graph-2:

Table 12
Parameter Estimates

Predictor		Predicted				
		Hidden Layer 1			Output Layer	
		H(1:1)	H(1:2)	H(1:3)	[Gender=1]	[Gender=2]
Input Layer	(Bias)	-1.722	-.056	-.539		
	Sc_Health	-.826	-.700	-2.899		
	Sc_Family_Rel	.842	-.289	.122		
	Sc_Time	-.955	-3.847	-4.501		
	Sc_Behavior	-.705	1.432	3.030		
	Sc_Acad_Per	-4.459	4.459	-.238		
Hidden Layer 1	(Bias)				-.823	1.180
	H(1:1)				-1.292	.994
	H(1:2)				-1.520	1.363
	H(1:3)				1.565	-.949



Hidden layer activation function: Hyperbolic tangent

Output layer activation function: Softmax

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 435-444

FACTORS INFLUENCING THE CHOICE OF UNIVERSITY AND FIELD OF STUDY: A CROSS-SECTIONAL STUDY BETWEEN PUBLIC AND PRIVATE SECTOR UNIVERSITIES OF LAHORE

Tahira Bano and **Faiza Mushtaq**
Department of Statistics, Lahore College for Women University
Lahore, Pakistan
Email: tahira.stats@gmail.com
faizamushtaq912@gmail.com

ABSTRACT

This paper sheds light on the factors that affect one of the crucial and decisive process in a student's educational life that is the selection of field of study and university. The objective of this research is to explore and study the effect of various factors including socioeconomic factors, family contextual factors like parent's education, number of siblings and gender. Moreover, the effect of personal interest, scope of field and advice are also part of study. To serve this objective a sample of 399 students, from five public and six private sector universities of Lahore using stratified random sampling, is taken and data is analyzed using both descriptive and inferential techniques. Descriptive analysis includes percentages, bar charts and test of association. The results show that both male and female students used to study professional fields in public institutes. High achievers always prefer to study professional fields no matter in public or private university. Students whose parents are not highly educated but are well established used to study in private universities. Also, test of association showed that scope of field had impact on the selection process. In addition, parents, teachers and sibling's advice also play an important role to opt professional field of study especially in private sector. Among these the most influential are parents. However, for inferential analysis multinomial logistic regression is used. Outcome/Study variable consists of four categories, a combination of two dichotomized variable i.e., university type (private/public) and field of study (non-professional/professional).Predictors include gender, family income, parents' education, interest in field, scope of field, advise and marks in intermediate exams. The major findings of this analysis are

1) Children of ill-bred parents with high family income tend to study professional fields in private universities.
2) There is a significant role of advice and scope of field of study in its selection.
3) The students with good academic record but with low family income prefer to study professional fields in public universities.

KEYWORDS

Field of study, university, multinomial logistic regression, factors affecting, socioeconomic factors.

1. INTRODUCTION

Education is one of the main instruments in developing human resource. It is expected that through education, a sufficient pool of literate, skilled and motivated labor force can be developed. Under this emerging value of higher education, a number of new institutions can be witnessed especially in private sector. In Pakistan's educational system the institutions mainly belong to either public or private sector. Similarly, when it comes to fields of study the two major divisions are professional (science, engineering, management, economics) and non-professional fields (social sciences, music, art etc.).

The selection of university major is a critical moment in ones educational career as it has a long lasting impact on professional future (Galloti, 1999).

Objectives

- To explore the factors influencing the selection of university and field of study.
- To estimate the effect of each factor.
- To determine the most influential factors.

2. DATA AND METHODS

Since, the study mainly focus on selection of field of study at university level so the target population comprised of the students acquiring education in both public and private universities of Lahore. Among all these universities, eleven HEC recognized universities, five from public sector and six from private sector were selected for sampled population. These include Lahore College for Women University (LCWU), Government College University (GCU), Punjab university (PU), University of engineering and Technology (UET), King Edward (KE), University of Lahore (UOL) Lahore Medical and Dental College (LMDC), Lahore University of Management Sciences (LUMS), Foreman Christian College (FC) and University of Management and Technology (UMT).

And the total sampled population, obtained by adding the individual student's strength of each university, was 154,628 students.

The **sample size** for the study was calculated by using Yamane formula at 95% level of confidence and was 399. Since number of students in each university was not equal so, the sample size was proportionally allocated to each selected institute.

The **data collection** was done by means of a questionnaire. It comprised of eleven questions regarding the gender of respondent, financial background, parent's education, intermediate examination marks, family size and factors which are taken into account while selecting the field of study i.e. interest, scope and advice respectively.

The **response variable** was combination of university type and field of study. Both the institution type and field of study were dichotomized as public, private and professional and non-professional respectively. The major fields that belong to professional field category were engineering, economics, computer science, business, law and medical science. The non-professional field category comprised of social sciences, arts, education and life sciences. Hence, the variable of interest was categorized into four categories i.e. non-professional/public, non-professional/private, professional/public and professional/private respectively.

Explanatory variables

- Gender: A dummy variable coded 0 for male and 1 for female.
- Income: 0 for high(above average) and 1 for low(average and below average) whereas the average family income is Rs. 30,000/-
- Parent's Education: 1 for low (both are under –graduate) and 0 otherwise.
- Interest in field: 0 for yes and 1 for no.
- Scope of field: 0 for yes and 1 for no.
- Advice: 0 for yes and 1 for no.
- Marks: Marks obtained in intermediate examination.
- No of sibling: Number of siblings including respondent.
- Advice by: 1 for parents, 2 for siblings and 3 for others.
- Birth order:

For the sake of analysis, the **statistical techniques** used mainly were chi-square test of association (to test whether there exist association between categorical predictors and the response variable) and multinomial logistic regression (since the response variable was a multi-category nominal variable). In addition, data summarization included cross-classification table of total counts and percentages while graphical representation comprised of bar charts.

3. ANALYSIS AND RESULTS

The descriptives of the study, consisted of cross-classification tables and bar charts, revealed that the choice of university and field of study is greatly affected by gender. Males have high percentage in private universities studying professional fields. However, females are unique in their tendency to study professional fields in both private and public universities. Students with low family income use to study both professional and non-professional fields in public universities while those with high family income have the highest percentage for studying lucrative fields in private universities. Students whose parents are not highly educated but are well established used to study in private universities. In selection of universities or field of study high percentage of student had considered the scope of field rather their own interest mattered the most. Advice has a great impact in selection of universities and field of study. However, for the selection of professional and non-professional field parent's advice is the most influential one. Students attaining high marks in intermediate examination used to study in both types of universities i.e. public and private but prefer professional fields.

Table 3 for test of association between response variable(university type and field of study) and predictors showed that birth order and interest in field of study did not show association with the choice of field of study and institution as their p-values are greater than the given level of significance i.e. 0.05. While all other variables (gender, family income, parents education, scope, advice, marks and number of siblings) were significantly associated with the dependent variable.

The **multinomial logistic model** was built using type of Institute and field of study as response variable comprises of four categories, among which non-professional field of study in public institute has been taken as a reference category. Since birth order and interest in field did not possess association with the response variable so these two variables are not included in the model. The explanatory variables included in the model were intermediate marks, number of siblings, gender of student, family income, parent's education, scope of field and advice respectively. As multinomial logistic regression assumes that the explanatory variables should be independent therefore the multicollinearity check was done and it was seen there was no multicollinearity among explanatory variables as VIF was less than 10 and Tolerance was greater than 0.10. So, the assumption of multinomial logistic regression was fulfilled.

The results for **model diagnostics** including, **classification table** which shows that how many cases for dependent variable have been correctly predicted. In the present study overall 52.1% cases are correctly predicted by the model, for **model fitting information** the p-value is 0.000 which indicate that the full model statistically significantly predicts the dependent variable better than the intercept the only model alone, **goodness of fit test** result yield a p-value of 0.101 which is greater than the given level of significance i.e. 0.05 and it indicates that model is adequately fitting the data and that the fitted model is good, **pseudo R-square** yielded value 0.322 which indicate that 32.2 percent of the variation in the response variable is explained by the independent variables included in the model and **likelihood ratio test** indicated that the all the independent variables except parent's education and number of siblings have significant effect on dependent variable and the model fits adequately.

Table 5 shows the factors that affects the dependent variable i.e. the choice of field and institution. The values of B shows relationship between explanatory and response variable whether the explanatory variable have positive or negative effect on dependent variable.

For opting **non-professional field of** study in **private institute** the regression coefficient for marks is approximately equals to 0 and odd ratio is 1, so there is no effect of marks in selecting non-professional field in private university. Number of siblings is also not influencing this selection. Gender has a significant effect also males have 1/0.204 = 4.90 times higher chance of selecting non-professional field than females. Also, the chance of selection for the non-professional / private institute is 1/0.273 = 3.66 times higher for students with high family income as compared to low family income. The odds of not taking an advice is 1.706 times higher than having an advice and the choice of a non-professional field of study and private institute without consideration of scope is 3.976 times higher than considering it. The chance of selection of a non-professional field of study in a private university for a student whose parents are not highly educated is 1.092 times higher than highly educated ones**.**

For choosing **professional study** in **public university** marks have significant effect and the chance of selection of public institute and professional field of study is 1.010 times higher as the marks increases. There is 1.20 times higher chance of selecting a non-professional study in public institute instead of a professional field of study in public institute as the number of siblings increases. Again gender do affect for selection of a public institute and professional study where the chance of males is 1/0.26 = 3.83 times higher than females for opting this type of field of study and institute whereas the chance for a student with low family income for selection of a public institute and professional field of study is 1.264 times higher than belonging to a high family income .The selection of a professional field of study in public institute, taking into account its scope is 1/0.603= 1.658 times higher than not considering it also the chance for this selection 1.103 times higher without having advice. For students whose parents are highly educated, the chance for respective selection is 1.215 times higher than less educated ones.

Finally, in selection of **professional field** and **private university** marks do not affect and there is 0.805 less chance for this kind of selection with the increase in number of siblings. There is 1/0.161 = 6.211 times higher chance of males in opting professional studies than females in private institutes. Family income significantly affects the choice and there is 1/0.197=5.076 time higher chance for opting professional fields and private institutes with high family income than low one. For selection of professional field of study in private institutes, there is 1/0.781 = 1.280 times higher chance of considering its scope. Children with less educated parents have 1.297 times higher chance for private university selection than highly educated parents.

4. COMMENTS AND CONCLUSION

Selection of a university and field of study is a crucial phase in a student's life. As universities not only provide a learning environment but these institutions are also responsible for self- grooming as well as these also provide opportunities to explore hidden abilities in oneself. As far as the selection of university major or field of study is concerned, it is also a most important decision because it led to decide future career.

Many factors affect this crucial decision, which may include cost factors, demographic factors, familial group influences and personal interest as well. With the passage of time financial outcomes, job opportunities and market value of the degree in respective field of study have also become considerable factors in this decision making process.

A number of researches had been conducted in the past to investigate the factors which influence the selection of university and field of study. All of them concluded with almost the same factors which affect this decisive process.

For this study, a sample of 399 students from 11 universities of Lahore has been taken and the analysis is done using multinomial logistic regression. The factors influencing the choice of field of study and institution include gender, family income, parent's education, advice and scope of field. Moreover, marks in intermediate examination and number of siblings (family size) also affect this choice.

The results revealed that the students with high family income and ill-bred parents tend to study professional fields in private universities. Females are unique in their tendency to study in a private institute both professional and non-professional fields of study. Students with good academic record but belonging to low family income category are more likely to study professional fields in public institutes. The factors advice and scope have significant effect on selection decision. Especially, parent's advice has great influence to opt professional fields of study. Scope of field is much preferred upon interest just because professional fields are considered to be a source to reproduce family economic capital. However, descriptives narrate that especially professional fields in private universities, students owe to opt, is because of their interest.

In the end, the selection or choice of field of study and institution is influenced by many factors among which gender, family income, parent's advice and scope are the most significant factors.

5. ACKNOWLEDGEMENT

First of all, I am thankful to **Allah Almighty** and his last Prophet Hazrat Muhammad (PBUH) whose blessings and grace enabled me to complete this study.

I would like to express my special thanks to Mrs. Mahnaz Makhdum, Associate Professor Chairperson, Department of Statistics LCWU, Lahore, who provided us such an opportunity to do this research work.

Words of thanks, surely insufficient and incomplete, are sited for my competent are learned supervisor Mrs. Tahira Bano for their diligent and sincere guidance. Their thought-provoking criticism and encouragement enabled us to complete this work. Her patience, encouragement and experience always serve as the stimulus for me. I am highly indebted to her as this task could not have been possible without her support.

It is my pleasure to acknowledge the assistance provided by all teachers of Statistics Department, LCWU, Lahore. Their cooperative behaviors also really very appreciable.

I want to thank my dearest parents, my loving siblings for their cooperation and throughout encouragement. They have always been supportive and they always motivate me in my studies.

I specially want to thank to my adoring friends for their enthusiastic support and cooperation throughout my research work. Without their help and back-up, it would not have been possible for me to accomplish this research.

Finally, I would like to thank all persons who helped me directly or indirectly in this research work.

REFERENCES

1. Abbasi, M.N. and Sarwat, N. (2014). Factors inducing career choice: Comparative study of five leading professions in Pakistan. *Pakistan Journal of Commerce and Social Sciences (PJCSS)*, 8(3), 830-845.
2. Ahmed, K.A., Sharif, N. and Ahmad, N. (2017). Factors Influencing Students' Career Choices: Empirical Evidence from Business Students. *Journal of Southeast Asian Research*. 2017, Article ID 718849.
3. Aylon, H. and Maccdossi, O. (2018). Economic Achievements of nonacademic parents and patterns of enrollment in higher education of their children: The case of Israel. *High Educ*. 77, 135-153.
4. Kazi, A.S. and Akhlaq, A. (2017). Factors Affecting Student's Career Choice. *Journal of Research and Reflections in Education*, 11(2), 187-196.

5. Chaubey, D.S., Subramanian, R. and Joshi, S. (2011). Factors Influencing Students' Choice of Institutions for Higher Learning: An Empirical Analytical Study. *Indira Management Review*, 5(1), 33-42.
6. Kitsawad, K. (2013). *An investigation of factors affecting high school Students choice of University in Thailand*. Doctoral Dissertation, University of Wollongong.
7. Leppel, K., Williams, M.L. and Waldauer, C. (2001). The impact of parental occupation and socioeconomic status on choice of college major. *Journal of Family and Economic Issues*, 22(4), 373-394.
8. Moakler Jr, M.W. and Kim, M.M. (2014). College major choice in STEM: Revisiting confidence and demographic factors. *The Career Development Quarterly*, 62(2), 128-142.
9. Ryan, Y.J., Rotich, T., Korir, B., Mutai, K., Kosgei, M. and Koech, J. (2010). Factors influencing the choice of college among undergraduate students in public Universities in Kenya. A case study of the University of Eldoret. *Australian Journal of Commerce Study SCIE Journals*.
10. Sabir, I.R., Ahmad, W., Ashraf, U.R. and Ahmad, N. (2013). Factors Affecting University and Course Choice: A Comparison of Undergraduate Engineering and Business Students in Central Punjab. *Pakistan Journal of Basic and Applied Scientific Research*, 3(10), 298-305.
11. Shumba, A. and Naong, M. (2012). Factors Influencing Students' Career Choice and Aspirations in South Africa. *Journal of Social Sciences*, 33(2), 169-178.
12. www.ijhsss.com
13. www.ilmkiduniya.com
14. www.leardstatistics
15. www.petersons.com
16. www.statisticssolution
17. www.wikipedia.org

APPENDIX 1

Response Variable (Field of Study/Type of University)						
Predictors		**Non-Professional/ Public**	**Non-Professional/ Private**	**Professional/ Public**	**Professional/ Private**	**Total**
Gender	**Female**	101	18	64	74	257
	Male	16	13	33	80	142
Family Income	**Low**	57	6	48	22	133
	High	60	25	49	132	266
Parents Education	**Low**	68	21	66	120	275
	High	49	10	31	34	124
Interest in Field	**No**	55	10	55	71	191
	Yes	62	21	42	83	208
Scope of Field	**No**	95	29	67	114	305
	Yes	22	2	30	40	94
Advice	**No**	17	8	18	53	96
	Yes	100	23	79	101	303
Advice by	**No One**	17	7	18	54	96
	Parents	59	11	48	48	166
	Siblings	25	3	14	24	66
	Others	16	10	17	28	71

Response Variable (Field of Study/Type of University) Percentage					
Predictors		**Non-Professional/ Public (%)**	**Non-Professional/ Private (%)**	**Professional/ Public (%)**	**Professional/ Private (%)**
Gender	**Female**	25.3	4.5	16.0	18.5
	Male	4.0	3.3	8.3	20.1
Family Income	**Low**	14.3	1.5	12.0	5.5
	High	15.0	6.3	12.3	33.1
Parents Education	**Low**	17	5.3	16.5	30.1
	High	12.3	2.5	7.8	8.5
Interest in Field	**No**	13.8	2.5	13.8	17.8
	Yes	15.5	5.3	10.5	20.8
Scope of Field	**No**	23.8	7.3	16.8	28.6
	Yes	5.5	0.5	7.5	10.0
Advice	**No**	4.3	2.0	4.5	13.3
	Yes	25.1	5.8	19.8	25.3
Advice by	**No One**	4.3	1.8	4.5	13.6
	Parents	14.8	2.8	12.0	12.0
	Siblings	6.3	0.8	3.5	6.0
	Others	4.0	2.5	4.3	7.0
Marks	**Low**	0.8	0.3	0.3	0.5
	Average	7.8	1.3	2.3	9.8
	High	20.8	6.3	21.8	28.3

Independent Variables	p-value
Gender	0.000
Marks	0.031
No of Siblings	0.027
Parents Education	0.009
Family Income	0.000
Interest in field	0.098
Scope of the field	0.019
Advice	0.001
Birth order	0.434

Likelihood Ratio Tests				
Effect	**Model Fitting Criteria**	**Likelihood Ratio Tests**		
	2-Log Likelihood of Reduced Model	**Chi-Square**	**df**	**Sig.**
Intercept	550.859	0.000	0	-
Marks	564.014	13.155	3	0.004
Number of Siblings	557.964	7.104	3	0.069
Gender	586.291	35.431	3	0.000
Family Income	594.184	43.324	3	0.000
Scope of Study	560.419	9.560	3	0.023
Advice	562.156	11.297	3	0.010
Parents Education	551.552	0.693	3	0.875

Institute/Field		B	Wald	df	Exp(B)
Non-Professional/ Private	Intercept	-3.646	1.392	1	-
	Marks	0.003	0.622	1	1.003
	No of siblings	0.040	0.083	1	1.040
	Gender(female)	-1.591	11.164	1	0.204
	Gender(male)	-	-	0	-
	Income(low)	-1.299	5.210	1	0.273
	Income(high)	-	-	0	-
	Scope(no)	1.380	3.087	1	3.976
	Scope(yes)	-	-	0	-
	Advice(no)	0.534	1.091	1	1.706
	Advice(yes)	-	-	0	-
	Parents Education(low)	0.088	0.032	1	1.092
	Parents Education(high)	-	-	0	-
Professional/ Public	Intercept	-7.517	7.397	1	-
	Marks	0.10	9.945	1	1.010
	No of siblings	-0.064	0.410	1	0.938
	Gender(female)	-1.344	13.700	1	0.261
	Gender(male)	-	-	0	-
	Income(low)	0.220	0.521	1	1.245
	Income(high)	-	-	0	-
	Scope(no)	-0.505	2.215	1	0.603
	Scope(yes)	-	-	0	-
	Advice(no)	0.991	0.065	1	1.103
	Advice(yes)	-	-	0	-
	Parents Education(low)	0.260	0.346	1	1.215
	Parents Education(high)	-	-	0	-
Professional/ Private	Intercept	0.913	0.223	1	-
	Marks	0.002	0.778	1	1.002
	No of siblings	-0.216	4.834	1	0.805
	Gender(female)	-1.828	28.863	1	0.161
	Gender(male)	-	-	0	-
	Income(low)	-1.626	23.725	1	0.197
	Income(high)	-	-	0	-
	Scope(no)	-0.247	0.536	1	0.781
	Scope(yes)	-	-	0	-
	Advice(no)	0.991	8.161	1	2.694
	Advice(yes)	-	-	0	-
	Parents Education(low)	0.260	0.603	1	1.297
	Parents Education(high)	-	-	0	-

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 445-452

SOCIO-DEMOGRAPHIC AND ECONOMIC RISK FACTORS ASSOCIATED WITH CHILD MORTALITY: EVIDENCE FROM PUNJAB-PAKISTAN

Hifza Zulfiqar and **Zaheer Abbas**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
Email: hifzazulifqar26@gmail.com; zaheerabbas@uog.edu.pk

ABSTRACT

The purpose of the research was to analyze socio-demographis and economic factors associated with child mortality in Punjab province of Pakistan. Infant mortality "tip of the iceberg" rates are very sensitive measures that are associated with many factors of child's health. To test these factors, we used the data from Multiple Indicator Cluster Survey (MICS), Punjab Bureau of Statistics, Pakistan. Log linear model was applied to analyze multiple relationships between indicators of socio-demographic and economic factors associated with child mortality along with descriptive statistics. Findings indicate that mother's educational level and wealth index had inverse effects on child mortality in Punjab. The study recommended that batter health services and necessary planned interventions should be provided to weakest areas that results significant change in child mortality in province Punjab.

KEYWORDS

Socio-demography, Economic, Factors, Child mortality, Health, Log linear model.

INTRODUCTION

Death is universal truth in human societies of the world that has been constantly trying to delay death since evolution (George and Marianne, 2003). Developed nations are achieving their targets in the context of child mortality and under developed nations are still facing the problem of high child mortality (Gopal and Mohammad, 2014). More then ten million children under the age of five years are dying around the world and fifty percent of them are passing away in Pakistan (Government of Pakistan, 2001c). Child mortality has taken a central location in nationwide discussion and vital issue for population researchers due to relation with lack of healthcare facilities and resources for crawling population (Government of Pakistan, 2001a). Healthy lifestyle and high socio-economic status decreases the chance of mortality, morbidity and infant mortality (Pollack and Perkin, 2006).

Child mortality rates in Pakistan is signigicantly decreasing but currently there is a big difference as developed and under developed countries (*Singariya, 2015*). Child mortality rates is nine per thousand in developed countries whereas at 153 per thousand in the under developed countries (World Bank, 1993). It is also reported that death rates are reasonable high in children than the adults in both developed and under developed countries (Sathar, 1987). Child mortality under five years of age is affected by socio-economic status, eduational level of parents, availability of health care services, sanitation and nutritional diet (Ross et al., 2007). Every tenth child in Pakistan is died due to non-avialabilty of health facilities and high disease rate in urban and rural areas (Cleland and Farooqui, 1998). Diarrhea and acute respiratory illness which cause malnutrition are vital causes of infant mortality and Pakistan placed at sixth number in population having infectious diseases (Bryce et al., 2005).

Infant mortality rates are highly significnat which determine the economic acheivement (United Nation, 1999) and child health indication is poor with unbearably high infant and child mortality rates in certain regions of Pakistan (Government of Pakistan, 2003). This acutally does not establish a better health explanation of the country and Pakistna have fast economic progress but this was not sufficient to define human progress (Government of Pakistan, 2001b). An amount of nationwide work has been specified that the income level is inconsistent in a society and mortality rate is also very high in these societies (Boldstad and Manda, 2001).

Klaauw and Wang, (2003)[7] suggested an adjusted parametric structure for study of child mortality and infant. This structure determines that infant and child mortality rates can be decreased significantly by improving the household's socioeconomic and environmental attributes. Singh and Yu, 1996[8] observed differences and trends including the socioeconomic and demographic variables in US child mortality from 1950 to 1993, the study showed that poor socioeconomic factors are increasing the child death also health care trend may bring decrease in the child mortality if increasing the socioeconomic factors.

Marianne, et al., 2003[9] measured the association of federal poverty with the states infant and child mortality. Results showed that the demography and basic needs were more closely related to infant and child mortality than states economic ranking. Singariya, 2015[10] studied between different major regions of India and described the determinants and trends of neonatal, infant and under five child mortality. Results showed that mortality with urban factors is negatively associated. Go pal, 2013[11] indicated the tendency in rural and urban areas all causes and cause specific mortality in the USA 1969 to 2009. Residencies in rural areas have large mortality than living in urban areas in last four years.

MATERIALS AND METHODS

To test association among the variables, we use data from nationally representative survey “Multiple Indicator Cluster Survey (MICS)” conducted by Bureau Statistics, Punjab in 2011. The sample of 74130 mothers was taken through multistage random sampling technique. Questionnaire was used to collect information from selected mothers who had experienced at least one live child birth and died before fifth birthday. The questionnaire was comprised on birth history, household attributes, care of illness, vaccinating, child health etc. We selected variables for analysis purpose includes; child survival (CM8), education of mother (me level), sex (HL4), area (HH6), and wealth index (Windex5).

The log linear analysis was introduced in 1970 to analyze the cross-classified data (Bishop, Finberg and Holland, 1975, Haberman, 1975). There models are the extension of generalized linear models (GLM) that showed better results for independent binary response and polychromous variables and focused on association of grouped data. Log linear model is mainly used to quantify the importance of main effects and interaction/association effects represented by linear equations. The saturated (unrestricted) three factor model that is initially consider for validation is

$$\ln(F_{ijkl}) = \alpha + \alpha^{A} + \alpha^{B} + \alpha^{C} + \alpha^{D} + \alpha^{AB} + \alpha^{AC} + \alpha^{BC} + \alpha^{ABC} + e_{ijkl} \qquad (1)$$

Here, F_{ijk} = Expected frequency and, (α)s are the relative weight of each variable (ABC). We can also consider any sub-model (restricted) of equation (1), that may batter represent the relationship between factors under consideration, consist of some main effects and two factor and higher interactions.

RESULTS

This study is mainly focused on the role of demographic and socio economic factors on child mortality. Data are taken from 74130 married women below 50 years of age who have experienced at least one live birth.

A total of 74130 (60% in rural area) live birth were recoded from the selected sample of Punjab since last five years, in which twenty percent (20.2%) of the children were died before their first birthday. More than fifteen percent (15.6%) of the mothers have age between 15-17 years at the time of delivery. The results in Table 1 represents percentage of child mortality by different factors taken under study. The result shows that more than eleven percent (11.5%) of the children in urban and sixteen percent (15.78%) in rural area were reported died before celebrating their fifth birthday. The results of cross tabulation indicate that 95 child deaths per 1000 live births happened in rural areas as against 46 deaths per 1000 in urban areas of Punjab. This severe difference is a sign that the rural areas in which over 70 percent of our population is living has facing lack of basic health facilities, proper roads, clean water, malnourishment and education. It is further found that almost fifty percent (47.65%) of the child births were associated with the mothers have no education. It also observed that mothers with no education have the higher child deaths 68 per 1000. There was a reduction in child deaths as mother's level of education has increased.

Table 1
Child Mortality by Demography and Socioeconomic Individualities in 2011

Factors	Categories	Mortality Percentage		
		Yes	No	%
Area	Urban	3410	26222	11.50
	Rural	7022	37476	15.78
Mother Education	No Education	5043	30278	14.27
	Primary	1860	12022	13.41
	Secondary	1058	6424	14.14
	Higher	2468	14971	14.15
Wealth Index Quintiles	Lowest	4998	18116	21.62
	Middle	5434	39331	12.13
	Highest	0	6251	0000
Sex	Male	5300	32604	13.98
	Female	5131	31091	14.16

Source: Computed from cross tabulation results from MICS datasets, 2011.

Dissimilarity in the circulation of wealth of the country this is also a reason for high child mortality in Punjab. The lowest wealth indexed families have almost twenty two percent (21.62%) mortality rate, on the other hand, highest wealth indexed families have reported zero child mortality. Children born in a family with poor wealth quintile recorded 67 child deaths against the rich wealth quintile of 0 deaths per 1000. As long as the wealth of the country is in the hands of the few opportunity ones child mortality reduction couldn't be possible. The Table 1 also shows that 71 per 1000 male death as against 69 per 1000 female deaths in Punjab. Singh and Yu, 1996 described poor socioeconomic factor increase the mortality.

On the basics of descriptive results, we further proceed to infer that which factor and interaction have the most significant effect on child mortality from four under consideration factors. We take Child mortality as function of Area of residence, Mother's Higher Education, Wealth index and Sex of Child.

In Table 2, on the basics of chi-square test of goodness of fit, we have select the more parsimonious (restricted) model because it will quiet describe the data equally well, as the saturated model which retains former associations between variables. We attain this over the use of backward selection method (Hierarchical log-linear modeling) of the log linear models. The method created with the saturated model for child mortality with the designs follow,

Design:

Constant + CM8 + HH6 + windex5 + HL4 + me level
*+ CM8 * HH6 + CM8 * HL4 + CM8 * me level + CM8 * windex5*
*+ HH6 * HL4 + HH6 * me level + HH6 * windex5 + HL4 * me level*
*+ windex5 * HL4 + windex5 * me level + CM8 * HH6 * HL4*
*+ CM8 * HH6 * me level + CM8 * HH6 * windex5*
*+ CM8 * HL4 * me level + CM8 * windex5 * HL4*
*+ CM8 * windex5 * me level + HH6 * HL4 * me level*
*+ HH6 * windex5 * HL4 + HH6 * windex5 * me level*
*+ windex5 * HL4 * me level + CM8 * HH6 * HL4 * me level*
*+ CM8 * HH6 * windex5 * HL4 + CM8 * HH6 * windex5 * me level*
+ CM8 windex5* HL4 * me level + HH6 * windex5 * HL4 * me level.*

where CM8: Is Child alive?; HL4 : Sex of Child; HH6: Type of Place of Residence; Me level: Mother's Highest Education; Windex5: Wealth Index Quintiles.

This is symbolically denoted as

$$\begin{aligned}\ln(P_{ijklm}) = {} & \mu + \alpha_i + \alpha_j + \alpha_k + \alpha_l + \alpha_m + \alpha_{ij} + \alpha_{ik} + \alpha_{il} + \alpha_{im} \\ & + \alpha_{jk} + \alpha_{jl} + \alpha_{jm} + \alpha_{kl} + \alpha_{km} + \alpha_{lm} + \alpha_{ijk} + \alpha_{ijl} \\ & + \alpha_{ijm} + \alpha_{jkl} + \alpha_{jkm} + \alpha_{ikl} + \alpha_{ikm} + \alpha_{ilm} + \alpha_{jlm} + \alpha_{klm} \\ & + \alpha_{ijkl} + \alpha_{ijkm} + \alpha_{jklm} + \alpha_{ijklm} \\ & \quad i = 1,2 \quad j = 1,2 \quad l = 1,2 \quad k = 1,2,3 \quad m = 1,2,3,4,5\end{aligned}$$

where µ is the logarithm of the whole mean.

The likelihood ratio and the Pearson chi-square for the model is also zero. These measures indicate that model is perfectly fits the data.

Table 2
Test for Partial Association

Effects	DF	Partial Chi Square	Sig.
CM8*HH6*windex5*HL4	2	.290	.865
CM8*HH6*windex5*me level	8	3.499	.899
CM8*HH6*HL4*me level	4	1.911	.752
CM8*windex5*HL4*me level	8	4.477	.812
HH6*windex5*HL4*me level	8	8.751	.364
CM8*HH6*windex5	2	10.147	.006
CM8*HH6*HL4	1	1.803	.179
CM8*windex5*HL4	2	3.319	.190
HH6*windex5*HL4	2	1.196	.550
CM8*HH6*me level	4	3.982	.408
CM8*windex5*me level	8	.961	.998
HH6*windex5*me level	8	6.214	.623
CM8*HL4*me level	4	4.948	.293
HH6*HL4*me level	4	1.125	.890
windex5*HL4*me level	8	3.409	.906
CM8*HH6	1	12.194	.000
CM8*windex5	2	2736.785	.000
HH6*windex5	2	14293.145	.000
CM8*HL4	1	.084	.772
HH6*HL4	1	.026	.871
windex5*HL4	2	7.004	.030
CM8*me level	4	3.237	.519
HH6*me level	4	1.884	.757
windex5*me level	8	31.059	.000
HL4*me level	4	.765	.943
CM8	1	42530.394	.000
HH6	1	3000.067	.000
windex5	2	32929.370	.000
HL4	1	38.177	.000
Me level	4	54937.146	.000

CM8: Is Child alive?; HL4: Sex of Child; HH6: Type of Place of Residence; Me level: Mother’s Highest Education; Windex5: Wealth Index Quintiles.

Table 3 contains test of significant for K-ways effects. It reveal the rejection of the 3^{rd}, 4^{th} and 5^{th} order interactions with p-values 0.508, 0.149 and 0.990 respectively. It means that above stated effects have no contributions to the saturated model at 5% level of significance.

Results further depict that all single effect and two order interaction effects has significant contribution in child mortality. After the elimination of the 3rd, 4th and 5th order higher interactions the most parsimonious model has generating class and can be represented as

$$\ln(P_{ijklm}) = \mu + \alpha_i + \alpha_j + \alpha_k + \alpha_l + \alpha_m + \alpha_{ij} + \alpha_{ik} + \alpha_{il}$$
$$+\alpha_{im} + \alpha_{jk} + \alpha_{jl} + \alpha_{jm} + \alpha_{kl} + \alpha_{km} + \alpha_{lm}$$
$$i = 1,2 \quad j = 1,2 \quad l = 1,2 \quad k = 1,2,3 \quad m = 1,2,3,4,5$$

The above model represent a restricted model a subset of the saturated model which best explain the various demographic and socioeconomic factors directly affecting child mortality in Punjab.

Table 3
Tests for K way and Higher Order effects

K ways and Higher Order		DF	Significance	Likelihood Ratio
Effects K	1	119	0.000	1.511E5
	2	110	0.000	17688.190
	3	81	0.978	57.514
	4	38	0.964	23.862
	5	8	1.00	0.691
K Ways Effect	1	9	0.000	1.334E5
	2	29	0.000	17630.676
	3	42	0.846	33.647
	4	30	0.808	23.176
	5	8	1.000	0.691

CONCLUSION

Socio-demographic and economic factors are the main attributes of child mortality in Punjab. The log linear model show that education and socio-economic characteristics have significant impact on child mortality and improved survival is found to happen in families having high educational level, high income, female children and residing in urban areas of Punjab. The child death in rural area is higher as compared to urban areas based on satisfactory facilities of health care, hygienic conditions, food quality and level of education. Income of family also matter in the case of child mortality in Punjab. High income families are providing extensive care to their children under the age of five years and supporting environment. Education of mother is also playing pivotal to promote child health in Punjab. Educated mothers are providing care and fulfilling nutritional diet to their newborn babies and till the age of five years. So, they have healthy child as compared to less educated mothers. The study recommends that there is still needs to improve health care facilities and programs to improve child health at local, national and international level to minimize child mortality.

RESEARCH IMPLICATIONS

The findings of the present research can be used to minimize child mortality by health practionners, policy makers, scholars and all the stakeholders working on child health and mother's health as well.

Future Research

Further research should be conducted on different research areas including basic health unit, health care centers, health care facilities, community factors, socio-religious factors, medicine as a politics, doctors and nurses role etc. to minimize child mortality and improve socio-demographic and economic conditions of the families in Punjab and Pakistan as well.

Funding

There was no funding available to conduct the research.

ACKNOWLEDGMENT

We are very thankful to Department of Bureau Statistics, Punjab who conducted "Multiple Indicator Cluster Survey (MICS)" for national interest.

REFERENCE

1. Bolstad, W.M. and Manda, S.O. (2001). Investigating child mortality in Malawi using family and community random effects: a Bayesian analysis. *Journal of the American Statistical Association*, 96(453), 12-19.
2. Bryce, J., Boschi-Pinto, C., Shibuya, K., Black, R.E. and WHO Child Health Epidemiology Reference Group (2005). WHO estimates of the causes of death in children. *The Lancet*, 365(9465), 1147-1152.
3. Cleland, J., Hakim, A. and Farooqui, M.N.I. (1998). Health and Survival: Chapter 8 in *Pakistan Fertility and Family Planning Survey 1996-97*. National Institute of Population Studies, Islamabad, Islamabad Center for Population Studies, London.
4. Cochran, W.G. (1977). *Sampling Techniques*, 3rd edition. John Willey and Sons, Inc. New York.
5. Dixon, W.J. and Massey, F.J. (1969). *Introduction to Statistical Analysis*, 3rd edition. McGraw-Hill, Kogakusha.
6. Hillemeier, M.M., Lynch, J., Harper, S., Raghunathan, T. and Kaplan, G.A. (2003). Relative or absolute standards for child poverty: A state-level analysis of infant and child mortality. *American Journal of Public Health*, 93(4), 652-657.
7. Singh, G.K. and Siahpush, M. (2014). Widening rural–urban disparities in all-cause mortality and mortality from major causes of death in the USA, 1969–2009. *Journal of Urban Health*, 91(2), 272-292.
8. Government of Pakistan (2000). *Pakistan Integrated Household Survey, Round 3, 1998-99*. Federal Bureau of Statistics, Islamabad.
9. Government of Pakistan (2001a). *1998 Census Report of Pakistan*. Population Census Organization, Statistics Division, Islamabad.

10. Government of Pakistan (2001b). *Population Policy of Pakistan*. A Document of the Ministry of Population Welfare, Islamabad.
11. Government of Pakistan (2001c). *A Document on the Population Stabilization*. Pakistan Population Policy, Ministry of Population Welfare, Islamabad.
12. Government of Pakistan (2003). *Pakistan Demographic Survey*. Federal Bureau of Statistics, Statistics Division, Islamabad.
13. Marckwardt, A.M. (1984). *Response Rates, Callbacks and Coverage. The WFS Experience*, World Fertility Survey Report 55, London.
14. Pollack, A.E., Ross, J. and Perkin, G. (2006). *Intrauterine Devices in Developing Countries: Assessing Opportunities for Expanding Access and Use*. The William and Flora Hewlett Foundation, Menlo Park, California.
15. Ross, J., Abel, E.A. and Abel, K. (2004). Plateaus during the rise of contraceptive prevalence. *Int. Family Planning Persp.*, 30, 39-44.
16. Ross, J., Stover, J. and Adelaja, D. (2007). Family planning programs in 2004: new assessments in a changing environment. *Int. Family Planning Persp*., 33, 22-30.
17. Sathar, Z.A. (1987). Seeking explanations for high levels of infant mortality in Pakistan. *Pakistan Dev. Rev.*, 26, 55-70.
18. Shryock, H.S. and Siegel, J.S. (1976). *The Methods and Materials of Demography*. Academic Press, Inc. (London) Ltd. UK.
19. Singariya, M.R. and Bhargava, R. (2015). Trends and Determinants of Neonatal, Infant and Under-Five Mortality in India. *Euro-Asian Journal of Economics and Finance*, 3(1), 34-43.
20. Tekce, B. and Shorter, F.C. (1984). Determinants of Child Mortality: A Study of Squatter Settlement in Jordan. *Popul. Dev. Rev*. (A Supplement to child Survival), 10, 257-79.
21. United Nations (1999). *Programme of Action of the 1994 International Conference on Population and Development*. United Nations Publication, New York.
22. Wolpin, K.I. (1997). Determinants and consequences of the mortality and health of infants and children. In: Rosenzweig, M.R. and O. Stark (eds.), *Handbook of Population and Family Economics*, Vol. 1A. Elsevier, Amsterdam.
23. World Bank (1993). *World Development Report on Investing in Health*, New York, Oxford University Press.
24. Zar, J.H. (2007). *Biostatistical Analysis*, 5th edition. Pearson Education Inc., India.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 453-466

DETERMINANT OF SATISFACTION OF YOUTH ON NEW GOVERNMENT POLICIES, A CROSS-SECTIONAL STUDY AT LCWU, GCU, PU AND KC LAHORE

Maham Ramzan and **Anum Fatima**
Lahore College for Women University, Lahore, Pakistan
Email: mahamramzan21@gmail.com

ABSTRACT

This work defines the fact-finding study to examine the factors which cause the satisfaction of youth on new government polices among students with special reference to students of LCWU, GCU, PU and KC, Lahore. The data was collected from 398 students through questionnaires to access student's satisfaction towards new government policies. A pre-test survey based on a sample of twenty students were carried out to check the reliability of the designed questionnaire, consisting of profile variables and 26 ordinal scale based statements. The reliability was checked through Cronbach's Alpha. Factor Analysis has been carried out to analyze the data. The factor which effect the overall satisfaction on new government policies among youth are Health issues, Law and order, economic growth and educational Research. It is estimated that whatever may be the efficient of new government policies; its un-satisfaction always score over its satisfaction. This study provides an evidence of saying that Health Issue is the major factor regarding the satisfaction of youth on new government policies.

KEYWORDS

Government Policies, Factor Analysis.

1. INTRODUCTION

A Government is the system or group of people governing organized community. A government is an organization having a dominant in physical force in a specific geographic location. Government is necessary to ensure the proper use of force. According to this reason, it must have the ability to enforce the laws and it also must specify objective law to clarify the use of force. Government punishes and corrects any law breaking citizen back to the civil life. Government assures the people to protect their rights. A principle or course of action proposed or implemented by a governing body is termed as "policy". the group of people that in unison to guide and support a community, unit, business, institutions and many more are Governing bodies. The rules or ideologies that optimistically better guides decisions, resulting in positive outcomes that augment the community is stated as Government Policies.

Economic policies are about ensuring the efficient functioning of government related to their finance, business environment, and stock exchange of Pakistan. The economic policies are used to moderate the ups and downs in the economy and redistributing economic resources between individuals. The government provides the educational and research policies to empower the youth so that they are able to work for the development. The policies are relate to higher education. These policies provide research facilities to students to discover unrevealed aspects of life earth and many more. These policies are based on the resources of country from which the educational institutes can be facilitated. As the life of a person is most important in this world. Health care includes the procedures taken to provide care for those who are ill and to make the development of health care services. The health care provided to each citizens equally, gender-equal and accessible, and giving equal treatment according to the need of sick person. Health and medical care means giving care is based on the needs of individual patients. According to law and order policy the facilities are provided to different field of institutes by a detail concern of laws and principles that are essential to follow them. The crimes that have to be controlled by introducing rules and laws at equal level are also included in these policies. This entails preventing and combating criminal activities, investigating crime, enforcing sentences and providing support to victims of crime.

This work defines the fact-finding study to examine the factors which cause the satisfaction of youth on new government polices among students.

The experience of the Chinese government in the process of reform and development by focusing on five major roles: (a) promoter of growth, (b) manager of economy, (c) distributor of income, (d) regulator of industry, and (e) protector of citizen and business (Kuotsai T L).

Trust in governmental institutions also varies significantly with political-cultural factors. Citizens who are integrated, involved, and engaged in the political system generally have a significantly higher level of trust in most governmental institutions than people who are less integrated, less involved, and less engaged. Outsiders and people who are politically distant, in an ideological sense, from public institutions, have less trust in those institutions. The same is true for political-cultural factors, when institutions and citizens are loosely integrated (Tom C & Per L).

It is understandable that performance of the public administration has a definite impact on trust in government, but existing levels of trust in government may also have an impact on perceptions of government performance (Geert B & Steven W D V, 2003).

As per our knowledge no research has been done to discover factors affecting the satisfaction of youth on new government policies of a new government. This work revealed the factors that were the reasons behind the satisfaction or un-satisfaction of youth on new government policies. This study not only elaborate the factors behind satisfaction but also explored on the factors of creating dis-satisfaction.

2. METHOD AND MATERIAL

Population under research consist of students learning at Bachelor’s and MPhil’s juncture in distinct field at Lahore college for women university, Lahore (LCWU), Punjab University (PU), Government college university Lahore (GCU) and Kinnaird College (KC). The population size under study is 75240. Data was collected through questionnaire which comprises of 27 questions among which variables are mostly on metric scale.

Acceptable sample size is decided by applying the Yamane method (1967)

$$n = \frac{N}{1+Ne^2}$$
$$= \frac{75240}{1+75240(0.05)^2}$$
$$= 398$$

where

N = size of population and

e = Marginal Error

All universities should be characterized in the sample, so stratified sampling is used. Sample size for different strata (universities) is obtained by proportional allocation.

Table 1
Sample Selection Strategy

S#	University	No. of Students	Sample
1	Lahore college for women university	15000	79
2	Punjab university	45678	242
3	Government college university	10062	53
4	Kinnaird college	4500	24
	Total	**75240**	**398**

Reliability of questionnaire is verified by using the Cronbach's which turned out to be 0.850. Factor Analysis is a statistical tool that measures the impact of a few un-observed variables called factors on a large number of observed variables. It is used as a data reduction method. It may be used to uncover and establish the cause and effect relationship between variables or to confirm a hypothesis. In our study we have used Factor analysis to extract the factors of Youth's satisfaction.

3. RESULTS

Table 2
Cross tab and Chi-Square Association of Significant Values

Variables	Pearson Chi-square value	Degree of Freedom	P-Value	Categories	Satisfaction		
					Yes	No	Total
Education for all	125.290[a]	4.0	0.0	Strongly Agree	2	4	6
				Agree	67	32	99
				Neutral	20	77	97
				Disagree	14	177	191
				Strongly Disagree	1	4	5
School Building	118.483[a]	4.0	0.0	Strongly Agree	1	1	2
				Agree	64	34	98
				Neutral	25	62	87
				Disagree	10	190	200
				Strongly Disagree	5	6	11

Variables	Pearson Chi-square value	Degree of Freedom	P-Value	Categories	Satisfaction		
					Yes	No	Total
Capability of teacher	107.046[a]	2.0	0.0	Strongly Agree	60	30	90
				Agree	6	14	10
				Neutral	20	56	76
				Disagree	12	150	162
				Strongly Disagree	16	44	60
Teacher's attendance	112.291[a]	3.0	0.0	Strongly Agree	18	18	32
				Agree	40	66	106
				Neutral	27	10	37
				Disagree	10	170	180
				Strongly Disagree	8	30	38
Enrollment of new students	112.812[a]	3.0	0.0	Strongly Agree	22	6	28
				Agree	25	64	89
				Neutral	28	28	56
				Disagree	13	180	193
				Strongly Disagree	16	16	32
Unemployment	172.637[a]	2.0	0.0	Strongly Agree	8	7	15
				Agree	70	20	90
				Neutral	6	30	36
				Disagree	10	220	230
				Strongly Disagree	10	17	27
Stock exchange	136.088[a]	2.0	0.0	Strongly Agree	10	1	11
				Agree	48	17	65
				Neutral	36	52	88
				Disagree	11	154	165
				Strongly Disagree	5	64	69
IMF	175.325[a]	2.0	0.0	Strongly Agree	9	7	16
				Agree	70	20	90
				Neutral	3	33	36
				Disagree	15	225	240
				Strongly Disagree	7	9	16
Business environment	121.875[a]	2.0	0.0	Strongly Agree	5	6	11
				Agree	20	50	70
				Neutral	30	55	85
				Disagree	10	199	209
				Strongly Disagree	8	15	23
Currency decline	117.638[a]	2.0	0.0	Strongly Agree	50	20	70
				Agree	25	37	58
				Neutral	20	5	25
				Disagree	14	220	234
				Strongly Disagree	18	12	30

Variables	Pearson Chi-square value	Degree of Freedom	P-Value	Categories	Satisfaction		
					Yes	No	Total
Free medicine	74.668[a]	4.0	0.0	Strongly Agree	15	9	24
				Agree	52	28	80
				Neutral	31	66	97
				Disagree	9	149	158
				Strongly Disagree	3	36	39
Viral disease	48.465[a]	4.0	0.0	Strongly Agree	18	3	21
				Agree	60	34	94
				Neutral	23	66	89
				Disagree	8	168	176
				Strongly Disagree	1	17	18
Cleanliness	60.784[a]	4.0	0.0	Strongly Agree	1	2	3
				Agree	35	74	109
				Neutral	43	41	84
				Disagree	16	170	199
				Strongly Disagree	7	6	13
Relative accommodation	44.071[a]	3.0	0.0	Strongly Agree	4	4	8
				Agree	48	80	128
				Neutral	30	33	63
				Disagree	20	165	185
				Strongly Disagree	6	8	14
Hospital services	28.394[a]	4.0	0.0	Strongly Agree	18	76	26
				Agree	32	50	68
				Neutral	8	2	92
				Disagree	40	150	161
				Strongly Disagree	6	16	22
Doctor's availability	16.386[a]	4.0	0.0	Strongly Agree	23	80	103
				Agree	25	45	70
				Neutral	2	2	4
				Disagree	35	141	155
				Strongly Disagree	19	26	45
Attending patients satisfactory	11.376[a]	4.0	0.0	Strongly Agree	24	85	109
				Agree	30	56	86
				Neutral	3	2	5
				Disagree	45	149	194
				Strongly Disagree	2	2	4
Nursing staff	16.071[a]	4.0	0.0	Strongly Agree	12	13	25
				Agree	33	52	85
				Neutral	1	1	2
				Disagree	31	134	165
				Strongly Disagree	27	94	121

Variables	Pearson Chi-square value	Degree of Freedom	P-Value	Categories	Satisfaction		
					Yes	No	Total
Drug regularity	11.936[a]	4.0	0.0	Strongly Agree	28	35	63
				Agree	23	52	75
				Neutral	2	8	10
				Disagree	23	131	154
				Strongly Disagree	28	68	96
Law and order	3.715[a]	2.0	0.156	Strongly Agree	21	7	28
				Agree	58	40	98
				Neutral	26	67	93
				Disagree	4	143	147
				Strongly Disagree	1	31	32
Accountability	.171[a]	2.0	0.918	Strongly Agree	15	69	84
				Agree	26	65	91
				Neutral	12	17	29
				Disagree	35	130	146
				Strongly Disagree	16	13	48
Justice prevalence	.740[a]	2.0	0.691	Strongly Agree	8	6	14
				Agree	25	75	100
				Neutral	24	68	92
				Disagree	40	130	170
				Strongly Disagree	7	15	22
Timely justice	2.539[a]	2.0	0.281	Strongly Agree	30	60	90
				Agree	20	65	85
				Neutral	4	2	6
				Disagree	40	150	190
				Strongly Disagree	10	17	27
Corruption	0.010[a]	2.0	0.995	Strongly Agree	20	52	72
				Agree	25	82	107
				Neutral	9	20	29
				Disagree	35	120	155
				Strongly Disagree	15	20	35
Kick-back	1.842[a]	2.0	0.398	Strongly Agree	9	4	13
				Agree	48	21	69
				Neutral	47	85	132
				Disagree	5	144	149
				Strongly Disagree	1	34	35
Money laundering	0.113[a]	2.0	0.945	Strongly Agree	10	20	30
				Agree	25	80	105
				Neutral	9	14	23
				Disagree	35	110	145
				Strongly Disagree	25	70	95

The above table describes that the variables except Accountability, Justice Prevalence, Corruption and Money laundering variables all other are associated with the satisfaction of youth on new government policies.

Table 3
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.951
Bartlett's Test of Sphericity	Approx. Chi-Square	9179.463
	df	325
	Sig.	.000

Table 4
Total Variance Explained

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	10.578	40.686	40.686	10.578	40.686	40.686	5.816	22.370	22.370
2	5.912	22.739	63.425	5.912	22.739	63.425	4.973	19.128	41.498
3	1.586	6.101	69.526	1.586	6.101	69.526	4.456	17.140	58.639
4	1.140	4.385	73.912	1.140	4.385	73.912	3.971	15.273	73.912
5	.628	2.414	76.326						
6	.544	2.094	78.419						
7	.456	1.754	80.173						
8	.445	1.711	81.884						
9	.418	1.606	83.490						
10	.374	1.438	84.928						
11	.361	1.390	86.318						
12	.360	1.383	87.701						
13	.339	1.304	89.005						
14	.328	1.261	90.266						
15	.311	1.198	91.463						
16	.282	1.084	92.547						
17	.272	1.045	93.593						
18	.250	.960	94.553						
19	.239	.920	95.473						
20	.227	.873	96.346						
21	.219	.841	97.187						
22	.198	.762	97.949						
23	.185	.713	98.662						
24	.166	.637	99.299						
25	.130	.501	99.800						
26	.052	.200	100.000						
Extraction Method: Principal Component Analysis.									

Eigen values denote the quantity of variance in the data, that is mentioned as the sum of variance that they illuminated by their associated factors. The factors are removed by the sum of variance that they explain. In Table 4, the first four factor explain the majority of the variance, while the remaining factors explain a very small proportion of variance. In this method, only factors with Eigen values larger than 1.0 are taken; the other factors are not involved in the model. The Eigen value greater than 1.0 results in four factors being explained. The factors extracted should explain at least 50% (i.e. ≥50%) of the total variance. the cumulative percentage of variance is considered for the first 4 factors accounted for 73.91%. in the table these 4 factors account for 23.370%, 19.128%, 17.140% and 15.273%. that is 73.91% approximately 73% of the total variance explained to these 4 factors. the remaining 22 factors together account for only approximate 26.09% of the variance. Thus, the model with these 4 factors may be appropriate to represent the data.

A scree plot is a graphical presentation of the variance of each component in the dataset which is used to control how many components should be taken in order to describe a high percentage of the variation in the data. The shape of the plot is used to control the amount of factors. Scree plot and the purpose based on the Eigen values are signifying the same numbers of factors.

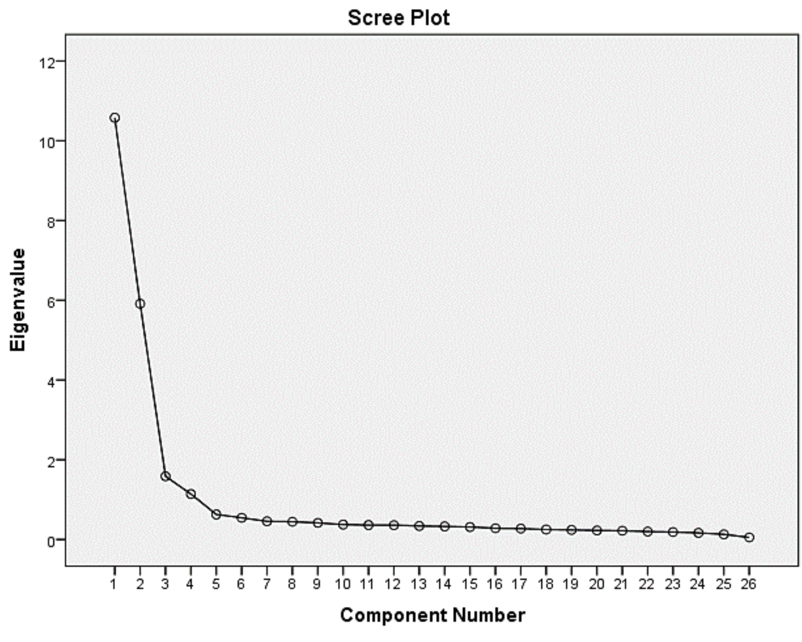


Figure 1

Communality, h^2 specify the sum of variance in each variable. The amount of variance in each variable accounted for by mutual factor is stated as communalities. The second column under "communalities" in Table 5 gives significant material after the desired quantity of factors has been extracted.

Table 5 defines that the variable concerning satisfaction of new government policies such as timely justice (Q22) only measures less than 50% variation of the variance explained by the factor.

The variable such as kick-back (Q3) law and order (Q7) relative accommodation (Q12) money laundering (Q16) nursing staff (Q17) drug regularity authority working (Q23) greater than 60% but less than 70% variation of variance explained by the factors.

The variables such as viral disease (Q2) education for all (Q4) free medicines (Q6) school buildings improved (Q8) cleanliness conditions in hospitals (Q10) justice prevailing (Q11) capability of teachers (Q13) hospital services (Q15) accountability at equal level (Q18) doctor's availability in OPD's (Q19) corruption control over (Q20) decline of currency (Q21) attending the patients satisfactory (25) greater than 70% but less than 80%.

The variables such as teacher's attendance (Q24) stock exchange (Q5) business environment (Q14) are greater than 80% but less than 90% remaining variables such as unemployment (Q1) IMF (Q9) are greater than 90% but less than 92%.

Table 5
Communalities

	Initial	Extraction
Q4: Education for all	1.000	.763
Q8: School buildings improved	1.000	.795
Q13: Capability of teacher	1.000	.787
Q24: Teacher's attendance	1.000	.803
Q26: Enrollment of new students	1.000	.815
Q1: unemployment	1.000	.902
Q5: Stock exchange	1.000	.858
Q9: IMF	1.000	.918
Q14: Business environment	1.000	.823
Q21: Currency	1.000	.796
Q6: Free medicines	1.000	.712
Q2: Viral disease	1.000	.727
Q10: cleanliness	1.000	.703
Q12: Relatives accommodation	1.000	.648
Q15: Hospital services	1.000	.711
Q19: Doctor's availability	1.000	.738
Q25: Attending patient satisfactory	1.000	.733
Q17: Nursing staff	1.000	.677
Q23: Drug regularity	1.000	.666
Q7: Law and order	1.000	.671
Q18: Accountability	1.000	.709
Q11: Justice prevalence	1.000	.701
Q22: Timely justice	1.000	.473
Q20: Corruption	1.000	.768
Q3: Kickback	1.000	.629
Q16: Money laundering	1.000	.689

The rotated component matrix is calculated by using varimax rotation. The factors thus obtained are given in the Table 6. With all the communalities of enough size, factor loadings arise out to be exceeding 0.50 presenting the one-half of the variance is accounted for by the loading on the single factor.

Table 6
Rotated Component Matrix[a]

	Component			
	1	2	3	4
Viral disease	.800			
Free medicines	.792			
Doctors availability	.770	.339		
Attending patient satisfactory	.765	.378		
Hospital services	.764	.326		
Relatives accommodation	.721			
Cleanliness	.718			
Nursing staff	.709	.407		
Drug regularity	.700	.405		
Corruption		.835		
Accountability	.312	.781		
Justice prevalence	.319	.772		
Money laundering		.769		
Law and order		.767		
Kick-back		.748		
Timely justice		.635		
IMF			.889	.334
Unemployment			.877	.341
Stock exchange			.845	.353
Business environment			.842	.311
Currency			.811	.363
Enrollment of new students			.339	.825
Teacher’s attendance			.352	.806
School buildings			.354	.805
Capability of teacher			.365	.792
Education for all			.386	.767

Presentation and Interpretation of Factors

Finally, the extracted factors with their respective Eigen values, explained percentage of variance and coefficient’s of reliability are presented in following tabular form. These factors are characterized by the definite declaration written to reflect the four attributes towards satisfaction of youth towards new government policies.

Table 7
Factor Analysis

Factors and Items	**Factor Loadings**	**Eigen Values**	**Explained Variance %**	**Cronbach's Alpha**
Factor 1: Health Issues		5.816	22.37	0.88047
Viral Disease	0.8			
Free Medicines	0.792			
Doctor's Availability	0.77			
Attending Patients Satisfactory	0.765			
Hospital Services	0.764			
Relative Accommodation	0.721			
Cleanliness	0.718			
Nursing Staff	0.709			
Drug Regularity	0.7			
Factor 2: Law and Order		4.973	19.128	0.96671
Corruption	0.835			
Accountability	0.781			
Justice Prevalence	0.772			
Money Laundering	0.769			
Law and Order	0.767			
Kick-Back	0.748			
Timely Justice	0.635			
Factor 3: Economic Issues		4.456	17.14	0.92599
IMF	0.889			
Unemployment	0.877			
Stock Exchange	0.845			
Business Environment	0.842			
Currency	0.811			
Factor 4: Education Issues		3.971	15.273	0.91384
Enrollment of New Students	0.825			
Teacher's Attendance	0.806			
School Buildings	0.805			
Capability of Teachers	0.792			
Education for All	0.767			

The first factor extracted is **Health Issues** accounts for the largest amount of variance which is 22.37. The highest factor loading is for viral disease which is 0.800. The value of Cochran's alpha coefficient is 0.88047 which shows the internal reliability of 1st factor **Health Issues.**

The second extracted factor **Law and Order** accounts for the second largest amount of variance which is 19.128. The highest factor loading for second factor is for control over corruption which is 0.835. The value of Cochran's alpha coefficient is 0.96671 that represents the internal reliability of **Law and Order** factor.

The third extracted factor is **Economic Issues** narrate for the third largest amount of variance which is 17.14. The highest factor loading for this factor is for IMF which is

0.889. The value of Cochran's alpha coefficient is 0.926 representing the internal reliability of third factor that is **Economic Issues.**

The last elicitation factor is **Education Issues** reveal for the smallest amount of variance from the other factors which is 15.273. The highest factor loading for the **Education issues** factor is for enrollment of new students in educational institutes of Pakistan which is 0.825. The value of Cochran's alpha coefficient is 0.9138 that describe the internal reliability of **Education Issues** factor.

4. CONCLUSION

The aim of this study was to give the information and to discern the factors, reasons and significance concerning the satisfaction on new government policies among youth, with access to the students of LCWU, GCU, PU and KC. The factors among the satisfaction of youth related to new government has also been studied. The purpose of this study was to provide a bridge between the satisfaction of youth and new government policies by keeping in touch with the factors of un-satisfaction also. This study conclude that this new government have to take magnificent steps to improve their new policies therefore it can easily build trust of nation and to satisfy the youth. A factor analysis is discussed about satisfaction towards new government policies and is followed by the reliability testing, which compute four major kind of attributes towards the new government policies. Those factors are presented below:

Health Care

The concept of satisfaction among youth related to health care is an important factor as it is based on the life of person. Health care policies are useful to interact with the youth that make an easier and efficient way to develop trust on the government working to serve nation. The government new health care policies should enable all the health institutes with better treatment facilities to provide comfort zone to its nation about their lives. Moreover, one other significant feature of health care policies is to create facilities for nation to give proper treatment and attention of doctors.

Law and Order

New policies have made to attain the objective of accountability at equal level. These policies are working with great effort, to providing the result in shape of accountability not only every citizen but also the political leaders. These policies develop bridge towards the gaining youth trust. The new policies are introduced for the decline of most prominent crimes, such as, money laundering, kick-back and corruption. If these new government policies succeed in gaining their objectives then the outcome of their effort is satisfaction of youth. These variables in our result leads towards the un-satisfaction of youth on new government policies.

Economic Growth

Youth find platform for their future concerns according to the economic condition of Pakistan. Since Pakistan is working with IMF from many years and according to new government policies to empower the economic growth, the government work with IMF organization through which every citizen of Pakistan is under the billions of debts. The government policies are not working efficiently to provide the better environment for business. The government should build effective policies for the decline on unemployment, so that the youth will get satisfied by government for the future of economic growth of Pakistan.

Education Issues

As we know that the education is the necessity of every person to survive and to earn better for their families. Seeking education is an important part of life of every person. The new government starting their education policies by the motto “Education for all”. The government empower the youth by providing better facilities to the educational institutes and facilities for the maintenance of school buildings through which children get better environment who are trying to seek knowledge. The working of government policies should be effective to improve the capability of teachers so that the youth will gain more knowledge and develop interest to their studies. If all these efforts can be done by the government and their policies then the outcome of un-satisfaction of youth may convert into the satisfaction among new government policies.

5. SUGGESTIONS

Government policies are necessary for the working of government world-wide. These policies should be for the welfare of nation and base on the factors that are the needs of nation. At the present condition of Pakistan government policies can make it easier and provide great help to the government to serve nation. Today every country’s government is using the policies to empower their position among citizens. Every kind of policies have a major impact on the current and also effect the upcoming years of their lives as these policies working may make their future better or worse. Hence, the new government policies should truly base not only for the betterment of Pakistan but also for the youth satisfaction. The government policies should develop an accuracy in their working for the satisfaction of their nation.

6. RECOMMENDATIONS

The significances of this study are based totally upon the research conducted among the students of LCWU, GCU, PU and KC and may not be correct to the youth of other cities as a justification of background and cultural factors. This survey can also be passed on a larger scale (nation-wide). Moreover, the population could also be extended by include elders (40 above) as the satisfaction on new government policies will be more elaborated by the elders.

REFERENCES

1. Anderson, C.J. and Guillory, C.A. (1997). Political institutions and satisfaction with democracy: A cross-national analysis of consensus and majoritarian systems. *American Political Science Review*, 91(1), 66-81.
2. Khushik, F. and Diemer, A. (2018). Critical Analysis of Education Policies in Pakistan: A Sustainable Development Perspective. *Social Science Learning Education Journal*, 3(09), 01-16.
3. Bouckaert, G. and Van de Walle, S. (2003). Comparing measures of citizen trust and user satisfaction as indicators of ‘good governance’: difficulties in linking trust and satisfaction indicators. *International Review of Administrative Sciences*, 69(3), 329-343.

4. Bouckaert, G. (2012). Trust and public administration. *Administration*, 60(1), 91-115.
5. Van de Walle, S. and Bouckaert, G. (2003). Public service performance and trust in government: the problem of causality. *International Journal of Public Administration*, 26(8-9), 891-913.
6. Kampen, J.K., De Walle, S.V. and Bouckaert, G. (2006). Assessing the relation between satisfaction with public service delivery and trust in Government. The impact of the predisposition of citizens toward Government on evaluations of its performance. *Public Performance & Management Review*, 29(4), 387-404.
7. Liou, K.T. (1998). The role of government in economic development: The Chinese experience. *International Journal of Public Administration*, 21(9), 1257-1283.
8. Mukumbareza, C. (2015). *Evaluating citizen satisfaction with the quality of e-government information services provided by Southern African Development Community governments* (Doctoral dissertation). Durban University of Technology.
9. Bowe, R., Ball, S.J. and Gold, A. (2017). *Reforming education and changing schools: Case studies in policy sociology*. Routledge.
10. Oomsels, P. and Bouckaert, G. (2012). Managing trust in public organisations: A consolidated approach and its contradictions. In *Paper for the XVI IRSPM conference: Contradictions in Public Management: Managing in Volatile Times* (p. 25).
11. Dildar, S.M., Saif, N.N. and Naz, A. (2016). Review of Educational Policies of Pakistan: Planning and Implication flows. In *First International Conference 5Es*, Department of Education, University of Sargodha, Pakistan.
12. Kim, S. (2010). Public trust in government in Japan and South Korea: Does the rise of critical citizens matter? *Public Administration Review*, 70(5), 801-810.
13. Kim, S. (2007). Democratic governance, government performance, and trust in government: China and South Korea. In *The public management research conference.* Tucson, Arizona (pp. 25-27).
14. Van de Walle, S., Van Roosbroek, S. and Bouckaert, G. (2008). Trust in the public sector: is there any evidence for a long-term decline? *International Review of Administrative Sciences*, 74(1), 47-64.
15. Bredgaard, T. (2011). When the government governs: Closing compliance gaps in Danish employment policies. *International Journal of Public Administration*, 34(12), 764-774.
16. Christensen, T. and Lægreid, P. (2005). Trust in government: The relative importance of service satisfaction, political factors, and demography. *Public Performance & Management Review*, 28(4), 487-511.
17. Christensen, T. and Lægreid, P. (2003, September). Trust in government–the significance of modernism, political cynicism and integration. In *EGPA Annual Conference, Workshop on Quality, Satisfaction, and Trust in Government: Reassessing Trust in a Reinvented Government, September* (pp. 3-6).
18. Van de Walle, S. and Bouckaert, G. (2003). Public service performance and trust in government: the problem of causality. *International Journal of Public Administration*, 26(8-9), 891-913.
19. Van de Walle, S. (2004). *Perceptions of administrative performance: The key to trust in government?* (Doctoral dissertation, Katholieke Universiteit Leuven).

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 467-480

CONTRIBUTIONS AND CONTROVERSIES OF STATISTICS IN RESEARCH, DEVELOPMENT AND PUBLIC DECISIONS

Shahjahan Khan
School of Sciences, Faculty of Health, Engineering and Sciences
International Centre for Applied Climate Science
Centre for Health Sciences Research
University of Southern Queensland, Toowoomba, Australia
Email: shahjahan.khan@usq.edu.au

ABSTRACT

In recent years, increasingly decision-makers and researchers are relying on an evidenced-based approach, using results extracted by analysing exponentially growing data. Statistical methods have become an integral part of analysing and interpreting simple or complex data to make valid and accurate decisions for society and science in the face of uncertainty. Statistics has found new momentum due to the emergence of big data and data analytics. Because of its power to uncover, otherwise unavailable, valuable information from data, statistics is increasingly being applied for both private and public benefit including business and trade, not to mention its crucial role in research and innovative technologies. The key role of statistics is engraved in the planning and development of every successful nation and international agency. In fact, the scope of statistics is not confined by borders as it disseminates through regional and international organisations and agencies driving social, economic, environmental, health, poverty elimination, education and other agendas for planned development. Obviously statistical methods are capable of analysing data to determine various indices, indicators, predictions, estimates, and perform tests to help identify any problems and take appropriate measures to remedy them. Because of the obvious power of statistics it has been used, misused and abused by those who are in authority and power.

1. INTRODUCTION

The world is increasingly moving towards evidence-based decision-making. As a result use of statistics in extracting evidence (information) from data has become inevitable. Almost all numerical evidences are based on data and hence analysis of data and interpretation of results produced by statistical procedure are essential. Systematic review and meta-analysis play a key role in synthesising results from multiple independent studies to produce pooled evidences using statistical methods (cf. Khan et al. 2019). While in most cases statistics is being used professionally by the decision-makers to deliver immense benefit to the society, but sometimes it is abused by people in authority.

To acknowledge and encourage the world about the importance of statistics, in promoting the World Statistics Day on 20 October 2015, the UN Secretary-General pointed out, "statistics are critical for evidence-based decision making across all cultural and historical backgrounds of countries and irrespective of their level of development." He also noted that UN General Assembly "acknowledged the fundamental importance of sustainable national statistical capacity to produce reliable and timely statistics and indicators." Khan (2016) highlighted various aspects of the celebration and wide ranging applications of statistics in the contemporary world.

In recognition of the emerging value and importance of statistics in the modern world a number of initiatives were undertaken globally. One of them is the Partnership in Statistics for Development in the 21st Century (PARIS21) since 1999. The assessment and monitoring of the progress of achieving each of the eighteen goal of the Millennium Dev. Goals (MDG) required calculation of various statistical rates, ratios and indices. This is also true for the seventeen target areas under the Sustainable Dev. Goals (SDG).

The recent emergence of data science and data analytics have further intensified the use of statistics. In 2012 the US President Obama announced Big Data Initiative. The White House press release described the goals of the Initiative: “to advance the state-of-the-art core technologies needed to collect, store, preserve, manage, analyze, and share huge quantities of data; harness these technologies to accelerate the pace of discovery in science and engineering; to strengthen our national security, and transform teaching and learning; and to expand the work force to needed to develop and use Big Data technologies.”

As the use of statistics becomes more wide spread the people in high position realise its power in changing people’s perception and influence it may have in mobilising public opinion on key public issues. Because of the inherent power of statistics to bring about changes in the society, powerful people wrongly abuse statistics to their personal gains.

2. STATISTICS AND ITS OLD HISTORY

The word statistics is used in difference contexts. There are three popular use of the word. Most frequently, the word statistics refers to an academic discipline, a piece of scientific information, or a value calculated from any sample data to represent a characteristic of the sample. Popularly the word statistics is referred to describe its key activities. It is in this sense, statistics is the science of learning from data. It deals with collection or production of data, and then compile and analyse data, and interpret outcomes of the results produced by the analyses.

Statistics is the numerical mirror of the society that reflects on the characteristics of collective life and is able to reveal the hidden jewel of facts which are there but unavailable otherwise. It helps to make appropriate decision in the face of uncertainty and limited information. Statistics has found its presence and crucial place in almost all areas of the contemporary knowledge society.

In ancient time, statistics was used by the kings and rulers for managing the state’s army and human resources, and hence it was known as the Science of Kings. In the bible, God asked Moses to conduct a census as in the Book of Numbers, Chapter 1: Verse 2-4:2 Take a census of all the congregation of the children of Israel, by their families, by their fathers’ houses, according to the number of names, every male individually, 3 from twenty years old and above—all who are able to go to war in Israel. You and Aaron shall number them by their armies. 4 And with you there shall be a man from every tribe, each one the head of his father’s house.

Interestingly, the Qur'an mentions the word Al-Ahsa (statistics) many times in different contexts and forms. First Household Survey was conducted by Omar Ibn Khattab (RA) in the 7th century AD to determine social security benefit for the citizens of his Khilafat.

Basra based scientist Yaqub bin Ishaq Al-Kindi (801-873) introduced the maximum likelihood idea when working with the frequency distribution of Arabic alphabets. The maximum likelihood principle leads a large part of statistical inference even today.

From Bukhara Abu Ali Ibn Sina Avicenna (980-1037) championed the medical science by writing famous Canon of Medicine. He introduced clinical trials for checking usefulness of interventions. This is still the best way to determine if any specific treatment is working or not.

In Baghdad, Muhammad ibn Mūsā al-Khwārizmī (780-850) introduced Arabic numerals and the concept of algorithm. The algorithm is the corner stone of computers and internet searches. The erstwhile USSR (Russia) Government printed a Memorial Postal Stamp in his honour in 1983.

3. MODERN STATISTICS

The origin of statistics in the modern time started in 1665 with the publication of Natural and Political Observations upon the Bills of Mortality by John Graunt in the Royal Society. Early applications of statistical thinking revolved around the needs of states to base policy on defence, demographic and economic data, hence its stat- etymology. Scope of statistics broadened in the early 19th century to include the collection and analysis of data from diverse field of modern society.

Today, statistics is widely employed in government, business, and the natural, health, engineering and social sciences (cf. Khan, 2013).Statistics plays a major role in shaping and providing scientific information that is useful in almost every aspect of human life, and beyond. Modern decision making, be it for an individual or a business or any national Government or an international agency, is increasingly using statistical methods to improve the quality of evidence/decision.

It has made its marks from astronomy to administration, business to biology, housing to health, engineering to environment, from commerce to community, manufacturing to ministry, marketing to management, from medicine to money, industry to infrastructure, politics to policy, tourism to trade-union, sports to strategy.

4. MAJOR CONTRIBUTIONS OF STATISTICS

Statistics has made huge contribution to the modern world. Key to these contributions are some of the frequently used statistical methods that has helped change the world in many positive ways. One such ancient, yet modern, method is the census involving 100% count of every element or item or individual in the population. It provides vital data for planning and development of any nation.

A census requires significant human, financial and technical resources, and takes long time (years) for the results to be available. A more quick and much less resource intensive, but scientifically valid, method is the sample survey that randomly selects a part of the population ensuring that it retains the typical characteristics of the population. Introduction of sample survey and the idea of probability based random sampling has changed the way science and society used to work.

The statistical method of Controlled Experiments and Randomised Clinical Trials (RCT) have enable humanity to establish cause and effect relationship of any interventions or treatments. The statistical process control and total quality management (TQM) have totally changed the industrial and production process.

Applications of statistical models such as linear and non-linear regression methods have become a common practice in the current data centric world. Statistical models play key role in economic, business, social, psychological, medical, agricultural and other scientific studies. Various statistical models are used in forecasting (weather, yield), and predictions as well as opinion and exit polls.

The use of statistical inferential techniques including confidence intervals and test of hypotheses have become almost routine in everyday research and decision-making. Most of the quantitative research methods are directly or indirectly statistical techniques.

In the contemporary world, many universities and research institutions offer courses without explicitly reflecting statistics in the title but the contents are essentially full of statistical methods and techniques. A brief list of such popularly used courses include, but not limited to, Actuarial Science, Bioinformatics, Biometry, Biostatistics, Climatology, Data Mining, Data Science, Data Analytics, Demography, Econometrics, Epidemiology, Environmetrics, Forecasting, Image processing, Management Science, Machine Learning, Meta-analysis, Quality/Process control, Quantitative Methods, Queuing methods, Reliability analysis, Research Methods, Survival analysis, Stochastic Process, Time Series Analysis etc.

4.1 Statistics Revolutionised Manufacturing Industry

Statistics has changed production technology in the USA and elsewhere. The key to this change is the work of William E Deming (1900 – 1993), an American engineer, statistician, author, and management consultant. Originally, an electrical engineer, specialized in mathematical physics, developed sampling techniques that are still used by U.S. Department of Census and Bureau of Labor Statistics.

In Japan the industrial miracle (1950-1960) was possible because of better design of products to improve service, higher level of uniform product quality, improvement of product testing in industry and lab, and greater sales through side markets. In US, he introduced the System of Profound Knowledge by appreciating a system, understanding variation and psychology as well as epistemology (theory of knowledge) in production industry. President Regan awarded him National Medal of Technology in 1987.

Statistics revolutionized the manufacturing industry in Japan due to the work of Genichi Taguchi (1924-2012), a Japanese engineer and statistician. From 1950s Taguchi developed a methodology for applying statistics to improve the quality of manufactured goods. He worked in the textile and public health where he became interested in design of experiments. He developed methods for enhancing quality and reliability. His contribution to loss function, robust design, and reduction of variation have influenced fields beyond product design and manufacturing, such as sales process engineering.

5. DIVERSE APPLICATIONS OF STATISTICS

In this section, we illustrate some real life examples of how statistics is widely used in various disciplines.

Environmental Application

In a study of the impact of extractable iron, extractable aluminum and potential of hydrogen (pH) on soil adsorption index the following data were reported.

Table 1
Soil Adsorption Data

Iron	Aluminum	pH	Adsorption
61	13	7.70	4
175	21	7.70	18
111	24	6.80	14
124	23	7.30	18
130	64	5.10	26
173	38	5.70	26
169	33	5.80	21
169	61	5.20	30
160	39	6.30	28
244	51	5.70	36
257	112	112	4.40
65	88	4.50	62
199	54	6.20	40

The simple matrix scatter plot of the four variables shows that there are linear relationship between soil adsorption index and the three explanatory variables.

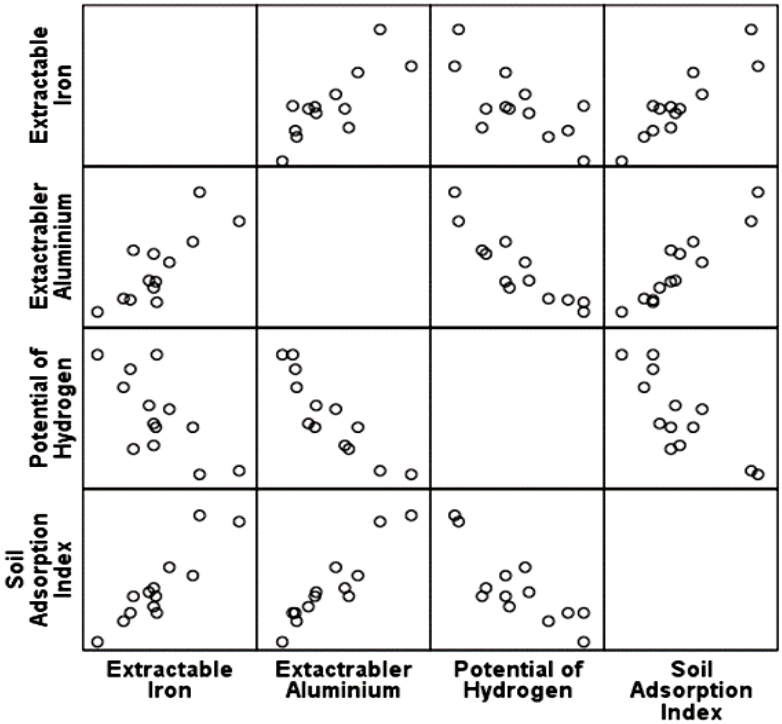


Figure 1: Matric Scatter Plot of Soil Adsorption Index, Iron, Aluminum and Potential o Hydrogen

A multiple regression model fitting produced high value of coefficient of determination, adjusted $R^2 = 0.933$ and the overall regression is highly significant (p-value 0). Further analyses reveal that potential of hydrogen (p-value 0.608) has no impact on the soil adsorption index, but both extractable iron (p-value 0.005) and extractable aluminum (p-value 0.008) significantly contributed to the soil adsorption index.

Table 2
Multiple Regression Analysis outputs

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.975[a]	.950	.933	4.545

a. Predictors: (Constant), Potential of Hydrogen, Extractable Iron, Extactrabler Aluminium

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3535.783	3	1178.594	57.057	.000[b]
	Residual	185.909	9	20.657		
	Total	3721.692	12			

a. Dependent Variable: Soil Adsorption Index

b. Predictors: (Constant), Potential of Hydrogen, Extractable Iron, Extactrabler Aluminium

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	-18.251	20.748		-.880	.402
	Extractable Iron	.112	.031	.446	3.638	.005
	Extactrabler Aluminium	.398	.118	.660	3.363	.008
	Potential of Hydrogen	1.421	2.663	.089	.534	.607

a. Dependent Variable: Soil Adsorption Index

Medical Application

Consider the study of Kent et al. (2015) on the orexigenic hormone acyl-ghrelin increases adult hippocampal neurogenesis and enhances pattern separation. This is a paper on health science, and from the title of the paper there is no indication of any involvement of statistics. But inside the paper statistics is highly prominent as evidenced by the fitted lines on the graph and use of ANOVA, correlation coefficient and test of hypotheses.

Correlation of neurogenesis in the rostral dentate gyrus (DG) with discrimination on the small and X-small separation tasks. Statistical analysis by ANOVA, Student's *t*-test and Pearson correlation analysis. $*p < 0.05$, $**p < 0.01$, $***p < 0.001$; $n = 12$ rats per group.

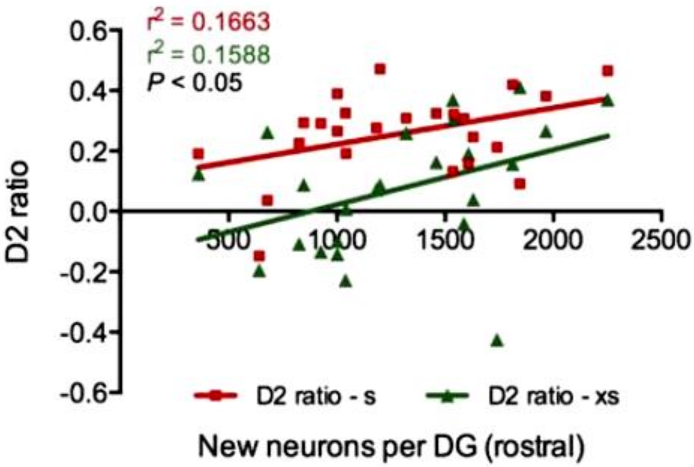


Figure 2: Fitted Model for D2 Ratio and New Neurons per DG (rostral)

Astronomical Application

There are 200,000 galaxies between 800m to 4b light-years away from the earth to be classified.

Figure 3: Part of the Galaxies in the Universe

Prof Simon Driver, International Centre for Radio Astronomy Research, Western Australia says, “These statistics can be used to build a model of how the entire galaxy population in the universe has evolved.”

Public Health Application

From the publication of Australian Institute of Health and Welfare (2018) statistics suggest that 1/3 Australian men and 1/4 women will have cancer by the age of 75. Incidence and mortality of all cancers combined in Australia is graphed in the following figure. There much more male having cancer than women in Australia. Similarly, the mortality from cancer is higher in the male population than their female counterpart.

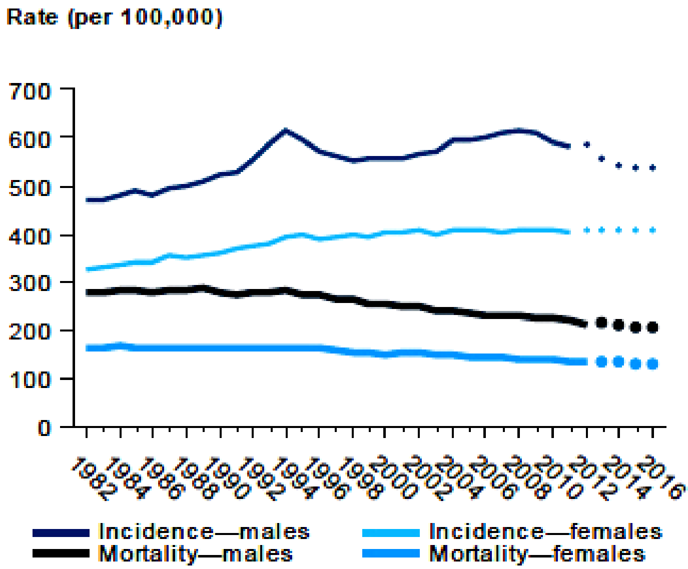


Figure 4: Incidence and Mortality of all Cancers Combined, 1982-2016

From the bar chart of the incidence of breast cancer below, it is evident that the incidence rate is much higher in the women population of age groups of 40-49, 50-59 and 60-69. Women at a higher age group 70 + also has significant number of incidence of breast cancer.

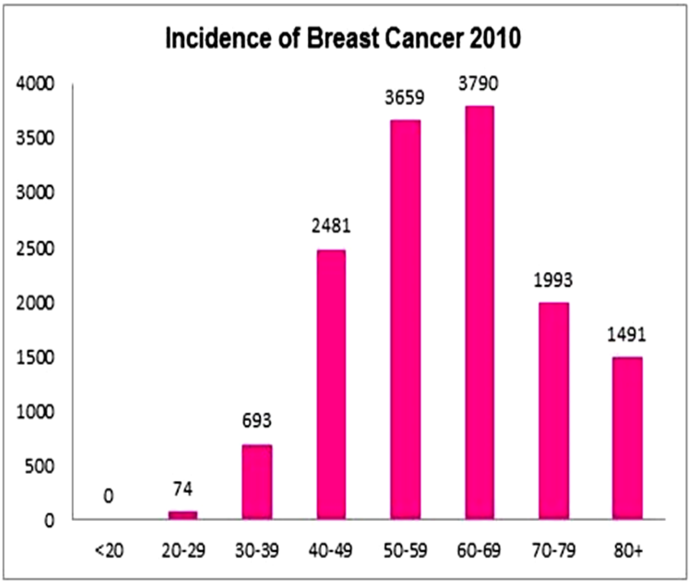


Figure 5: Incidence of Breast Cancer by Age for Australian Women

Decision-makers and policy-drivers often use above statistics to create awareness and device policies to rectify or improve various priority issues with high impact on the people, economy and environment.

6. CONTROVERSIAL P-VALUE ISSUES

In statistics, p-value is a difficult concept involving probability of rejecting null hypothesis under certain conditions. It measures the credibility of the null hypothesis based on the extremeness of observed sample. If the sample is extreme, the p-value is small and provides stronger evidence against the null hypothesis. As the sample changes, the p-value also changes and so does the evidence against the null hypothesis.

Any research involve hypothesis tests, p-value has been inevitably used. In many cases it has been misused or abused intentionally or due to using it as a sole criterion to decide on hypothesis test without taking in to account many other factors relevant to the problem. Often the researchers do not distinguish between 5% and 4.99% as long as the p-value is less than 5% they rush to conclude in favour of significant results ignoring many other key factors impacting on the results and practical significance. Some argue to have a more stringent level of statistical significance than 5%. Particle physicists have taken to 5 sigma result or 0.0000003, or 1 in 3.5 million.

In the use p-value the following factors should be taken in to account:

a) Methodological quality of the research: No amount of statistics can make up for a study that is the wrong design for the question you care about – or that's poorly done. Check internal and external validity.
b) Effect size: You need to understand exactly what is being measured and how big the apparent effect is to get a result into perspective.

c) Understanding the uncertainty: If there’s a p-value, you need to know exactly what it was, not only that it is under 0.05 – is it just under, or are there more zeros? You need better ways to understand the uncertainty of the estimate; and that means standard deviations, margin of error, or confidence/credible intervals.
d) More than one study: Certainty doesn’t come from a one-off – and especially not from a surprising one. This lead to systematic reviews and meta-analysis.

6.1 ASA's Statement on *p*-Value

The American Statistician in 2016 published a statement of the American Statistical Association (ASA) highlighting the importance of Context, Process, and Purpose of the study in interpreting and using the p-value.

a) P-values can indicate how incompatible the data are with a specified statistical model.
b) P-values do not measure the probability that the studied hypothesis is true, or the probability that the data were produced by random chance alone.
c) Scientific conclusions and business or policy decisions should not be based only on whether a p-value passes a specific threshold.
d) Proper inference requires full reporting and transparency P-values and related analyses should not be reported selectively.
e) A p-value, or statistical significance, does not measure the size of an effect or the importance of a result. Statistical significance is not the same as scientific/practical significance.
f) By itself, a p-value does not provide a good measure of evidence regarding a model or hypothesis.

Wassersterin (2016) explain the context, process, and purpose of the ASA Statement on p-value as at https://amstat.tandfonline.com/doi/full/10.1080/00031305.2016.1154108

7. USE AND ABUSE OF STATISTICS IN POLITICS

Politicians use statistics for varieties of reasons, especially in assessing the popularity of their personal, party or Government position and approval of their policies using various surveys. In the democratic system, often decisions on when the Government would go for the election is very much dictated by the outcomes of the opinion polls, often without admitting it in public. Normally, in the democratic nations, government would not call an election if the opinion poll is not strongly in its favour.

Re-election of President Obama in 2012

The most pressing issue in the US election in 2012 was high unemployment rate which was over 8% up until August 2012. On 5 October 2012, the US Bureau of Labor Statistics reported a 0.3% drop in the unemployment rate for the month of September to 7.8%. This was a make or break figure for the re-election of President Obama.

Table 3
Monthly Unemployment Rate in USA reported on 5 October 2012

Year	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
2008	5.0	4.9	5.1	5.0	5.4	5.6	5.8	6.1	6.1	6.5	6.8	7.3
2009	7.8	8.3	8.7	9.0	9.4	9.5	9.5	9.6	9.8	10.0	9.9	9.9
2010	9.8	9.8	9.9	9.9	9.6	9.4	9.4	9.5	9.5	9.4	9.8	9.3
2011	9.1	9.0	9.0	9.1	9.0	9.1	9.0	9.0	9.0	8.8	8.6	8.5
2012	**8.3**	**8.3**	**8.2**	**8.2**	**8.2**	**8.2**	**8.2**	**8.1**	**7.8**	7.8	7.7	7.9

As expected, the Republican camp claimed that the drop of the **unemployment rate was 'cooked'** by the influence of the administration. Steve Haugen of the Bureau of Labor Statistics: "flatly dismissed the idea that there was any way the White House or Obama campaign could have had a hand in how the numbers turned out." "The data are not manipulated for political reasons. I've been involved in the process myself for almost three decades. There's never been any political manipulation of the data, period,"

Independence of statistics office is absolutely crucial for the objectivity of official statistics and ensuring public trust. If there is no public trust in the data produced by statistics office, there is no need for it and no will benefit from it. Legal and organisational protection of statistics workers in government offices is essential for advoiding any undue influences of the people in power.

Abuse of Statistics in USA

November 2015, Donald Trump tweeted "Whites killed by blacks – 81%", citing "Crime Statistics Bureau – San Francisco". The US fact-checking site ***Politifact*** says – "Bureau" did not exist and the true figure is 15%. When confronted, Trump shrugged and said, "Am I going to check every statistic?"

Noble Laureate Joseph Stiglitz on commenting about President Trump said, "A couple of things are most disturbing – the attack on the press and the attack on the foundations of knowledge which goes beyond the press."

In the face of indiscriminate false statements and claims it is essential to challenge the motive behind 'alternative facts', 'witch hunt' and 'fake news' to uphold facts.

US diplomat and politician, John Adams who assisted Thomas Jefferson in drafting the Declaration of Independence in 1776, unequivocally stated not to play with 'facts' saying, "Facts are stubborn things; and whatever may be our wishes, our inclinations, or the dictates of our passion, they cannot alter the state of facts." For details see https://www.brainyquote.com/quotes/john_adams_134175

8. A FALSE BLAME ON STATISTICS

Some people attribute a very misleading statement ridiculing statistics to Mark Twain. But there is no evidence to suggest that he ever said, "There are three kinds of lies: lies, damned lies and statistics." As at the url http://www.twainquotes.com/Statistics.htmlin his

Autobiography Mark Twain himself noted that the quote is form Benjamin Disraeli. However, Stephen Goranson wrote on the Mark Twain Forum in a post-dated 31 July 2002: Twain's Autobiography attribution of Benjamin Disraeliis generally not accepted.

As found at http://www.york.ac.uk/depts/maths/histstat/lies.htm the origin of the above quote is still uncertain, but if it originated with any one well-known figure, the most likely candidate is Sir Charles Dilke. A 1895 article by Leonard H. Courtney published in The Bristol Mercury, 19 Oct 1891, claimed that Sir Dilke said, “False statements might be arranged according to their degree under three heads, fibs, lies, and statistics.” Interestingly, almost all those who had tried to associate lies with statistics in those days are ‘politicians’. One of the reasons for this may be the way statistics was produced or crafted or manipulated by the people in power without adhering to any methodological, moral, ethical or professional standard.

Obviously, those people loosely used the word statistics before statistics evolved as a scientific discipline. Those days the scientific rigor of statistics was not well established. What may have been known as statistics at that time were nowhere near it is in the contemporary world with the scrutiny of design of studies, quality and randomness of data, analytical and computing appropriateness and modelling and prediction accuracy.

Regardless of the origin of the quote some people conveniently use it to undermine statistics because statistics can expose the unpleasant truth and reveal the incontestable facts. It is anyone’s guess, why the politicians would discredit statistics. They were often uncomfortable with facts presented by statistics that go against their interest and popularity. This is also true for the powerful politicians of our time, not much has changed in this profession. The introduction of misleading phrases such as ‘alternative fact’ is even worse than blaming statistics to be lies. So, regardless of the origin of the quote, it would be more appropriate to modify it in the following way to represent the facts on the ground: “There are three kinds of lies: lies, damned lies and politics.” And there are plenty of evidences for this.

Anyone will agree that it is not fault of statistics if some people abuse or misuse it. Similar to a knife, statistics can be used by professionals objectively to benefit people and society or abuse it by dishonest individuals to mislead/harm society. If anyone to blame, it is not statistics, rather who misuses this powerful tool for selfish reasons.

The fact remains, ‘***figures don't lie, but*** **liars can** ***figure***’ *and/or* ***‘figures fool,*** **when** ***fools figure*****’.**

REFERENCES

1. Australian Institute of Health and Welfare (2018). Cancer in Australia: Actual incidence data from 1982 to 2013 and mortality data from 1982 to 2014 with projections to 2017. *Asia Pacific Journal of Clinical Oncology*, 14(1), 5-15. https://doi.org/10.1111/ajco.12761
2. Kent, B.A., Beynon, A.L., Hornsby, A.K., Bekinschtein, P., Bussey, T.J., Davies, J.S. and Saksida, L.M. (2015). The orexigenic hormone acyl-ghrelin increases adult hippocampal neurogenesis and enhances pattern separation. *Psychoneuroendocrinology*, 51, 431-439. https://doi.org/10.1016/j.psyneuen.2014. 10.015

3. Khan, S., Memon, B. and Memon, M.A. (2019). Meta-analysis: a critical appraisal of methodology, benefits and drawbacks. *British Journal of Hospital Medicine*, 80(11), 636-641. https://doi.org/10.12968/hmed.2019.80.11.636
4. Khan, S. (2016). Statistics on Spotlight: Reflection of World Statistics Day. *Journal of ISOSS*, 2(2), 175-184.
5. Khan, S. (2013). Statistics in Planning and Development. *Pakistan Journal of Statistics*, 29(4), 513-524
6. Khan, S. and Memon, M.A. (2017). Statistics for Evidence-based Decisions – Use, Misuse and Abuse. *Third International Conference on Science, Engineering & Environment (SEE), USQ, Brisbane, Australia*, Nov. 13-16, 2017, 22-30, ISBN: 978-4-9905958-9-0 C3051.
7. Wasserstein, R.L. and Lazar, N.A. (2016). The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129-133, DOI: 10.1080/00031305.2016.1154108

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 481-486

CONSEQUENCES FACED BY WORKING WOMEN (SURVEY STUDY)

Sidra Razzaq, Rizwan Ali and **Nagina Zulfiqar**
Department of statistics, University of Wah, Wah Cantt, Pakistan
Email: sidra.razzaq@uow.edu.pk
rizwan.ali@uow.edu.pk
nagina.zulfiqar@uow.edu.pk

ABSTRACT

The main objective of this survey is to investigate the factors which plays a role behind the paid working women, consequences faced by women at work place, comparison of trends in women's labor force participation rate in past years and now a day and it also analyze the preaching of ISLAM about the working of women.

Our sampling frame contains both male and female students, employs working at different organization and people who are resident of Islamabad. The sample consists of Quaid-e-Azam University, Islamabad, COMSATS Islamabad, NUML and SZABIST Islamabad. The technique used for selecting the sample is Simple Random Sampling. Questionnaire approach is used for the collection of data. X^2 –test and Z-test are used for the estimation of results. The permissible error was taken to be as 5%.

1. INTRODUCTION

The present study focuses on the working women who have faced and are still facing numerous work problems. Social, cultural and religious factors have played a vital role for women entering the job market. The main objective of the survey is to investigate the factors which play a vital role behind the working women.

Throughout the history of Pakistan, Muslim women have suffered a great deal of unnecessary restrictions due to the misconceptions of Islam. Though women play a vital role in economic development of society, yet their contribution is rarely acknowledged. Persuading career out of home does not normally find encouraging responses. The first and the most common problem which a working women has to face is harassment, including abusive remarks, offers for lunch or dinner dates, unnecessary physical contact and lustful stares.

Much has changed in the last decade but even in urban areas, when a woman travels to work, she still has to be ready for stares and rude remarks from strangers. For years, women in rural areas of Pakistan who have chosen to work have been the subject of books and papers supporting and explaining their problems, but no work has been done in support of working women in urban Pakistan.

2. LITERATURE REVIEW

Not much of work has been done for the working women of Pakistan. The articles which were studied before the survey was conducted are the Problems of Working

Women in Karachi (Hussain, 2009). Another article which helped understand the problems faced by working women is, Working Women in Pakistan: Analysis of Issues and Problems (Pervez et al., 2015). Other articles which helped understand the topic were, Critical Analysis of Life of Pakistani Women: Views of Educated Women (Akhter and Akbar, 2016). Policy brief on female labor force participation in Pakistan (Serrat, Park and Yoshino, 2016). Balancing work and family through social support among working women in Pakistan (Malik et al., 2010). Employment situation of women in Pakistan (Sadaquat, 2011). A multi-level perspective on equal employment opportunity for women in Pakistan (Syed, Pio and Ali, 2013). Moving Forward with the Legal Empowerment of Women in Pakistan (Weiss, 2012).

3. OBJECTIVE OF THE STUDY

The major objective of the study are:

- To analyze the factors, that plays a role behind the paid working women.
- Women face double pressure of work, one at her job place and the other at her house.
- To find out the factors of depression, in the paid working women.
- Effect on socialization of children when their mother's are working.
- Harassment and gender discrimination.
- Consequences faced by women at work place.
- Preaching of ISLAM about working of women.

4. VARIABLES UNDER STUDY

The variables which are studied in this survey are Harassment, economic pressure, dual responsibility, depression and use of drugs, High divorce rate or khula rate.

Harassments:

Working women has to face different type of harassments. Harassments in different manifestation starting from staring and stalking to sexual advances. Women do not generally feel intellectually challenged in their pursuits; rather they face psychological pressures, which restrict their active participation in work. The fear of police is more common in women than men. On account of corruption and inhumane attitudes, prevailing in police department and judicial environment, women show great reluctance in contacting police, to claim their rights or seek remedy. Family laws, and other disputes, which may involve women, are perceived to be failed in safeguarding women rights. Complicated and lengthy court procedures and uncomfortable environment, are among those reasons, which complement the problems of working women in official affairs.

Economıc Pressure:

The economic role of women is clearly defined, and is well-acknowledged, as far as faith and tradition is concerned, but the developments in contemporary society have created new issues, to be considered in this respect. However, the share of female population, in economic activity, is not proportionate with their number. Due to tough economic conditions and other social reasons, women are entering the economic arena in large number, without any identification of those areas, which can prove to be more productive for them. In certain sectors, women are taken as cheap labor and are paid fewer wages than their male counterparts. Females workplace is not often adequately designed where they could feel at ease during work and break timings. Such unfriendly and oppressive behavior of employers is a problem in our society in which the real goal of any entrepreneur is profit maximization, and not human welfare.

Poor transport system is also adding to the problems of working women. Public transport is costly, scarcely available and unpleasant. Government is still unable to provide adequate transport facilities to the commuters, even in the major cities.

Dual Responsıbılıtıes:

Working women has to face dual responsibilities, one at work places and other at their homes. Their family life, especially children, suffer a lot because they have to perform dual jobs in form of fulfilling their family obligations along with the job. Working women find themselves in another dilemma. They always feel that home is their original domain, which they have to sustain in all circumstances. Women generally prefer to stay at home, and normally, do not opt for the job out of choice, but out of necessity. They always have a feeling of guilt for sparing less time for family and maternal responsibilities. This dual responsibility proves the double burden on her, and resultantly, she has to fight simultaneously on two fronts.

Depression and Use of Drugs*:*

A woman with depression already has a slowed and dulled nervous system. They may drink to escape worrisome thoughts or emotional pain, but they are actually making their entire situation worse over time. Women usually begin an addiction through an emotional pathway, such as feelings of depression or anxiety. These feelings could range anywhere from mild anxiety to a prolonged clinical depression. Whether a woman turns to an addiction depends on many factors.

High Divorce/Khula Rate:

Divorce and khula rates are high in working women due to due to their dual responsibilities. Financial problems are a main cause of divorce. The huge financial difficulties would shortly, threaten a couple's life. Early marriage is one of the main causes of divorce. According to one study, 13.5% of divorced women and 23.7% of divorced men married early. Moreover, some families ask for dowry which is a common reason for divorce, and this factor also forces women to enter in the labor market.

5. SURVEY METHODOLOGY

Survey Design:

This study was a survey work which was designed to enquire into and provide information about the consequences faced by working women.

Area of the Study:

The area of the study was Islamabad. The working people of organizations, NGO’s, hospitals etc., were used for the survey.

Sampled Population:

The population of the study consists of all the working people of age greater than 22.

Sampling Technique:

The population understudy is homogenous so we have adopted the Simple Random Sampling Technique.

Sampling Unit:

Our sampling unit is a single person of age greater than 22.

Sample Size:

The sample size we have selected is of 100.

Method of Inquiry:

Keeping in mind the requirement of our survey we have chosen the questionnaire approach.

6. HYPOTHESIS

The hypothesis which are under study are as follows:

Chi-square (χ^2) Test:

The chi-square statistic is commonly used for testing relationships between categorical/qualitative variables. The null hypothesis of the chi-square test is that no relationship exists on the categorical variables in the population; they are independent.

Hypothesis Test # 1:

H_0: Trend of working women is independent of the economic pressure.
H_1: Trend of working women is not independent of the economic pressure.

Hypothesis Test # 2:

H_0: Working of women due to large family members is independent on the pressure from their men.
H_1: Working of women due to large family members is dependent of pressure from their men.

Hypothesis Test # 3:

H_0: Increased divorced/ khula rate independent on rigid behavior of working women.
H_1: Increased divorced/ khula rate dependent on rigid behavior of working women.

Hypothesis Test # 4:

H_0: Harassment of working women is independent of societies.
H_1: Harassment of working women is dependent of societies.

Hypothesis Test # 5:

H_0: Economic development of society is independent of increased number of working.
H_1: Economic development of society is dependent of increased number of working.

7. RESULTS

Hypothesis	Level of Significance	Calculated Value	P-Value	Decision
H_0: Trend of working women is independent of the economic pressure. **H_1:** Trend of working women is not independent of the economic pressure.	$\alpha = 0.05$	$X^2_{cal} < X^2_{(\alpha.v)}$ $X^2_{cal} < X^2_{(0.05,\ 12)}$ $15.343 < 23.3367$	$p>0.10$	Do not reject the null hypothesis (not significant)
H_0: Working of women due to large family members is independent on the pressure from their men. **H_1:** Working of women due to large family members is dependent of pressure from their men.	$\alpha = 0.05$	$X^2_{cal} > X^2_{(\alpha.v)}$ $X^2_{cal} > X^2_{(0.05,\ 16)}$ $36.53 > 28.84$	$p<0.01$	Reject the null hypothesis (Strongly significant)
H_0: Increased divorced/khula rate independent on rigid behaviour of working women **H_1:** Increased divorced/khula rate dependent on rigid behaviour of working women	$\alpha = 0.05$	$X^2_{cal} < X^2_{(\alpha.v)}$ $X^2_{cal} < X^2_{(0.05,\ 16)}$ $24.8 < 28.84$	$p>0.05$	accept the null hypothesis (not significant)
H_0: Harassment of working women is independent of societies. **H_1:** Harassment of working women is dependent of societies.	$\alpha = 0.05$	$X^2_{cal} > X^2_{(\alpha.v)}$ $X^2_{cal} > X^2_{(0.05,\ 16)}$ $39.67 > 28.84$	$p<0.01$	Reject the null hypothesis (strongly significant)
H_0: Economic development of society is independent of increased number of working **H_1:** Economic development of society is dependent of increased number of working	$\alpha = 0.05$	$X^2_{cal} > X^2_{(\alpha.v)}$ $X^2_{cal} > X^2_{(0.05,\ 08)}$ $26.961 > 17.5343$	$p<0.01$	Reject the null hypothesis (strongly significant)

Z-Test:

Z-test is a statistical test where normal distribution is applied and is basically used for dealing with problems relating to large samples when $n \geq 30$.

Data points should be independent_from each other. One data point isn't related or doesn't affect another data point.

Data should be randomly selected from a population, where each item has an equal chance of being selected.

Hypothesis Test # 6:

H_0: Do you think that ISLAM allows a woman to work?

Hypothesis test # 7:

H_0: Today, a woman of our society is working according to ISLAMIC values?

Hypothesis	Level of Significance	Calculated Value	P-Value	Decision
H_0: Do you think that ISLAM allows a woman to work?	$\alpha = 0.05$	$I\ Z_{cal}\ I \geq I\ Z_{tab}\ I$ $3.4 \geq 1.96$	$p<0.003$	Reject H_0

Hypothesis	Level of Significance	Calculated Value	P-Value	Decision
H_0: Today, a woman of our society is working according to ISLAMIC values?	$\alpha = 0.05$	$I\ Z_{cal}\ I \geq I\ Z_{tab}\ I$ $0.6 \leq 1.96$	$p<0.2743$	Accept H_0

8. COMMENTS AND CONCLUSION

Women are facing many consequences in their personal and professional life. The results shows that working women are respectable in society. Following are the conclusion which we have drawn from this survey.

The trend of working women is dependent of the economic pressure. Another reason for working of women is the large number of family members.

Increased divorced/khula rate is independent on rigid behavior of working women. Harassment of working women is dependent of societies. On the question that does Islam allows a woman to work, majority people said No. People agreed that women of our society are working according to Islamic values.

REFERENCES

1. Hussain, I. (2009). *Problems of working women in Karachi, Pakistan*. Cambridge Scholars Publishing.
2. Pervez, S., Zia-ur-Rehman, M., Javed, J. and Raza, I. (2015). Working Women in Pakistan: Analysis of Issues and Problems. *Pakistan Journal of Social Sciences (PJSS)*, 35(2), 997-1011.
3. Akhter, N. and Akbar, R.A. (2016). Critical Analysis of Life of Pakistani Women: Views of Educated Women. *Journal of Elementary Education*, 26(1), 111-121.
4. Serrat, O., Park, D. and Yoshino, N. (2016). Policy Brief on Female Labor Force Participation in Pakistan. *ADB Briefs*, 80, 100.
5. Malik, M.I., Saif, M.I., Gomez, S.F., Khan, N. and Hussain, S. (2010). Balancing work and family through social support among working women in Pakistan. *African Journal of Business Management*, 4(13), 2864-2870.
6. Begum Sadaquat, M. and Sheikh, Q. (2011). Employment situation of women in Pakistan. *International Journal of Social Economics*, 38(2), 98-113.
7. Syed, J., Pio, E. and Ali, F. (2013). A multi-level perspective on equal employment opportunity for women in Pakistan. *Equality, Diversity and Inclusion: An International Journal*. 32(3), 289-309.
8. Weiss, A.M. (2012). *Moving forward with the legal empowerment of women in Pakistan*. US Institute of Peace.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 487-492

INFORMATION SEEKING BEHAVIOR AND ACADEMIC PERFORMANCE OF UOG STUDENTS

Faqeeha Noor, Zaheer Abbas[§] and **Taskeen Sajid**
Department of Statistics, University of Gujrat, Gujrat, Pakistan
[§]Email: zaheerabbas@uog.edu.pk

ABSTRACT

Information seeking is defined in such a way, in which the individual seek, evaluate, analyze, select, and utilize information. In information seeking process, people change their level of knowledge by exposure of information. In order to check the information seeking behavior and academic performance of students of university of Gujrat, the sample size of 400 students was selected by using stratified random sampling.

A well-structured questionnaire having six factors as "library, internet, classroom, emotions, peers and information seeking" was designed to measure the information seeking behavior and academic performance of students of university of Gujrat. Descriptive analysis, normality, Kruskal Walli's test, Confirmatory factor analysis and structural equation modelling was applied to find the objective of the study. There is no difference in seeking behavior of all the faculties students of UOG, their seeking behaviors are same. Structural equation modelling results showed that the relationship between seeking behavior and library, internet, classroom, emotions and peers is highly significant and goodness of fit shows that the model is good fitted.

KEYWORDS

Information seeking, Academic behavior, library, internet, emotions, peers.

1. INTRODUCTION

The term information comes from "formation" which means indication of the construction of any pattern. In today's world, most of the development depends on the proper availability and utilization of the information. Without valid and reliable information, it is difficult to progress in every aspect of life (Bajpai, 1990).

Actually information is the action of thoughts of humans. It can be concrete or in abstract form. Information is data which is manipulated in the form that is meaningful for the individual and it is valuable in prospective actions and decisions. It is the basic necessities of the individual. For doing personal and academic concerns it provides basics to the individual to participate in new activities and interests.

Information seeking is defined as the ways in which individual seek, evaluate, analyse, select and utilize information. For seeking new information people contacts with each other, share ideas and organize the computer based information system (Wilson, 2000).

In information seeking process, people change their level of knowledge. They interpret and manipulate their knowledge by gaining of new information. Seeking information deals with our cognitive functioning like problem solving, learning and thoughts. These are higher cognitive process (Marchionini, 1995).

Behaviour refers to how an individual acts and behaves in a particular situation which factors affects their ways to express. According to King “A manner in which a user conduct himself in relation to a given information environment is information seeking behaviour”.

Objective of Study

Objective of the study are

- To explore the sources that students used to gain information relative to their studies.
- To measure academic performance of UOG students.
- To compare the information seeking behaviour of different faculties of UOG.
- To develop a structure equation model for UOG student’s academic performance on the base of their information seeking behaviour.

2. LITERATURE REVIEW

Das et al. (2016), conducted study on information seeking behavior of law students in the digital age. Stratified random sampling was used to collect data by students through questionnaire, the target population was university of Dhaka and sample size was one hundred and eighty-three (183). The study showed that the students highly preferred electronic resources instead of print documents and library played a vital role in seeking information of law students. Furthermore, the legal information resources are also most used resources by law students. The main purpose of seeking information was preparing assignment and for examination.

Nagar et al. (2014), the study carried out that information seeking is very essential part of the researches. Selected population was the research scholars of faculty of Arts and faculty of Science of Aligarh Muslim university, sample size was one hundred and twenty (120) and questionnaire was used to collect data. The results showed that the researchers of faculty of Arts used library on daily basis while researchers of faculty of Science used it occasionally. They further conclude that the scholar of faculty of Science were more satisfied with library staff as compare to Arts Scholars. First source of both the faculties was internet in information seeking and most of the scholars were satisfied with internet, library services and library staff but some of the research scholars are not satisfied with these resources.

Sankari et al. (2011), explored that libraries are the main center of seeking or retrieving information. Studied population was the users of Vinayaka Mission’s Kirupananda Variyar Engineering college library in Solem and sample size was one hundred and sixteen (116). Simple random sampling was used to collect data through questionnaire. The percentage analysis and chi-square test was applied. Results showed that to find the relative material about information problem is not an easy task. Library played a major role in seeking information. But there should be need to conduct seminars and workshops about library using.

3. MATERIAL AND METHOD

Materials and methods are used to complete the study in a particular area. A cross sectional study was conducted. The study was conducted in University of Gujrat. The primary data were collected by using the well-defined questionnaire. In order to check the information seeking behavior and academic performance of students of university of Gujrat, the sample size of 400 students was selected by using stratified random sampling from UOG. Six factors were considered as a measure of Information Seeking Behavior. These factors are library, internet, classroom, emotions, peers and information seeking. Each factor has different indicators and is measured on 5-point Likert scale. Later available software SPSS 21.0V, STATISTICA AND AMOUS was used in data analysis.

Data Analysis Techniques

Data analysis techniques of this study are

- Descriptive statistics
- Normality
- Kruskal walli's test
- Confirmatory factor analysis (CFA)
- Structure equation model (SEM)

4. RESULTS AND DISCUSSION

Reliability of Questionnaire

Questionnaire is said to be reliable if it constantly quantity the aims for which it is planned. Reliability is a requirement for the validity of a research. Any measuring tool which does not measure some features constantly has a slight chance of being considered a valid measure of that feature. Reliable data brings the consistent and significant outcomes. Cronbach's alpha is used to measure reliability. For reliability of questionnaire the value of Cronbach's alpha should be greater than 0.7.

Reliability Analysis

Table A-1 shows the factor wise reliability analysis. The values of Cronbach's alpha of factors library, internet, classroom, emotions, peers and information seeking is 0.77, 0.79, 0.74, 0.78, 0.87 and 0.72 respectively. Since all these values are greater than 0.7, so, data is reliable and we proceed for further analysis.

Descriptive Statistics

Descriptive Statistics discuss summaries of sample data. Table A-2 involves Descriptive statistics (minimum, maximum, mean, standard deviation,) of quantitative variables of respondent of this study like age and CGPA. The average age of the respondents of this study is 21.27 with standard deviation 2.378. The minimum and maximum age of the respondents is 17 and 35 respectively. The average CGPA of the respondents is 3.209 with standard deviation 0.396. Also, the minimum and maximum CGPA of the respondents is 1.70 and 4.00 respectively.

Kruskal Wallis test

Table A-4 displays the group means of faculties and factors of information seeking behaviour. As the p-value of Kruskal walli's is greater than 0.05 so the null hypothesis is accepted. And conclude that the seeking behaviour of students of all faculties is same.

Confirmatory Factor Analysis

Table A-5 shows the goodness of fit of model. Chi-square value is used to measure the goodness of fit measure in CFA. The recommended value for chi-square/df is less than 3. But in this study the values of chi-square/df of library, internet, classroom, emotions, peers and information seeking are 3.2, 5.3, 6.10, 17.7, 11.04 and 5.8 respectively. Since, all these values are greater than 3, so, these values not support the estimated model. The GFI value of all the variables are greater than 0.90 so these values support the estimated model.

Structure Equation Model

Table A-6 shows the estimates for structure equation modelling. As p-values of all variables are less than 0.05, so all values are significant. The estimates of all relations of endogenous and exogenous variables are significant.

Table A-7 shows the goodness of fit of model for structural equation model. Chi-square value is used to measure the goodness of fit measure in SEM. The recommended value for chi-square/df is less than or equal to three. In our study the value of chi-square/df is 2.6 which indicate that it supports the estimated model. The GFI value is 0.932 and AGFI value is 0.902. Both the values are greater than 0.9 which shows that model is good fitted. The RMSEA value should be less than 0.08. In this model the value of RMSEA is 0.064 which means model is good fitted.

5. COMMENTS AND CONCLUSION

It is recommended that

- University should improve the quality of library services.
- The availability of books for students should be sufficient.
- There should be group discussion between peers which helps to seek information.
- Question answering environment should be created by teacher in classroom.
- Facility of internet should be provided for study.
- Students should be aware of using internet and library services.

CONCLUSION

The study is about information seeking behavior and academic performance of UOG students. The study is held to check that which factors effect the ISB of students and which sources were used in gaining information. Several factors which are used to conduct the study are library, internet, classroom, emotions, peers and information seeking. It is concluded that there is no difference in SB of the students of all faculties. Mostly students prefer internet to seek information and also peers are the major source of seeking information. The study shows that the relationship between information seeking and all other factors is significant and the SEM model is good fitted.

REFERENCES

1. Das, R.K. and Jadab, A. (2016). Information Seeking Behaviour of Law Students in the Digital Age. A user study at university of Dhaka. *International Research Journal of Library and Information Science*, 6(3), 381-396.
2. Nagar, P. and Munshi, S.A. (2014). Information Seeking Behaviour of Female Research Scholars of Faculties of Social Science and Arts in Aligarh Muslim University: A comparative study. *Journal of Library and Information Science*, 4(1), 459-474.
3. Sankari, L., Chinnasamy, K. and Venkatachalam, A.M. (2011). A Study of Information Seeking Behaviour of Users of Vinayaka Mission's Kirupananda Variyar Engineering College Library in Salem. International Journal of Library and Information Science, 3(8), 179-186.
4. Ajiboye, J.O. and Tella, A. (2007). University Undergraduate Students Information Seeking Behaviour: Implications for Quality in Higher Education in Africa. *The Turkish Online Journal of Educational Technology*, 6(1), 40-52.
5. Fasola, O.S. (2013). Information Seeking Behaviour of Students of Ajayi Crowther University, Oyo, Nigeria. *Brazilian Journal of Information Science,* 7(2), 47-60.

APPENDIX-A

Table A-1
Reliability Analysis

Factors	Cronbach's Alpha	N of Items
Library	0.776	9
Internet	0.797	7
Classroom	0.741	7
Emotions	0.785	5
Peers	0.870	6
Information Seeking	0.722	6

Table A-2
Descriptive Statistics of Age and CGPA

Variables	Minimum	Maximum	Mean	Std. Deviation
Age	17	35	21.27	2.378
CGPA	1.70	4.00	3.209	0.396

Table A-3
Normality

Shapiro-wilk		
Statistic	Df	Sig.
.835	400	.000

Table A-4
Kruskal Wallis

	Sum-factor
Chi-square	5.388
Df	6
Assymp sig.	.495

Table A-5
Measure of Goodness of Fit of Model for Confirmatory Factor Analysis

Factor/Variable	**$\chi2$**	**d.f**	**p-value**	**$\chi2$ / d.f**	**GFI**	**AGFI**	**RMSEA**
Library	88.14	27	0.000	3.2	0.95	0.91	0.07
Internet	74.65	14	0.000	5.3	0.94	0.88	0.112
Classroom	85.52	14	0.000	6.10	0.93	0.87	0.121
Emotions	88.70	5	0.000	17.7	0.92	0.76	0.20
Peers	99.38	9	0.000	11.04	0.91	0.78	0.17
Information Seeking	52.66	9	0.000	5.8	0.95	0.89	0.112
Recommended				**≤ 3**	**≥ 0.90**	**≥ 0.90**	**≤ 0.08**

Structure Equation Modelling

Table A-6
Estimates for Structure Equation Modelling

Variables	**Estimates**	**P value**
Classroom ← Internet	0.646	0.000
Emotions ← Classroom	0.747	0.000
Peers ← Classroom	0.391	0.001
Peers ← Emotions	0.423	0.000
Information Seeking ← Internet	0.398	0.000
Information Seeking ← Emotions	0.42	0.000
Information Seeking ← Peers	0.131	0.004

Table A-7
Measure of Goodness of Fit of Model for structure equation model

Model	**$\chi2$**	**d.f**	**p-value**	**$\chi2$ / d.f**	**GFI**	**AGFI**	**RMSEA**
SEM	216.937	83	0.000	2.6	0.932	0.902	0.064
Recommended				**≤ 3**	**≥ 0.90**	**≥ 0.90**	**≤ 0.08**

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 493-498

PHYSICAL AND MENTAL HEALTH OF BETA-THALASSEMIA PATIENTS AT SUNDAS FOUNDATION SIALKOT

Sana Shakeel, Zaheer Abbas and **Wajeeha Khalil**
Department of Statistics, University of Gujrat
Email: zaheerabbas@uog.edu.pk

ABSTRACT

Thalassemia is inherited claret disorder in which body make abnormal form of hemoglobin. Beta thalassemia is thalassemia major it can be defined as if one or more genes of beta globin are faulty in a person than Beta-thalassemia occurs. Some clinical factors and demographic criteria's affect the physical and mental state of patients. This study is carried out to measure Physical and Mental health of such patients. Data collected from two hundred (200) patients those are treated at Sundas Foundation Sialkot through purposive sampling by using SF-36 Questionnaire.

Results explore that most of the patients' physical and mental health is at medium level because, thalassemia is fatal disease that does not cure properly. Independent sample t-test reveals that mental health is significantly affected by sibling status of thalassemia, history of thalassemia and other claret disease splenectomy on the other hand Physical health is not affected by them except splenectomy. Association explains that if physical health improves than mental health is also improved. Odds ratio illustrate that there is nine times higher possibility that if parents do not have disease than there is no history of disease in their siblings. The results of Binary logistic regression indicate that there is sufficient evidence that age in years and blood transfusion rate have negative impact on QOL, but splenectomy has positive effect on QOL.

KEYWORDS

Thalassemia, hemoglobin, beta thalassemia, quality of life, physical health, disease

1. INTRODUCTION

Thalassemia disease widely spread throughout the world, it is more compulsory to aware people with the awful disease. Firstly, it is compulsory to know its history.

Historical Background of Thalassemia

Cooley and Lee in (1925) defined a serious form of anemia with splenmegaly and bone changes. It is originally called as thalassemia anemia. It is compressed to thalassemia from "the sea" by Whipple and Bradford in (1932). In (1925), Riettl explicated a mild form of hemolytic jaundice; it is the disorder in which the red cells increased osmotic resistance. First proof of Cooley's anemia is resolute by Caminopetros in (1938) and Neel in (1950) gave idea that this is a homozygous state of a latent trait resulting in decreased intracellular hemoglobin contented (Hypochromic) and small sized red cells (Microcytosis). Hereditary disorder of hemoglobin is the monogenic disease which was defined by weather all in (2000).

Thalassemia

This disease is inherited blood sickness now which body produces nonstandard form of hemoglobin. Hemoglobin is protein molecule in red blood cells that carries oxygen. If there is abnormality in hemoglobin levels then it became cause of anemia. Thalassemia patient's bone marrow does not able to produce enough healthy hemoglobin or red blood cells. (Victore et al., 1999). Alpha and beta are two main types of thalassemia

Alpha Thalassemia

A person need four genes to make alpha globin protein, two on each chromosome 16, he get two from each parent. Alpha thalassemia develops in a person if any one or more than one genes is missing (Akhtar, 2016). If a patient has one faulty gene than he has no sign of anemia. This condition is known as "Silent carrier state". If a patient has two faulty genes, then he has no symptoms of thalassemia. We can say that this is called thalassemia minima". In this condition patient has hemoglobin H disease. This kind of patient has need regular blood transfusion. Alpha thalassemia major is the most unadorned form of alpha thalassemia. It is a severe condition in which fluid accumulates in part of "fetus" body. It becomes cause to produce abnormal Hemoglobin. Even blood transfusion is not play role in life of such kind of thalassemia patient (Akhtar, 2016).

Beta Thalassemia

Beta globin chain make by two globin genes that can be defined in such manners that if one or more genes of beta globin are faulty in a person than Beta thalassemia occurs (Akhtar, 2016).

Beta Thalassemia Minor

Beta thalassemia minor occur when one gene is faulty. A person does not see any problem if he has this kind of thalassemia (Akhtar, 2016).

Beta Thalassemia Major

Beta thalassemia major occurs if two genes are faulty. It is also called Cooley's anemia because it was described by Cooley in 1925 (Gardenghi *et al*., 2010).

In beta thalassemia condition alpha globin chain increased in body and it became cause to decrease beta globin chain. It increases synthesis of hemoglobin without beta chain. Severe kind of anemia is the core signs of beta thalassemia major that is due to ineffectual erythropoiesis, hemolysis and hypersplenium. Erythrocyte precursor's destruction moves to reduction in erythrocyte which release into the blood stream. Hemolytic action in the reticulo-endothelial systems (RES) rises due to the damage of addition bodies containing mature erythrocytes, leading to hyperplasia of the RES, erythroid system, and extramedullary hemopoeisis (occurs in the liver and spleen) (Mahmoud, 2015).

Thalassemia Treatment and its Complications

Medical advancement helps to increase survival rate of thalassemia patients. It comes to know that patients can alive more than 4 or 5 years with blood transfusion and chelating therapy. But frequently blood transfusion became the cause of iron overload and patient has faced many complications because of it. They face lot of behavioral and neurotic problems, growth failure and liver disease etc.

To decrease iron in Iron overload patients Iron Chelation Therapy (ICT) is used. Desferoximation drug is used as ICT. It is effective in reducing hepatic and extra hepatic adverse effect. Desferasirox are also available for this purpose but it is more expensive.

Objective of Study

The following are the aims of this study:

a) To evaluate current physical health and mental health of thalassemia patients.
b) To check relationship between Physical and Mental Health.
c) To find out difference between physical and mental health level on the basis of categories of sibling status, family history and blood transfusion rate.
d) To find amount of risk in siblings of the patients due to family history.
e) To explore Quality of Life of thalassemia patients.
f) Impact of blood transfusion rate and removal of spleen on the quality of life.

Significance of Study

This study will help to thalassemia patients and the role of Sundas foundation in this field. It will help in improving health of patients. It also helps to know blood transfusion rate and removal of spleen are highly effect the patient life in positive and negative manner.

2. LITERATURE REVIEW

Jafari *et al*., (2008) carried out a survey on Quality of life in Thalassemia major, Reliability and validity of Persian version of SF-36 Questionnaire. The data collected from (200) thalassemia patients of center Dastgheib hospital. They included Physical functioning, role physically, bodily pain, general health, vitality, social functioning, role emotional, mental health scales as well as some demographic factors in study. The result determined that there had no significant association of age and gender on all of eight scales that used to measure QOL. Each scales had low correlation with each other.

Jameel *et al*., (2016) conducted a study on the Compromised Quality of Life in β-Thalassemia Major Children in Nonurban Setup in a Developing Country in Day care Thalassemia center. Data was collected from transfusion dependent patient’s (35) males and (15) females (7-18 years) of urban and rural areas. Chi square test demonstrate the significance of association between demographic and clinical data. The result explores that maximum of them were either uneducated or under primary education (92%).

Consanguinity of parents was mentioned in (88%) of patients. Females achieved low scores as compared to males in virtually all the parameters.

Ansari *et al.,* (2014) conducted a study on Quality of life in victims of thalassemia major. In study, (250) patients of Ali-Asqar hospital and Zafar clinic of Tehran Iran and (51) healthy people included. Researcher used six dimensions to measure health (overall health, physical, psychological, social and environmental relationship) of patients and some demographic information also got by patients. Result revealed that six dimensions of quality of life was greater in control group and better QOL scores associated with age and higher education level. Iron chelation therapy and other hand cardiac disease, hepatitis and history of any blood disease associated with QOL.

3. MATERIAL AND METHODS

This study used quantitative approach to measure physical and mental health of beta Thalassemia patients. Study explore that the clinical and demographic factors affected on thalassemia patient's life.

Sample Selection Strategy

Purposive sampling technique is used, because all patients are not available at Sundas foundation all the time. According to purpose, I got information from those patients that are present in clinic at the time when I visit the clinic. I got information from two hundred patients.

Questionnaire

The S-F 36 form used to measure physical and mental health of thalassemia patients. This questionnaire contains eight dimension of QOL which relates to physical and mental health.

The domains 'physical functioning' (10 objects), 'Role physical' (4 objects), 'bodily pain' (2 objects), and 'general health' (5 objects) are more related to physical health. The domains 'vitality' (4 items), 'Social functioning' (2 items), 'role emotional' (3 items) are more related to measure mental health. There are 40 scales used to measure QOL of patients but in this questionnaire eight scales provide more information about QOL of patients.

Data Analysis

- Visual Binning
- Descriptive Statistics
- Association using Kendall's Tau b
- Independent Sample Test (Mann Whitney and Kruskal Wallis test)

Relative Risk and Odds Ratio

- Recode Data (General heath Rating Index)
- Binary Logistic Regression

4. COMMENTS AND CONCLUSION

Thalassemia disease is one of the major diseases that affect physical and mental fitness of patients and QOL of patients depends on both. I conclude that this questionnaire is more reliable in finding the QOL of patients. I conclude that in rural area, more people affected by this disease and they do not have awareness about it. Results explore that most of the patients physical and mental health is at medium level because, thalassemia is fatal disease that does not cure properly but Sundas foundation provide blood to the patients that helps patients to live life in a specified manner. Stages of physical health are related with levels of mental health. It shows that if levels of physical health are increased than mental health’s stage is also increased.

Overall physical health is same across the categories of sibling status of thalassemia and family history of thalassemia while their overall mental health is different among the categories of sibling status of thalassemia and family history of thalassemia. Both Physical and mental health are different among the categories of rate of blood transfusion rate and splenectomy which means that physical and mental health of those whose spleen was not removed is different from those whose spleen was not removed.

There is sufficient evidence that age in years, blood transfusion rate and splenectomy has effect on QOL.

REFERENCES

1. Akther, R. (2016). *The value of different red cell parameters in the diagnosis of microcytic hypochromic anaemia* (Doctoral dissertation, BRAC University).
2. Aljeesh, Y.I. (2016), Quality of life among thalassemia children patients in the Gaza strip. *American Journal of Nursing Science*, 5(3), 106-113.
3. Alzahrani, R.A., Almutairi, O.M., Alghoraibi, M.S., Alabdulwahed, M.S., Abaalkhail, M.K., Alhawish, M.K. and Alosaimy, M.T. (2017). Quality of life in transfusion-dependent thalassemia patients. *Journal of Taibah University Medical Sciences*, 12(5), 465-470.
4. Amani, F., Fathi, A., Valizadeh, M., Farzaneh, E. and Fattahzadeh-Ardalani, G. (2017). Quality of life among Ardabil patients with beta-thalassemia major. *International Journal of Research in Medical Sciences*, 3(11), 3308-3312.
5. Ansari, S.H., Baghersalimi, A., Azarkeivan, A., Nojomi, M. and Rad, A.H. (2014). Quality of life in patients with thalassemia major. *Iranian Journal of Pediatric hematology and Oncology*, 4(2), 57-63.
6. Azarkeivan, A., Hajibeigi, B., Alavian, S.M., Lankarani, M.M. and Assari, S. (2009). Associates of poor physical and mental health-related quality of life in beta thalassemia-major/intermedia. *Journal of Research in Medical Sciences: The Official Journal of Isfahan University of Medical Sciences*, 14(6), 349-355.
7. Baraz, S., Miladinia, M. and Mosavinouri, E. (2016). A comparison of quality of life between adolescences with beta thalassemia major and their healthy peers. *International Journal of Pediatrics*, 4(1), 1195-1204.

8. Caocci, G., Efficace, F., Ciotti, F., Roncarolo, M.G., Vacca, A., Piras, E., Littera, R., Markous, R.S.D., Collins, G.S., Ciceri, F. and Mandelli, F. (2012). Health related quality of life in Middle Eastern children with beta-thalassemia. *BMC Blood Disorders*, 12(1), 6.
9. Gardenghi, S., Ramos, P., Marongiu, M.F., Melchiori, L., Breda, L., Guy, E., Muirhead, K., Rao, N., Roy, C.N., Andrews, N.C. and Nemeth, E. (2010). Hepcidin as a therapeutic tool to limit iron overload and improve anemia in β-thalassemic mice. *The Journal of Clinical Investigation*, 120(12), 4466-4477.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 499-510

AN EVALUATION OF THE PERFORMANCE OF VOCATIONAL TRAINING INSTITUTES BEING RUN BY THE PAKISTAN BAIT-UL-MAAL

M. Rafiq Khan[§] and **Mirza Fahd Haroon**
Lahore School of Economics, Lahore, Pakistan
[§]Corresponding author Email: drrafiq@lahoreschool.edu.pk

ABSTRACT

The work presented in this article was undertaken to evaluate the performance of Vocational Training Institutes being run by the Pakistan Bait-ul-Mal (PBM). The requisite data were collected from both secondary and primary sources. The secondary data includes the data from Pakistan Bait-ul-Maal on the budgets, number and addresses of the beneficiaries and locations of the vocational training institutes. The primary research data includes the responses gathered through questionnaires served to the respondents including management, existing students and passed out students. The primary data collected was analyzed using SPSS software which helped in running the regressions to show the relationships between the dependent and independent variables involved in this study that indicated the performance of vocational training institutes under the control of PBM. The analysis also displayed frequencies of independent variables as frequency tables which have been discussed.

KEYWORDS

Evaluation, Vocational, Institutes, Bait-ul-Mall, Pakistan.

INTRODUCTION

There are so many people who are educated but sitting idle because they do not have job or some work to do and lack resources to study to make their lives better. Being a part of this world, they also are entitled to all the rights which others have. Thus, something must be worked out for them. Luckily, vocational training institutes are promising places to address their problem because these institutes provide technical education to those who are out of job and wasting their skills to utilize their abilities in the process of national development. These institutions employ job seeking girls and poor and needy women such as widow, etc. They offer special courses including embroidery work, textile printing, tailoring, pen work, teacher's training, handicraft making, designing, interior decorating, computer training, cooking, baking and others. Such training centers have created interest among women due to their conducive work environment and quality of work they impart. In rural areas, these institutions are a source of relief for those who are unable to get higher education from colleges and universities. Government is playing a very vital role for the effective work of these institutions and has opened many centers which are providing assistance to a significant number of people. A large number of girls are learning some skills and utilizing them to earn their livelihood and this is a major contribution to the women empowerment. This activity also tends to increase production in the country because people are engaging in some kind of work that adds value to the national output.

Pakistan Bait-ul-Maal provides financial assistance and rehabilitation to the deserving, poor, needy, orphans, widows and students. It also provides assistance to registered NGO’s which are working for the development of social projects and thus are a source of income for many deserving people. Government has been a major provider of funds to PBM. PBM also gets huge amount of Zakat that is the compulsory tax on Muslims at the rate of 2.5% per annum of the value of the ownerships beyond their essential needs.

PBM is presently running five important social welfare programs: National Centers for the Rehabilitation of Child Labor (NCRCL), Institutional Rehabilitation, Individual Financial Assistance Program, Food Support and Vocational Training Schools. The task of evaluation of performance of all five was undertaken by Khan et al in 2008. Results of two evaluations Khan and Shaukat (2009) and Khan and Ali (2009) have been published and processing of articles from remaining three is on the way

Although no evaluation of PBM programs seems to have been undertaken in Pakistan, some studies highlighting the importance of vocational training with reference to the development and social welfare in the developing countries have been conducted.

Wanie, (1963) described VTIs to understand economic growth in the developing countries and stated that an upgrading of the education was essential to achieve economic targets and highlighted that in majority of the developing countries there were no proper technical personnel available. H concluded that the whole process of uplifting education in the region would be relevant to the cultural and ultimate economic and social objectives of the country. William (1970) discussed the impact of different training programs on wages. While discussing vocational training, he remarked that the returns to vocational training did depend on the program size. Freeman (1974) had discussed about the occupational training in proprietary schools and technical institutes and their impact on wages. In his opinion, these trainings had a positive effect on wages. Robins. (1987) demonstrated the relationship between various economic out-Comes and vocational training in high school for those who had completed exactly twelve years of schooling. According to the researcher the relationships between economic outcomes and vocational training in high school for young people who had completed exactly twelve years of schooling appeared to be robust. Veum (1995) described the wage impact of various training sources and Parent (1999) collected statistical data from the National Longitudinal Survey of Youth (NLSY). The result of employer-provided training was studied in the article by analyzing the impact of the aforementioned training on the wage profile and on the mobility of the young workers. It was concluded that training with an employer had a definite positive effect on the wage. The work reported in this article is on the assessment of the performance of vocational training institutes under PBM.

The main goal for carrying out this research was to evaluate the performance of vocational training institutes and to know what is needed to be done in case if these schools if these are not being properly run. The objectives of research were as follows:

1. To sort whether the students are getting benefits which are told by PBM
2. To check the effect of the distance of the school on number of applicants
3. To check the effect of training on the literacy rate

4. To check whether the income of the parents affects the number of applications to these schools.
5. To determine the after training financial benefits received by all the applicants.

Different hypotheses framed are listed in the Result section.

METHODOLOGY OF RESEARCH

Different elements of the plan of research work are described below:

Sources of Data:

Data for this research were collected from both secondary and primary sources.

Primary Data:

A questionnaire was designed to collect the primary data. The sample population was interviewed and their responses were recorded. A separate questionnaire was designed for administration, existing students and passed out students. The addresses of passed out students were obtained from PBM and the data were collected by visiting scheduled respondents. It was a very hectic work since the target population was spread all over Lahore. The questionnaires for the administration were filled up by the PBM officials at their main branch and vocational training institutes. It made it easier to compare the responses of students and administration. Data from the existing students were collected by visiting different schools, distributing questionnaires and helping them to understand it.

Secondary Data:

The secondary data were collected from the regional office of Bait-ul-Maal which is situated in Lahore. The data which were taken from the Office included the addresses of the students, budget allocation for vocational training institutes, policies of this program and the enrollment system at the schools. The website of Bait-ul-Maal and other websites were also visited for getting the information regarding the programs which were being run by it to get more information about vocational training institutes. The secondary data also included the articles and previous studies related to this research which helped to make things easier for carrying out the research study.

Population:

The population included all those students who were currently studying and those who had passed out.

Sampling:

The data of all the students were not available, so only 40 students were taken and continued the research.

Variables:

The variables which include both dependent and independent are listed below with a little explanation:

***Dependent Variables*:** The dependent variable was Number of Applicants. This variable measured all the students who applied for the school.

Independent Variables**:** These are illustrated below

- Income of Parents**:** The income of parents which affects the decision to get enrolled in a school was an important variable. If the income is enough, then the applicant will not apply for the schools or vice versa.
- Distance to School from Residence**:** It is the distance from the residence of the applicant which could be very important for deciding whether to go to school or not.
- Previous Education Received by an Applicant**:** This is the qualification which the applicant had achieved before. If an applicant had got a less education then he would go for it.
- Application Procedure**:** If the application procedure was complicated and lengthy then the number of applicants would be less.
- Cost Incurred**:** This was the cost incurred by the applicants while coming to schools to get the training. It included traveling and other expenses.
- Age of the Applicants**:** It was the age of an applicant which is required for the training.
- Nature of Course**:** There were different types of courses which might affect the applicant's decision. It included short term and long term courses.
- Variety of Courses**:** It was the variety of courses which each school offered to students. If it offered different and well worth courses then the number of applicants would increase.
- Student/Teacher Ratio**:** This variable showed the number of students taught by 1 teacher. If it was less, then one could say that it would help to get better training because a teacher would be able to concentrate on a smaller strength. If this ratio is high, it would cause training to go down and applicants would be reluctant to join the centre.
- Number of VTI**:** This variable meant the total number of VTI in a particular area. Higher the number of vocational training institutes, more training was provided to the people. This would help to increase the student base and provide training to many people which would improve the lives of many people when they would earn after getting training. It had a positive relation to the dependent variable.
- Funds provided by the Government**:** This variable measured the amount allocated by the government for the opening of these schools.
- Official Behavior**:** This variable showed the cooperation from the officials while getting admitted.

Data Analysis

The data after tabulation in MS Excl were analyzed by running an SPSS software program that displayed the results of analysis in the form of regression and frequency tables.

Interpretation of Results

The general information gathered the interviews, and other PBM sources were compiled as results of descriptive research. The data collected through primary research was interpreted using the statistics technique SPSS which made it possible to know the correct relationships and achieving the goals.

RESULTS

The analysis of the result has two parts. The information collected through the interviews of the managements concerned is reported as results of qualitative analysis. Those obtained by the application of statistical methods are reported as results of quantitative analysis.

Result of Qualitative Analysis: As told by the officials, the VTIs have been established throughout the country including the northern areas with the objective to reduce poverty. PBM has pledged to open at least one vocational school in every district. The current strength of these schools as mentioned in PBM Journal is 141 operating on two shift basis. Currently, 8358 people are getting training from these schools and around 37,200 have passed. The yearly budget of each school is Rs.925, 000 and a total amount of Rs.288.2333 million has been spent for these schools. Vocational training is multi-sect oral in nature. The officials were of the opinion that for a nation whose population has crossed 160 million, there cannot be a better source of living than self-employment. These training courses potentially provide women with the opportunity to redress the inequalities, which they currently suffer in employment. The courses would provide them a chance to enter in traditional and non- traditional occupations which will make them capable to earn. This will help eradicate poverty and current inequalities will disappear. In this way, the women of Pakistan will have a high quality of life.

One of the respondents from a school told about the activities which were being carried out and the facilities available at the school. She was asked the amount of budget which the school was getting each year in order to know whether the officials at the PBM Head Office were correct. She showed the budget book printed by the PBM and all the expenses incurred and praised all the people working in this field.

Results of the Quantitative Analysis: This analysis includes the graphs, frequency tables, explanation and regressions.

Results of Frequency Analysis**:** The result of the frequency analysis is shown below with the help of tables.

Table 1
Frequencies of the Income of the Parents

	Income	**Frequency**	**Percent**	**Valid Percent**	**Cumulative Percent**
Valid	Less than Rs. 3000	4	10	10	10
	Between Rs. 3000 and 6000	20	50	50	60
	Over Rs. 6000	16	40	40	100
	Total	40	100	100	

Table 1 indicates that 10% parents have a family income under Rs. 3000, 50% have from 3000 to 6000 and 40 % have more than Rs.6000. This shows that people send their children to schools run by PBM to get skills and earn something.

Table 2
Frequencies of Distance Travelled

	Distance from Residence to School	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Less than 10 km	12	30	30	30
	Between 10 to 30 km	14	35	35	65
	Between 30 to 50 km	8	20	20	85
	Over 50 km	6	15	15	100
	Total	40	100	100	

Table 2 indicates that 30% of the students travel less than 10 km to get to the school whereas, 35% travel between 10 to 30 km and 15% travel more than 50 km which shows that students come even from far off places in order to study in schools.

Table 3
Frequencies of Official Help

	Official Help	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	28	70	70	70
	No	12	30	30	100
	Total	40	100	100	

Table 3 indicates that 70 % of the students think that the officials are helpful and 30% say that they are not which is not desirable but overall situation seems to be satisfactory.

Table 4
Frequencies of Age of the Applicants

	Age of Applicants	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Less than 15	12	30	30	25
	Between 15 and 25	20	50	50	70
	Over 25	8	20	20	100
	Total	40	100	100	

Table 4 indicates that 30% of the interviewed students were under 15 years of age while 50% were between 15 and 25 and 20% were over 25 years. This shows that the students of all ages are getting education without discrimination of gender.

Table 5
Frequencies of Opinion about Ease of Enrollment Process

	Ease of Enrollment	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	24	60	60	60
	No	16	40	40	100
	Total	40	100	100	

Table 5 indicates that 60% of the students say that the enrollment process was easy as they did not find any major problem while applying there, whereas, 40% say that they faced some problems that need to be solved.

Table 6
Frequencies of Cost Incurred

	Cost Incurred	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Between Rs. 200 to 500	26	65	65	40
	Between Rs. 500 to 1000	10	25	25	80
	Over Rs. 1000	4	10	10	100
	Total	40	100	100	

Table 6 indicates that 65% of the total students incurred between Rs.200 to Rs.500 while getting enrolled. 25 % incurred between Rs.500 to Rs.1000 and 10% incurred above Rs.1000.

Table 7
Frequencies of Spending in School

	Spending in School	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	30	75	75	75
	No	10	25	25	100
	Total	40	100	100	

Table 7 indicates that 75% of the students said that they had to buy certain things required at work whereas 25 % said that they did not have to buy anything as they were provided all the required material

Table 8
Frequencies of Opinions about Length of Program Term

	Opinion about Length of Term	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Long term	28	70	70	70
	Short term	12	30	30	100
	Total	40	100	100	

Long term is a course which lasts for more than 6 months whereas short term lasts for less than 6 months. Table 8 indicates 70% joined the school to study long term courses while 30% joined it for short term courses.

Table 9
Frequencies of Financial Benefits after Getting the Training

	Financial Benefits to Students	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	26	65	65	65
	No	14	35	35	100
	total	40	100	100	

Table 9 indicates that 65% think that they will get financial benefits after getting training from these schools and 35% say that they are not confident to get any financial benefits as they will be getting married or staying at home.

Table 10
Frequencies of secondary education (SE) after training

	Continuity of SE	Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Yes	18	45	45	45
	No	22	55	55	100
	Total	40	100	100	

Table 10 indicates that 45% of the students will continue secondary education and 55% think that they wont be able to continue studies.This shows that some students want to persue their education but due to some problems they wou;d not be able to do so.

Results of Regression Analysis**:** The multiple regression was run by keeping the number of applicants as dependent variable and some of the variables listed below as independent variables. The regression could not be run by using all the variables (Frequency analysis). The regression equation is given below.

$$\mathbf{Y = \alpha + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) \ldots..}$$

where Y is the dependent variable and **X** stands for the independent variables.

Table 11
Variables and their Description

Variables	Description
X_1	Income
X_2	Financial benefits
X_3	Distance from home
X_4	Cost incurred
X_5	Age
X_6	Secondary education

The results are shown in Table 12 to 14

Table 12
Model Summary

Model	R^2	Adjusted R^2	Std. Error of the Estimate
1	0.999	0.998	.378

Table 13
Coefficients

Model		Estimate	Std. Error	t- Statistic	P- Value
1	(Constant)	-15.804	0.796	-19.831	0.0025
	X_1	8.071	0.755	10.678	0.0087
	X_2	-3.713	0.729	-5.086	0.0365
	X_3	1.865	0.745	2.501	0.1295
	X_4	4.843	0.869	5.567	0.0308
	X_5	3.300	0.541	6.029	0.0259
	X_6	1.919	0.646	2.967	0.0973

Table 14
ANOVA

Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Model	959.713	6 or 5 (check)	159.952	1114.67	.0009
	Residual	0.286	2	0.143		
	Total	960	8			

R-square = 99.9701 percent
R-square (adjusted for Df) = 99.8804 percent
Standard error of estimate = 0.378811
Mean absolute error = 0.13154
Durbin-Watson statistic = 2.45179

The output shows the results of fitting a multiple linear regression model to describe the relationship between number of applicants and 6 independent variables. The equation of the fitted model is

No. of applicants = –15.8049 + 8.07175 * Income – 3.713 * Financial Benefits + 1.86547 * Distance from home + 4.84305 * Cost incurred + 3.30045 * Age + 1.91928 * Secondary education

Since the P-value in the ANOVA Table 14 is less than 0.01, there is a statistically significant relationship between the variables at the 99% confidence level.

The R-Square statistic indicates that the model as fitted explains 99.9701% of the variability in number of applicants. The adjusted R-square statistic, which is more suitable for comparing models with different number of independent variables, is 99.8804%. The standard error of the estimate shows the standard deviation of the residuals to be 0.378811.

The mean absolute error (MAE) of 0.13154 is the average value of the residuals. The Durbin-Watson (DW) statistic tests uses the residuals to determine if there is any significant correlation based on the order in which they occur in the data file. Since the DW value is greater than 1.4, there is probably not any serious autocorrelation in the residuals. In determining whether the model can be simplified, notice that the highest P-value of the independent variables is 0.1295, belonging to distance from home. Since the P-value is greater or equal to 0.05 so it must be excluded from the model.

DISCUSSION

The objective of the work reported here was to evaluate the performance of vocational training institutes (VTIs) being run by PBM. The results of the study may be quite helpful to assess the performance of VTIs and know what can be been done further to improve their working.

Both qualitative and quantitative results may be integrated to derive the true picture. It sounds pertinent to quote qualitatively that a lady was reluctant to answer some questions during the enquiry. Thus it could be assumed that she feared his seniors at the PBM Head Office and she was hiding the correct answers to save her position. When asked about the facilities which were being provided to the students there, she told that they had all the required facilities. She took the researchers round the Institute particularly to the rooms where girls were working and where all the equipment was placed. Some of the class rooms did not even have fans. The girls were working in hot and humid conditions and some of the equipment was also not in the working order. When she was asked about the reason, she answered that the requests had been sent to the government; so that would take some time to solve all these problems. She did not even know the budget allotted to them. She opened the PBM book and read it out and also read answers to other questions from the book.

It was felt that the Director PBM had a fairly good knowledge about the work she had been doing. She answered all the questions without hesitating and kept praising the government for its role for the betterment of women in Pakistan. She knew all the budget details of VTI’s and all the works which were being run further to improve their performance. She showed the new projects which were being undertaken. She had been working for the last 5 years and had a lot of knowledge about all the welfare programs. She helped to understand the main problems which VTI’s were facing and what was being done and what should be done in future.

The regression analysis is of great help to establish a statistical relationship between the dependent variable that is the number of students who joined the institutes and the independent variables listed in Table 11 that were factors which affect it. The attempt to run regression with independent variables revealed that the regression could not be run with all variables hence; it was run with 6 variables only. The analysis revealed that all variables had different impact on the dependent variable. For example, if we look at the income variable, it shows that if income of parents increases by 1 %, the chances for a child to go to vocational training institute increase by 8.07. If the financial benefits increase by 1 %, the chances for a child to go to VTI decrease by 3.713. The distance from home variable has different impact on dependent variable. It shows that if distance from home increases by 1 km then the chances for a child to go VTI increases by 1.86. Similar is the case with other variables. They have a positive relationship with the dependent variable as if they increase, the dependent variable also increases. When the result was drawn, 4 out of 6 hypotheses proved to be true which showed that the variables had a significant effect on the dependent variable. The results show that the value of R square is 99 % which shows that 99% of changes in dependent variable are caused by the independent variables. Since all the variables have a low significance value, they affect the dependent variable more. The study shows that PBM and Government of Pakistan are working well for the welfare of the civil society through opening and running of VTIs. The qualitative observations revealing some negativity of the program when integrated with quantitative results based on statistical frequencies and regression analysis presented a satisfactory picture about the performance VTI Program of PBM. As there is always a room for elimination or minimization of weaknesses and deficiencies encountered during the enquiry, these may be overlooked and authorities may be suggested to address negativity by pinpointing its existence anywhere.

The researchers were convinced that VTIs are performing well but there is still a lot of work to be done for the betterment of women in our society. The projects like this allow women to work and earn their own livelihood and reduce burden on their families. Thus the range of these projects may be extended by PBM and Government of Pakistan.

CONCLUSION

The research study was conducted to evaluate the performance of vocational training institutes being run by PBM. The results show that these institutes are working well and the policies followed by PBM in context of VTIs are very effective. The government, of course, needs to do much more for the welfare of the society and should allocate more funds. It has to give them all the facilities which are available at private institutions. That would help create competition among them which will improve the standards. It has to open more institutes like VTIs where the deserving girls and women may come to learn skills and use them to earn their livelihood.

REFERENCES

1. Freeman, R.B. (1974). Occupational training in proprietary schools and technical institutions. *The Review of Economics and Statistics*, 56(3), 310-318.
2. Khan, M.R. and Ali, U. (2009). An Analysis of Individual Finance Assistance Program under the Pakistan Bait-ul-Maal. Proceedings of *Fifth International Conference on Statistical Sciences: Mathematics, Statistics and Applications* Organized by Islamic Countries Society of Statistical Sciences in Lahore on January 23-25, 347-353.
3. Khan, M.R. and Shaukat, A. (2009). Evaluation of the Food Support Program of Pakistan Bait-ul-Mal, Proceedings of *Fifth International Conference on Statistical Sciences: Mathematics, Statistics and Applications* Organized by Islamic Countries Society of Statistical Sciences in Lahore on January 23-25, 2009, 339- 345.
4. Pakistan Bait-ul-Mal Official Website www.pbm.gov.pk
5. Parent, D. (1999). Wages and Mobility. *Journal of Labor Economics*, 17(3), 298-317.
6. Robins, P.K. (1987). Training Programs and wages. *The Journal of Human Resources*, 22(1), 113-125.

7. Veum, J.R. (1995). Sources of training and their impact on wages. *Industrial and Labor Relation Review*, 48(4), 812-826.
8. Wanie, W.J. (1963). The role of education in the development of underdeveloped countries. *The Canadian Journal of Economics and political Science/Revue canadienne d'Econoque de science politique*, 29(4), 437-445.
9. William, R.S. (1970). Nonwage benefits of vocational training. *Industrial and Labor Relations Review*, 23(4), 500-509.

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 511-529

APPLYING ISM ON NON-TARIFF BARRIERS FOR THE INVESTORS OF OIL AND GAS SECTOR IN PAKISTAN: CPEC PERSPECTIVE

Ifra Aziz Khan Niazi[§], Amna Hassan Khan, Zabihullah Nishat, Aqila Yousify and **Zain Mustafa**
COMSATS University Islamabad, Lahore Campus, Lahore, Pakistan
[§]Email: ifraniazi460@yahoo.com

ABSTRACT

There are the chances of increasing investments in Oil and Gas sector of Pakistan after CPEC (China Pakistan Economic Corridor) and the foreign investors are going to face non-tariff barriers, that government usually impose on foreign investors to restrict trade or to alter the prices and policies of products to limit the trade, to the significant extent. The purpose of this study is to find out the relationship, dependability and linkage of those non-tariff barriers, and to classify and prioritize them according to their importance with respect to their driving-dependence relationship. Data is collected through a matrix type questionnaire and further elaborated with the help of Interpretive Structural Modeling (ISM) technique followed by Matriced' Impacts Croise's Multiplication Applique 'ea' un Classement (MICMAC) and Structural Self-Interaction Matrix (SSIM) to validate the results. ISM model has been built using the elementary concepts of directed graph theory. Methodology identifies the most and least important factors and from the iterations, results reveal that there are five levels in the model in which factors are divided according to their driving-dependence relationship. This is a seminal study that will help policy makers, investors, government and society at large to understand the complex inter-dependent relationship between barriers and deeper understanding of the phenomenon. This data and results are based on the responses of local respondents and economists, data can be collected through real foreign investors to cross-check the results.

KEYWORDS

Interpretive Structural Modeling (ISM), Oil and Gas, Investments, Barriers, Pakistan, MICMAC, SSIM.

INTRODUCTION

Non-tariff barriers are the ways through which governments alter prices and policies of products to limit their trade. Basically, a nontariff barrier is a way to restrict excess trade. As part of political and economic policy, countries use some barriers to control the amount of trade and paying extra tax on imported goods, they conduct with other countries. These non-tariff barriers are used not only in Pakistan but in all over the world. And these are the hindrances that investors of Oil and gas sector can face while investing. These barriers particularly include Technology Barrier, Health and Safety Issue, Liability Issue, Licensing, Natural and Environmental Issue, High Costs/Investments, Lack of Confidence, Lack of Governmental Support, Terrorist Activities, Fear of Failure, Lack of Support from Local Community etc. (Roger, 1999).

China Pakistan Economic Corridor (CPEC) is project containing roads/ railways and mega energy projects through Pakistan worth billions of dollars. It has been merely three to four years since project started and the project has shown promising results. CPEC is destined to attract new investments in region and revive our mutilated economy and reduce unemployment to a significant level. Through this project, there are lots of chances to increase investments and prospective investors of CPEC projects foresee lot of barriers, especially non-tariff barriers to invest, after CPEC project, in Pakistan. This study will particularly talk about the barriers for the investors of Pakistan's oil and gas industry. The barriers in oil and gas sector of Pakistan will include environmental risks, health and safety risks, ultimate reputational risks of management, lack of confidence, technology and expertise issue, human errors, cost on society, cost of pollution, injuries and deaths, licensing and permitting requirements and policies, lack of local support, terrorist activities etc. The other non-tariff quantity control barriers include: quotas, voluntary export restraint, "by local" legislation, standards and labels, specific permission requirements, administrative delays, reciprocal requirements, counter trade, restrictions on services, essentiality defense, not for-profit services, standards and immigration. After CPEC, these barriers will be a little problematic for new investors. It has become imperative to identify and set priorities to eliminate these barriers. This study will be helpful to prospective investor, regulators and society at large to understand the complex inter-active relations among barriers. It will also help the stake holders to set policy preferences based on relative importance. It will be a seminal study since no such study has already been conducted.

This study has following objectives:

1. To identify non-tariff barriers in oil and gas sector of Pakistan, in context of investment in CPEC
2. To rank and prioritize them
3. To develop an interpretive model of their contextual relations
4. To classify them based on their driving-independence power through ISM.

ISM (Interpretive Structural Modeling) is a technique for identifying relationships among specific items or variables which define a specific problem or an issue. Finding direct and indirect inter-relations between the factors describe the situation more accurately than studying them in isolation. It starts with the identification of relative problems and solving them with the technique and goes on towards making an element set and a Structural Self-Interaction Matrix (SSIM) which is then converted to Reachability Matrix (RM). This methodology is interpretive as it decides when and where the factors are inter-related by portraying it in a diagraph model. In SSIM, group of experts' responses and after consensus, results are finalized. ISM would be applied of non-tariff barriers for the investors of oil and gas sector in Pakistan in the context of CPEC. Population will be prospective investors in CPEC and other stakeholders like government, society at large and academia. This methodology uses non-probability sampling i.e. Panel of experts. For eliciting opinions of experts, a matrix type questionnaire is used. Rest of the paper is arranged as literature review, methodology, results and conclusion.

LITERATURE REVIEW

It's pleasure to have an opportunity to contribute in the research about Pakistan's oil and gas sector and its non-tariff barriers and their relatability. World is in the need of more oil and gas reserves to meet the needs and demands of the residents in this world. Oil and gas exploration and production is majorly done by gulf countries till now but now-onwards, more exploration and investment in oil and gas sector will be needed (Insight Turkey, Vol. 4, No. 1 (January-March 2002). There are some proceedings in Pakistan's oil and gas sector and the govt. is paying heed towards oil exploration and inviting companies to invest in Pakistan's oil and gas sector. CPEC also seems to be helpful in attracting investments. Nadeem Babar, head of Prime Minister Imran Khan's Task Force on Energy Reforms, told Reuters, that there are 20-30 discoverable onshore blocks of gas in Pakistan and the government is thinking to change the shale gas and licensing policy to attract new investors (Jorgic, 12 March 2019). Babar said that there is a shortage of LNG and its expensive, so Islamabad wants foreign interested companies like Saudi Arabia's Aramco and Exxon Mobil and Russian Gazprom, to extract and invest in Pakistan (Jorgic, 2019). "A mechanism like what was done in CPEC will be developed," Babar said, referring to a 15,000-strong army division set up to safeguard Beijing-funded infrastructure projects in the China-Pakistan Economic Corridor (CPEC) (Jorgic, Gloystein and Hogue, 2019). The companies that would invest in this sector will obviously face non-tariff barriers that will be discussed ahead with their relatability with the help of ISM (Interpretive Structural Modeling) technique.

Technology is the most important barrier that will be faced by the companies investing in our country. Oil and gas require much heavier infrastructure then any industry. It needs well pads, pipelines, compressor stations, processing plants, roads, water treatment facilities, drilling plants, heavy transport, fracking techniques etc. (De Rijke, (2013). These technologies are not available in Pakistan. Enhancing technical cooperation and improving the capacity for technical renewal is one means of promoting sustainable development in the region. In the words of the Eminent Persons Group (EPG), the flow of new and high technology accelerates economic development and enhances scientific and technological capability, promotes trade and investment liberalization and reduces disparities in the level of economic development (Xiaobing and Jianglin, 2012).

Oil and Gas industry involves use of heavy machinery and materials which comes with many risks like machine hazard, ergonomic hazards, electronic and other hazardous energy risks which are critical to leading to injuries and death. Injuries and death are also a barrier for companies not only in Pakistan but also in foreign countries. "Major incidents and worker deaths at California refineries", this document estimates the probabilities of worker deaths for refineries in California. This incident leads to major impact on the state economy (Gonzales et al., 2016).

The other major risk is regarding safety which is most important for the investors. Accidents can seriously impact the smooth energy supply procedure (Doukas et al., 2011). Though the law and order situation in Pakistan has progressed in past few years but investors still feel insecure while operating in areas like Baluchistan (Jafri, 2016). The recent example of it is an attack on Chinese consulate in Karachi. Similarly, in Aug 2000, Explosion of one interstate pipeline in New Mexico in which loss was about 1.2 billion cubic feet, and in March, 2001, Tanker BALTIC CARRIER accident, with a loss of 2,700 heavy fuel tons (Doukas et al., 2011) are the examples of safety and accidental risks involved in it.

Companies also face high liability risks like product liability, financial liabilities, professional liability, political instabilities, quality assurances etc. while operating offshore (Doukas et al., 2011). Unique challenges exist in the product development and distribution and Management of the product while working in host country. And, Pakistan's business environment is difficult, the country is listed 147 in least pro-business economies in the world by world bank (World Bank, 2018) channel through which recent economic and financial crises have impacted the world petroleum markets is through consumer confidence (Sadorsky, 2001). Investors fleeing risky investments move money into safe havens like US treasury bonds and this had severe consequences for hedge funds, most notably Long-term Capital Management (LTCM) (Sadorsky, 2001).

The province of Baluchistan is rich in resources but has unskilled population due to the negligence of policymakers is an interference in using those resources to the utmost. The failure of provincial governments to improve the capacities of the population and its institutions has also considerably contributed to the underdevelopment of the province (Bashir et al., 2019). According to the same study, the development programs that were executed in past did not bring advantageous change, thus which speared disappointment between the residents of the region as their interest for more noteworthy financial improvement was not replied. To tackle the objections of people, completely different governments tried, however, all of them have failed. However, the provincial government lacks the capacity, moreover because the resistance from the social group leaders, has barred the method. Due to this, the anger of local people of Baluchistan against government or state has shaped over these matters. A number of people especially nationalist parties and youth have warned the government and state against more development in province unless they are taken into assurance and given equitable share (Frederic, 2006). For the last eight years the Baloch nationalists have been actively involved in an armed struggle against the government of Pakistan, which in turn has been trying to suppress the rebel movement by various means (Aslam, 2011). And, companies may need to hire specialized local support to make sure compliance. Special efforts to simplify and improve the simplicity of such administrative requirements could meet regulatory objectives while encouraging investment in the energy sector (Gonzales et al., 2016).

Different natural and environmental issues are related with the generation of gas and oil for example palm oil cause issues including deforestation, biodiversity misfortune, ozone depleting substance emanations and change of high preservation esteem and peat land. The accidents in energy supply that occur due to extreme weather conditions are not negligible. With reference to the natural gas market, IEA mentions the examples of Canada during the winter of 1992–1993 and the USA in January 1994 (IEA, 1995). The province of Baluchistan has peculiar natural and environmental issues like tropical climate, hills, scarcity of water etc. Social issues incorporate concerns with respect to lawful and conventional land use rights, land securing, and treatment of nearby and indigenous networks (Gazdar et al., 2007). As of late, development of oil palm has developed exponentially which has brought about expanded consideration on the natural and social supportability of the division, especially in Indonesia and Malaysia. (Page 8 Version 6.0 March 2015 © Barclays Bank PLC).

In Pakistan, the monetary achievement has always been gained at the cost of environment. Since all production and manufacturing units don't consider how much environmental harm they are causing. The environment has been affected adversely in the past few years. On the other hand, the consumption of Oil and Gas in Pakistan has increased by 9.64 percent in the year 2016-17. The major drivers of increased oil and gas consumption are energy sectors and transportation companies and Pakistan's demand for oil and gas is expected to double in next 10 years which is a challenge for Pakistan as our own natural resources are not enough to meet the needs (Arifeen, 2018).

The other issue is human error. Human error is an improper decision or behaviour of any worker in the workplace which leads to a negative impact on the workplace (Alkhaldi et al., 2017). The worker working in the company of gas and oil also make error which cause and accident such as loss of lives, harms property and condition, and hampers the monetary exercises of any working environment. Heinrich is a leader of Industrial Safety Engineer (Mattia, 2013; Patel, Sherratt and Farrell, 2012) and he described human error as a poor behavior or an inadequate risk perception (Patel et al., 2012). Zhu and Xiao Ping (2009) found from the perspective of human mind that human error happens due to the artificial mistakes, whether intentional or unintentional. As well as, Zeng et al. (2008) claimed that human error that potentially causes an accident can be referred as an unsafe act that depart from hazard control or job procedures to which the person has been trained or informed and in turn this act leads to unnecessary exposure of a person to hazards.

The Pakistani government makes it difficult for investors to come and invest here. Oil and gas agreements and the associated fiscal rules establish the "price" of the resource in terms of the bonuses, royalties, taxes or other payments the investor will make to the government over the life of the project (Sunley et al., 2003). They impose on them certain taxes and duties which is why they don't want to invest here. Whereas, if the oil and gas tax regime is more onerous than the standard tax regime, the taxpayer could seek to have certain project-related activities treated as down-stream activities outside the ring fence (Sunley et al., 2003). Furthermore, the government of Pakistan also does not really support the investors in a good way. High startup costs mean that very few companies even attempt to enter the sector. This lowers potential competition from the start. Proprietary technology forces even those with high startup capital to face an immediate operating disadvantage upon entering the sector. High fixed operating costs make companies with startup capital wary of entering the sector. Local and foreign governments force companies within the industry to closely comply with environmental regulations. These regulations often require capital to comply, forcing smaller companies out of the sector (Macfadyen and Watkins, 2014). Moreover, most of the countries cannot confidently allow others to dig out its resources very easily that becomes a barrier when it comes to the agreement for this kind of investment. This investment barrier, according to Dominique Simarad, slows down and reduces investment in the economy and weaken the economy's productive capacity and restrain long-term economic growth.

In developing countries terrorism has more dangerous effects than it has on developed nations, it affects the social and economic texture of the country Terrorist attacks, though not so often, can cause serious damages in the energy supply procedure, either due to actual destructions or because of the state of terror that a possible terrorist act creates (H. Doukas et al., 2011). In case of developed countries terrorism may affect the transaction of resources from the venerable sections of economy to the more secure sections. Involvement in war of Afghanistan, made Pakistan a victim of terrorism since last three decades, which resulted a tardiness in economic growth of the countries. Security is basic necessity for creating a bay for economic and business activities, therefore; unfavorable security condition in Baluchistan province and border areas of Pakistan, is known to be one of the major barriers to the oil and gas industry (Hyder, Akram and Padda, 2015). From the start, China has been concerned about the threats to CPEC in Pakistan in general and particularly in Balochistan. The southern province of Pakistan is home to a long-running insurgency by various ethnic Baloch militants seeking independence. (Notezai, 2018). China’s own geocentric and geo-economic interests in Pakistan are crystal clear. Therefore, China wants CPEC to be successful at all costs. On the other hand, the Baloch militant’s groups opposed the CPEC project are extremely concerned with China. Chinese analyst has recommended that the Chinese government take care to build local project support to ensure the success of the project (Notezai, 2018). The reasons of duplicity are sufficiently clear. Pakistan is uncomfortable with the prospect of becoming the focus if US-China economic confrontation that threatens to escalate into the fold of the recently launched counter belt and road fund between US and Japan and that the European Union has unveiled its plans to resist China’s expansion plans in one belt one road (Hussain 2018).

Since CPEC's launch in 2015, projects worth some US$ 18 billion have either been completed or are in advanced execution stages, helping to boost Pakistan's GDP growth to nearly 6%. But Chinese machinery imports ' corresponding burden has greatly contributed to a runaway deficit that has drained foreign exchange reserves to dangerously low levels, sending the rupee's exchange rate this year against the dollar plunging 19 percent (Hussain 2018). Moreover, the local community also prevents and creates some issues regarding the phenomenon by sometime charging high prices, living cost, local energy consultants suggest new tariff and non-tariff costs etc. This research paper focuses and discusses technological barrier, health and safety issues, liability issues, licensing, natural issues and environmental barriers, high costs in investment, lack of confidence, lack of governmental support, terrorist activities, fear of failure, lack of support from local community etc. that a company can face to invest in the oil and gas sector of Pakistan.

METHODOLOGY

This study follows qualitative paradigm of research. Overall design of the study consists of literature survey, data collection, analysis and interpretation of results. For literature survey relevant databases, websites, newspapers and other archival data has been explored thoroughly. The data has been collected from panel of experts using a matrix type questionnaire filled by the panel of experts as the respondents. Size of panel of experts is 11, all with the minimum experience of 10 years and the minimum education of PhD. Panel of experts include professors, economists and people from the related industry. The n(n-1)/2 matrix type questionnaire, suitable for ISM is used to elicit the data. The questionnaire is attached as annexure I. For the purposes of analysis ISM coupled with MICMAC is used in the study. Interpretive structural Modeling (ISM) is a technique for identifying relationships among factors or variables which define a specific problem or an issue. It is more suitable for studying the causal and complex relationships among multitude of factors. ISM is used in combination with Matriced' Impacts Croise's Multiplication Applique 'ea' un Classement (MICMAC) analysis.

ISM

ISM is proceeded stepwise which comprises of following steps.

Step 1: SSIM:

SSIM is prepared from questionnaires using majority rule/approval vote.

Table 1
Structural Self Interaction Matrix (SSIM)

Sr.	Barriers	1 Technology Barrier	2 Health and Safety Issue	3 Liability Issue	4 Licensing	5 Natural and Environmental Issue	6 High Costs/ Investments	7 Lack of Confidence	8 Lack of Governmental Support	9 Terrorist Activities	10 Fear of Failure	11 Lack of Support from Local Community
1	Technology Barrier		V	O	A	V	X	V	A	O	A	A
2	Health and Safety Issue			A	O	X	A	O	A	O	O	A
3	Liability Issue				O	O	X	V	A	A	X	O
4	Licensing					O	A	O	A	A	A	O
5	Natural and Environmental Issue						X	O	X	O	O	X
6	High Costs/ Investments							O	X	O	O	A
7	Lack of Confidence								A	A	X	A
8	Lack of Govern-mental Support									V	V	X
9	Terrorist Activities										V	A
10	Fear of Failure											A
11	Lack of Support from Local Community											

Table 2
Structural Self Interaction Matrix Completed for Preparing Initial Reachability

S#	Barriers	1 Technology Barrier	2 Health and Safety Issue	3 Liability Issue	4 Licensing	5 Natural and Environmental Issue	6 High Costs/ Investments	7 Lack of Confidence	8 Lack of Governmental Support	9 Terrorist Activities	10 Fear of Failure	11 Lack of Support from Local Community
1	Technology Barrier		V	O	A	V	X	V	A	O	A	A
2	Health and Safety Issue	A		A	O	X	A	O	A	O	O	A
3	Liability Issue	O	V		O	O	X	V	A	A	X	O
4	Licensing	V	O	O		O	A	O	A	A	A	O
5	Natural and Environmental Issue	A	X	O	O		X	O	X	O	O	X
6	High Costs/ Investments	X	V	X	V	X		O	X	O	O	A
7	Lack of Confidence	A	O	A	O	O	O		A	A	X	A
8	Lack of Governmental Support	V	V	V	V	X	X	V		V	V	X
9	Terrorist Activities	O	O	V	V	O	O	V	A		V	A
10	Fear of Failure	V	O	X	V	O	O	X	A	A		A
11	Lack of Support from Local Community	V	V	O	O	X	V	V	X	V	V	

Step 3: Initial Reachability Matrix

Reachability matrix prepared by using rule v=1, a=0, x=1, and o =0, for i*j relations whereas j*i relations were inferred vise-versa. Diagonals filled with inferred 1.

Factors	1	2	3	4	5	6	7	8	9	10	11
1	1	1	0	0	1	1	1	0	0	0	0
2	1	1	0	0	1	0	0	0	0	0	0
3	0	0	1	0	0	1	1	0	0	1	0
4	0	0	0	1	0	0	0	0	0	0	0
5	1	1	0	0	1	1	0	1	0	0	1
6	1	0	1	0	0	1	0	1	0	0	0
7	1	0	1	0	0	0	1	0	0	1	0
8	0	0	0	0	0	1	0	1	1	1	1
9	0	0	0	0	0	0	0	1	1	1	0
10	0	0	1	1	0	0	1	1	1	1	0
11	0	0	0	0	0	0	0	1	0	0	1

Step 4: Final Reachability Matrix:

From the initial reachability matrix, final reachability has been prepared by removing transitivity.

Table 4
Final Reachability Matrix

S#	Barriers	1	2	3	4	5	6	7	8	9	10	11	Deriving
1	Technology Barrier	1	1	0	0	1	1	1	0	0	0	0	5
2	Health and Safety Issue	1	1	0	0	1	0	0	0	0	0	0	3
3	Liability Issue	0	0	1	0	0	1	1	0	0	1	0	4
4	Licensing	0	0	0	**1**	0	0	0	0	0	0	0	1
5	Natural and Environmental Issue	1	1	0	0	1	**1**	0	1	0	0	1	6
6	High Costs/ Investments	1	0	1	0	0	1	0	1	0	0	0	4
7	Lack of Confidence	1	0	1	0	0	0	1	0	0	1	0	4
8	Lack of Governmental Support	0	0	0	0	0	1	0	1	1	1	1	5
9	Terrorist Activities	0	0	0	0	0	0	0	1	1	1	0	3
10	Fear of Failure	0	0	1	1	0	0	1	1	1	1	0	6
11	Lack of Support from Local Community	0	0	0	0	0	0	0	1	0	0	1	2
Dependence		5	3	4	2	3	5	4	6	3	5	3	43

From the final reachability matrix, driving and dependence power of each factor/barrier has been determined by counting no of 1's in row, and in column respectively. Driving and dependence power has subsequently been used for MICMAC analysis.

Table 5
Deriving and Dependence Power

Factors	1	2	3	4	5	6	7	8	9	10	11	Deriving
1	1	1	1*	0	1	1	1	1*	0	1*	1*	**9**
2	1	1	0	0	1	1*	1*	1*	0	0	1*	**7**
3	1*	0	1	1*	0	1	1	1*	1*	1	1*	**9**
4	0	0	0	1	0	0	0	0	0	0	0	**1**
5	1	1	1*	0	1	1	1*	1	1*	1*	1	**10**
6	1	1*	1	0	1*	1	1*	1	1*	1*	1*	**10**
7	1	1*	1	1*	1*	1*	1	1*	1*	1	0	**10**
8	1*	0	1*	1*	1*	1	1*	1	1	1	1	**10**
9	0	0	1*	1*	0	1*	1*	1	1	1	1*	**8**
10	1*	1*	1	1	0	1*	1	1	1	1	1*	**10**
11	0	0	0	0	0	1*	0	1	0	1*	1	**4**
Dependence	**8**	**6**	**8**	**6**	**6**	**10**	**9**	**10**	**7**	**9**	**9**	

Step 4: Iterations:

Iterations have been performed to determine the levels of factors in the model. For iterations elementary concepts of set theory has been used. For each factor reachability set (set consisting of factor itself and the factors to home it achieves), antecedent set (factor itself and by home it is achieved) and intersection (the set of common factors in reachability and antecedent) have been calculated as iteration 1. Where their reachability set, and intersection set are identical, that factor occupy level one. In this way the reachability and intersection set of all factors have been inspected and level one is determined. After determination of level one, the factors, the level of which is determined, have been eliminated from the iteration table resulting into iteration 2. Table of iteration 2 has again been inspected for finding reachability and intersection identical for determining level 2. This process has been repeated until the last level is found. The iterations reveal that there are five levels in the model. From the iterations, ISM model has been built using the elementary concepts of directed graph theory.

Table 6
Iteration I

S#	Barriers	Reachability Set	Antecedent Set	Intersection Set	Level
1	Technology Barrier	1,2,3,5,6,7, 8,10,11	1,2,3,5,6,7,8,10	1,2,3,5,6,7,8,10	
2	Health and Safety Issue	1,2,5,6,7,8,11	1,2,5,6,7,10	1,2,5,6,7	
3	Liability Issue	1,3,4,6,7,8,9, 10,11	1,3,5,6,7,8,9, 10,11	1,3,6,7,8,9,10	
4	Licensing	4	3,4,7,8,9,10	4	I
5	Natural and Environmental Issue	1,2,3,5,6,7,8, 9,10,11	1,2,5,6,7,8	1,2,5,6,7,8	
6	High Costs/ Investments	1,2,3,5,6,7,8, 9,10,11	1,2,3,5,6,7,8, 9,10,11	1,2,3,5,6,7,8, 9,10,11	I
7	Lack of Confidence	1,2,3,4,5,6,7, 8,9,10	1,2,3,5,6,7,8,9,10	1,2,3,5,6,7,8,9,10	
8	Lack of Governmental Support	1,3,4,5,6,7,8, 9,10,11	1,2,3,5,6,7,8,9,10 ,11	1,3,5,6,7,8,9, 10,11	
9	Terrorist Activities	3,4,6,7,8,9,10,11	3,5,6,7,8,9,10	3,6,7,8,9,10	
10	Fear of Failure	1,2,3,4,6,7,8, 9,10,11	1,3,5,6,7,8,9,10	1,3,6,7,8,9,10,11	
11	Lack of Support from Local Community	6,8,10,11	1,2,3,5,6,8,9, 10,11	6,8,10,11	I

Table 7
Iteration II

S#	Barriers	Reachability Set	Antecedent Set	Intersection Set	Level
1	Technology Barrier	1,2,3,5,7,8,10	1,2,3,5,7,8,10	1,2,3,5,7,8,10	II
2	Health and Safety Issue	1,2,5,7,8	1,2,5,7,10	1,2,5,7	
3	Liability Issue	1,3,7,8,9,10	1,3,5,7,8,9,10	1,3,7,8,9,10	II
5	Natural and Environmental Issue	1,2,3,5,7,8,9,10	1,2,5,7,8	1,2,5,7,8	
7	Lack of Confidence	1,2,3,5,7,8,9,10	1,2,3,5,7,8,9,10	1,2,3,5,7,8,9,10	II
8	Lack of Governmental Support	1,3,5,7,8,9,10	1,2,3,5,7,8,9,10	1,3,5,7,8,9,10	II
9	Terrorist Activities	3,7,8,9,10	3,5,7,8,9,10	3,7,8,9,10	II
10	Fear of Failure	1,2,3,7,8,9,10	1,3,5,7,8,9,10	1,3,7,8,9,10	

Table 8
Iteration III

S#	Barriers	Reachability Set	Antecedent Set	Intersection Set	Level
2	Health and Safety Issue	2,5	2,5,10	2,5	III
5	Natural and Environmental Issue	2,5,10	2,5	2,5	
10	Fear of Failure	2,10	5,10	10	

Table 9
Iteration IV

S#	Barriers	Reachability Set	Antecedent Set	Intersection Set	Level
5	Natural and Environmental Issue	5,10	5	5	
10	Fear of Failure	10	5,10	10	IV

Table 10
Iteration V

S#	Barriers	Reachability Set	Antecedent Set	Intersection Set	Level
5	Natural and Environmental Issue	5	5	5	V

Step 5: Building ISM Model

According to iterations, factor 4,6,11 occupy level one, factor 1,3,7,8,9 occupy level two, factor two occupies level three, factor 10 occupies level four and factor 5 occupies level five (fig. 1). The factor which occupies the bottom is the most important factor whereas the factors occupy the top level are the least important factors. Hence as a result, factor 5 i.e. Natural and Environmental Issues is the most important and independent factor. According to the model, most important factor that is more prioritized is natural and environmental issues and least prioritized factors Licensing, High cost and Investments and Lack of Support of local community. Other than that, are linkage factors.

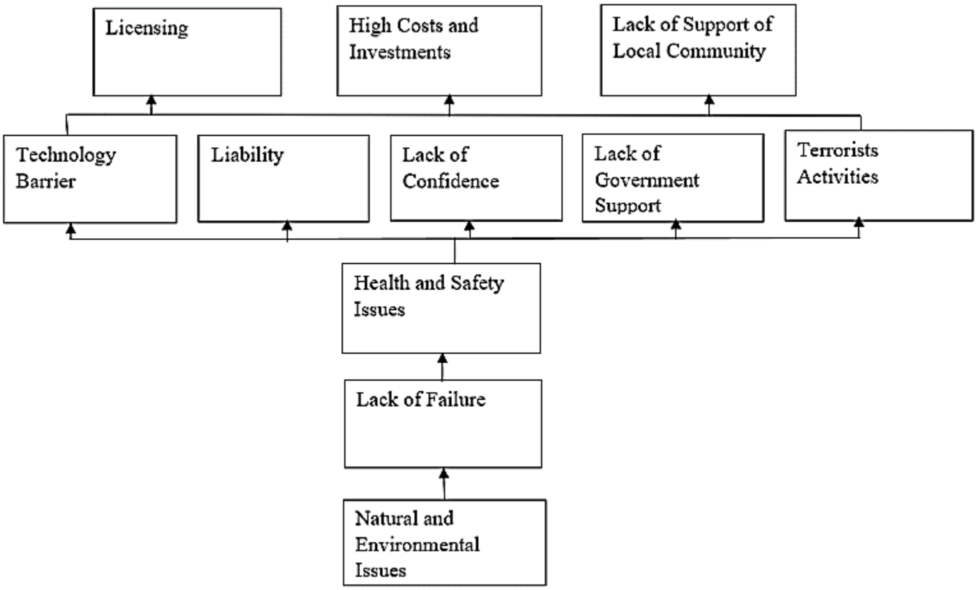


Figure 1: ISM Model

MICMAC Analysis:

Micmac is a cross impact matrix multiplication applied to classification using concepts of Boolean algebra. In this diagram factors are divided into four quadrants i.e. independent, autonomous, dependent and linkage. In fact, this diagram plots dependence power, factors on X axis and driving power on Y axis. The factor numbers are written against their driving-dependence power in respective quadrants. From inspection of MICMAC factor number 5,2 are independent, factor 4 is autonomous, factor 11 is dependent and factor 1,3,6,7,8,9,10 are linkage (Fig. 2).

Driving Power											
11											
10		Independent				5		Linkage 7,10		6,8	
9								1,3			
8							9				
7						2					
6											
5											
4									11		
3											
2		Autonomous						Dependent			
1						4					
	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**

Dependence Power

Figure 2: MICMAC Analysis

DISCUSSION

Bottom of the model is occupied by factor "Natural and Environmental issues", the logic of which maybe using the data set elicited by way of model exchange isomorphism (i.e. mental models of economists converted into binary, conical, ISM, and MICMAC). Since the experts comes of Pakistan and presumed to be familiar with the fact that possibilities of exploration of oil and gas are likely to be held in Balochistan. The province of Balochistan has peculiar natural and environmental issues like tropical climate, hills, scarcity of water etc. Therefore, the importance of natural and environmental issues predominantly appears to be overemphasized by experts. Had the data been collected from foreign investors, the results might have probably bit different.

CONCLUSION

This study identifies the relationship of non-tariff barriers for the investors of oil and gas sector of Pakistan. It is a seminal study since no such study has already been conducted. It explains non-tariff barriers in oil and gas sector of Pakistan, in context of investment in CPEC, ranks and prioritize them, develops an Interpretive model of their contextual relations and classifies them based on their driving-dependence power through ISM. Data is collected through matrix questionnaire, a survey responded by the panel of experts, relationship between the barriers (factors) is then identified and elaborated by ISM and MICMAC. Literature reveals eleven barriers (factors) that can be faced by the prosper investors, data collection and methodology identifies that the most important factor is the Natural and Environmental Issue that is identified after a statistical study on the Reponses of the respondents that are economist, professors and some of which belong to industry. Top factors that are least important contain Licensing, High Cost of Investment and Lack of Local Support. The study contributed list of factors that can be faced by the investors investing in the Oil and Gas sector of Pakistan, interpretive structural modeling technique and driving-dependence diagram of the factors. The study has implications for researcher, for policy makers and for society at large. Model provides deeper understanding of the phenomenon. It will be helpful to prospective investor, regulators and society at large to understand the complex inter-active relations among barriers. It will also help the stake holders to set policy preferences based on relative importance. The study has certain limitations like: 1) The list of non-tariff barriers have been generated from review of limited number of studies. The future researchers may prepare rather exhaustive list. 2) Data has been collected from economists of Pakistan; future researchers may collect the data from real foreign investors of this sector. 3) The study uses qualitative approach. Future researchers may use quantitative approaches.

REFERENCES

1. Acquah-Andoh, E., Ifelebuegu, A.O. and Theophilus, S.C. (2019). Brexit and UK Energy Security: Perspectives from Unconventional Gas Investment and the Effects of Shale Gas on UK Energy Prices. *Energies*, 12(4), 600.
2. Ali, A.A., Selvam, D.D.D.P., Paris, L. and Gunasekaran, A. (2019). Key factors influencing knowledge sharing practices and its relationship with organizational performance within the oil and gas industry. *Journal of Knowledge Management*, 23(9), 1806-1837.
3. Alkhaldi, M., Pathirage, C. and Kulatunga, U. (2017). The role of human error in accidents within oil and gas industry in Bahrain. In *13th International Postgraduate Research Conference (IPGRC): Conference Proceedings* (pp. 822-834). University of Salford.
4. Arifeen, M. (2018). Review of oil and gas consumption. *Pakistan and Gulf Economist*: http://www.pakistaneconomist.com/2018/08/06/review-of-oil-and-gas-consumption-last-year/
5. Aslam, R. (2011). Greed, creed, and governance in civil conflicts: a case study of Balochistan. *Contemporary South Asia*, 19(2), 189-203.
6. Bahremani, A. H. (2002). Oil and gas policies of the Islamic Republic of Iran. *Insight Turkey*, 4(1), 31-37.
7. Bashir, S., Arshad, M. and Barech, S. (2019). Local Communities Participation in China-Pakistan Economic Corridor (CPEC): The Case of Balochistan. *Pakistan Journal of Applied Social Sciences*, 9, 71-87.
8. Binnette, R. (1999). *International business* 2nd edition. Pearson education India.
9. Coelho, S.T. (2005). Biofuels–advantages and trade barriers. In *United Nations Conference on Trade and Development* (UNCTAD).
10. De Rijke, K. (2013). Hydraulically fractured: Unconventional gas and anthropology (Respond to this article at http://www.therai.org.uk/at/debate). *Anthropology Today*, *29*(2), 13-17.
11. Doukas, H., Flamos, A. and Psarras, J. (2011). Risks on the security of oil and gas supply. *Energy Sources, Part B: Economics, Planning, and Policy*, 6(4), 417-425.

12. Export Entreprises SA (2019). Country Risk of Pakistan: Investment. Retrieved from Import-export: https://import-export.societegenerale.fr/en/country/pakistan/country-risk-in-investment
13. Ford, J.A., Steen, J. and Verreynne, M.L. (2014). How environmental regulations affect innovation in the Australian oil and gas industry: going beyond the Porter Hypothesis. *Journal of Cleaner Production*, 84, 204-213.
14. Gazdar, H., Budhani, S.J., Mallah, H.B. and Khan, N.M.I. (2007). Balochistan Economic Report Background Paper on Social Structures and Migration. *TA4757-Pak: Balochistan Economic Report, Collective for Social Science Research 173-I Block, 2*.
15. Girko, P.D.S.I., Gubareva, P.D.A.V. and Kovalenko, P.D.K.E. (2017). Application of anti-dumping duties on imports of steel pipes for oil and gas industry: view from abroad. *Universidad y Sociedad*, 9(5), 309-312.
16. Gonzales, D., Gulden, T.R., Strong, A. and Hoyle, W. (2016). Major Incidents and Worker Deaths at California Refineries *(Chapter-5) in Cost-benefit Analysis of Proposed California Oil and Gas Refinery Regulations*. Rand Corporation.
17. Hussain, T. (2018). Why is the US trying to bully Pakistan with empty China threats? *This Week in Asia*. Online available at https://www.scmp.com/week-asia/politics/article/2158234/why-us-trying-bully-pakistan-empty-china-threats
18. Hyder, S., Akram, N. and Padda, I.U.H. (2015). Impact of terrorism on economic development in Pakistan. *Pakistan Business Review*, 839(1), 704-722.
19. Learner, E. (1990). The structure and effects of tariff and nontariff barriers in 1983. *The Political Economy of International Trade*, *New York: Basil Blackwell*, 224-260.
20. Liu, J. and Kemp, A. (2019). Forecasting the sign of US oil and gas industry stock index excess returns employing macroeconomic variables. *Energy Economics*, 81, 672-686.
21. Macfadyen, A. and Watkins, G. (2014). Alberta's Conventional Oil Resources. In *Petro Politics: Petroleum Development, Markets and Regulations, Alberta as an Illustrative History* (pp. 83-102). Canada: University of Calgary Press. Retrieved from http://www.jstor.org/stable/j.ctv5rf5q2.12
22. Morgera, E., Tsioumani, E. and Buck, M. (2015). Technology Transfer, Collaboration and Cooperation. In Morgera, E., Tsioumani, E. and Buck, M. (Eds.) *Unraveling the Nagoya Protocol: A Commentary on the Nagoya Protocol on Access and Benefit-sharing to the Convention on Biological Diversity* (pp. 314-321). Leiden; Boston: Brill.
23. Nag, T. (2019). Barriers to Cross-border Energy Cooperation and Implications on Energy Security: An Indian Perspective with Reference to Energy Trade in South Asia. *Global Business Review*, 0972150919826380.
24. Nkhata, B., Mosimane, A., Downs borough, L., Breen, C. and Roux, D. (2012). A Typology of Benefit Sharing Arrangements for the Governance of Social-Ecological Systems in Developing Countries. *Ecology and Society*, 17(1). Retrieved from http://www.jstor.org/stable/26269016
25. Notezai, M.A. (2018). Why Balochs Are Targeting China: CPEC, Balochistan, and unresolved political issues are a potent mix. *The Diplomate*. Online available at https://thediplomat.com/2018/11/why-balochs-are-targeting-china/

26. Overseas Pakistani friends (2019). *Historical chart of Pakistan rupee exchange rate vs US dollar*. Retrieved from https://www.opfblog.com/7786/historical-chart-of-pakistan-rupee-exchange-rate-vs-us-dollar/
27. *Pakistan - oil and gas equipment*. (2018). Retrieved April 9, 2019, from https://www.export.gov/article?id=Pakistan-Oil-and-Gas-Equipment
28. Patunru, A.A. and Rahardja, S. (2015). Trade protectionism in Indonesia: Bad times and bad policy. *Think-Asia*, https://think-asia.org/handle/11540/6410
29. Pirog, R. (2007). Russian Oil and Gas Challenges. *Connections*, 6(3), 82-99.
30. Sadorsky, P. (2001). Risk factors in stock returns of Canadian oil and gas companies. *Energy Economics*, 23(1), 17-28.
31. Samuel, U.T. (2019). The Myths and Realities of International Trade in the Age of Globalization: A Critical Assessment of Nigeria. *European Business & Management*, 5(5), 70.
32. Scholvin, S., Breul, M., Mello, P., Françoso, M. and Revilla Diez, J. (2017). Gateway cities in global production networks: exemplified by the oil and gas sector. *UNICAMP Texto para discussão*, 307.
33. Schuenemann, F. and Kerr, W.A. (2019). European union non-tariff barriers to imports of African biofuels. *Agrekon*, 1-19.
34. Sunley, E.M., Baunsgaard, T. and Simard, D. (2003). Revenue from the oil and gas sector: Issues and country experience. In Davis, J.M., Ossowski, R. and Fedelino A. (Eds), *Fiscal Policy Formulation and Implementation in Oil-Producing Countries*, International Monetary Fund, 153-183.
35. Vásquez, P. (2014). Tracing Oil- and Gas-Related Conflicts (Chapter-1). In *Oil Sparks in the Amazon: Local Conflicts, Indigenous Populations, and Natural Resources* (11-35). University of Georgia Press
36. Walter, I. (1971). Nontariff Barriers and the Export Performance of Developing Economies. *The American Economic Review*, 61(2), 195-205.
37. Warfield, J.N. (1973). On arranging elements of a hierarchy in graphic form. *IEEE Transactions on Systems, Man, and Cybernetics*, 2, 121-132.
38. Webb, M., Strutt, A. and Walmsley, T. (2019). Regulatory harmonization in the ASEAN region: The effects of applying different types of non-tariff measures. Presented at the *22nd Annual Conference on Global Economic Analysis*, Warsaw, Poland
39. Xiaobing, Z. and Jianglin, Z. (2012). Economic and Technical Cooperation: Creating the Environment to Remove the Barriers. In Drysdale P., Yunling Z. and Song L. (Eds.), *APEC and liberalization of the Chinese economy* (pp. 207-222). ANU Press.

QUESTIONNAIRE

Applying ISM on Non-Tariff Barriers for the Investors of Oil and Gas Sector, in Pakistan: CPEC Perspective

Notes:

- We are conducting research regarding the issues faced by the oil and gas investors, after CPEC, in Pakistan.
- Your input will be a great contribution in our research work and this questionnaire will be used for research purpose only in combined statistical statements.

Section-1

PERSONAL INFORMATION

Name (Optional): ______________________________

Designation: ______________________________

Organization: ______________________________

Date: ______________________________

DEMOGRAPHICS

Gender

☐ Male ☐ Female

Age Group

☐ 21-30 ☐ 31-40
☐ 41-50 ☐ Above 50

Income (in thousands)

☐ <40 ☐ 40-80
☐ 81-100 ☐ 101-200
☐ 201-300 ☐ Above 300

Marital Status

☐ Married ☐ Single

Qualification

☐ Less than 14 years ☐ 14 years
☐ 16 years education ☐ Above 16

Experience

☐ up to 5 year ☐ 5-10 Years
☐ 10-15 Years ☐ Above 20

Section 2:

RESEARCH QUESTIONNAIRE

1. Contextual Relationship = leads to
2. What to enter in the white cells:
   - Enter **V** when the row influences the column
   - **A** when the column influences the row
   - Enter **O** when there is no relation between the row and the column
   - Enter **X** when row and column influence each other

S#	Barriers	1 Technology Barrier	2 Health and Safety Issue	3 Liability Issue	4 Licensing	5 Natural and Environmental Issue	6 High Costs/ Investments	7 Lack of Confidence	8 Lack of Governmental Support	9 Terrorist Activities	10 Fear of Failure	11 Lack of Support from Local Community
1	Technology Barrier											
2	Health and Safety Issue											
3	Liability Issue											
4	Licensing											
5	Natural and Environmental Issue											
6	High Costs/ Investments											
7	Lack of Confidence											
8	Lack of Governmental Support											
9	Terrorist Activities											
10	Fear of Failure											
11	Lack of Support from Local Community											

Proc. 15th Islamic Countries Conference on Statistical Sciences
Lahore, Pakistan – December 21-24, 2019, Vol. 34, pp. 531-536

THE EFFECTS OF ACADEMIC DISMISSAL POLICY AND STUDENTS GRADUATION RATE

Narmeen Faisal and **Ammara Nawaz Cheema**[§]
Department of Mathematics, Air University, Islamabad, Pakistan
Email: [§]ammara.au@gmail.com

ABSTRACT

Academic dismissal polices are implemented to support and encourage academic success. These probations are used to signal students for improvement. With existing empirical evidence mostly restricted to short-run outcomes. This study examines long-term academic outcomes of academic dismissal for bachelor students in engineering at air University. Using administrative records, descriptive estimates suggest that academic dismissal does not relate to a difference in the propensity of graduation, nor to a change in study delay, when comparing students around the academic dismissal threshold. Not meeting this credit threshold forces students to leave, and most decide to re-enroll in the same program.

1. INTRODUCTION

Students who fail to meet their academic goals may subject to academic dismissal. Consistent satisfactory academic performance is needed to meet the goals. Fletcher and Tokmouline (2017) used discontinuity regression approach to evaluate effect of academic probation in college success. They also find pre-determined student characteristics based on different responses. Lindo et al., (2010) suggested that simple model is suitable for their research work. Moreover, Martorell and McFarlin (2011) studied that academic dismissal polices provides no positive results on students from Texas University. Carlos Calcagno and Long (2008) concluded that academic probation increases short term consistency and decreases rate of graduation among graduates.

Substantial lists of university websites stating academic dismissal policies can be find out by a simple google search. Fletcher and Tokmouline (2017) compared that average college learners results of students on academic dismissal, with those graduates who are not on probation (see, Scalice et al. (2000)). For use of regression discontinuity design, comparison of undergraduates provides several factors that lead to reasonable devoid of truth discussed by Lee and Lemieux (2010). Fletcher (2017) explored that consequences of academic dismissal are only concluded by Canadian university and finds that his study covers broaden aspects of academic probation.

Kelley (1996) examined distinct phases of academic probation which includes cognitive, emotional, behavioral and environmental factors. These factors inhibit student performance, their reaction for being placed on probation and how they cope up with probation. All factors are associated with controllable and un-controllable behaviors. Heider (1958) stated that the perception of causality is a union of personal and environmental forces.

(Murray et al. 1993) proposed that an individual's learning-style has influence on classroom participation. Brinckerhoff (1989) states that many students encounter academic difficulties because they have an unassessed learning disability. Bandura and National Inst. of Mental Health (1986) claims that students with high self-efficiency have favorable outcomes and perform multiple tasks needed for success. (Multon et al. 1991) propose that students with low self-efficiency lack persistence, are not fully satisfied and have unsatisfactory performance. De Charms (1968) suggested that determined attempts are made to get favorable outcome. Students take measurable steps and get command over different circumstances.

In academic career, students decide how to manage their time. Researchers also suggested that when values are clarified, then students are motivated (see, Katz and Hass, 1988; Schwartz and Howard, 1982). Students are highly motivated to interpret the reasons of their efficient academic performance (Jones and Davis, 1965; Kelley, 1971). Bradley (1978) explored that students want to maintain a good public image.

2. METHODOLOGY

The main focus of this proposed study is in the statistical analysis of the academic dismissal cases. To illustrate the proposed methodology, we take the data on academic dismissal of students from AU. We apply basic statistical analysis.

Descriptive statistics are used to describe the basic features of the data in a study. They provide simple summaries about the sample and the measures. Together with simple graphics analysis, they form the basis of virtually every quantitative analysis of data.

Descriptive statistics are typically distinguished from inferential statistics. With descriptive statistics you are simply describing what is or what the data shows. With inferential statistics, you are trying to reach conclusions that extend beyond the immediate data alone. For instance, we use inferential statistics to try to infer from the sample data what the population might think. Or, we use inferential statistics to make judgments of the probability that an observed difference between groups is a dependable one or one that might have happened by chance in this study. Thus, we use inferential statistics to make inferences from our data to more general conditions; we use descriptive statistics simply to describe what’s going on in our data.

Descriptive Statistics are used to present quantitative descriptions in a manageable form. In a research study we may have lots of measures. Or we may measure a large number of people on any measure. Descriptive statistics help us to simplify large amounts of data in a sensible way. Each descriptive statistic reduces lots of data into a simpler summary.

The central tendency of a distribution is an estimate of the “center” of a distribution of values. There are three major types of estimates of central tendency:

- Mean
- Median
- Mode

The **Mean** or average is probably the most commonly used method of describing central tendency. To compute the mean all you do is add up all the values and divide by the number of values

The **Median** is the score found at the exact middle of the set of values. One way to compute the median is to list all scores in numerical order, and then locate the score in the center of the sample.

The **Mode** is the most frequently occurring value in the set of scores. To determine the mode, you might again order the scores as shown above, and then count each one. The most frequently occurring value is the mode.

Dispersion refers to the spread of the values around the central tendency. There are two common measures of dispersion, the range and the standard deviation. The **range** is simply the highest value minus the lowest value.

The **Standard Deviation** is a more accurate and detailed estimate of dispersion because an outlier can greatly exaggerate the range

Graphing is a pictorial way of representing relationships between various quantities, parameters, or measurable variables in nature. A **graph** basically summarizes how one quantity changes if another quantity that is related to it also changes.

3. RESULTS AND DISCUSSION

The relative performance of graduate rate we observe that, 62% successfully completed their degree, 17% have refunded, 15% is the dismissal rate, and 5% are migrated to other discipline of study. Whereas, poverty level indicates 75% of students can easily afford their education expenses and 25% belongs to low income class.

Table 1
Gender, High School Rank and Credit Hours

Descriptive	Gender	High School Rank/ High School Percentage	Credit Hours Earned in Fall 13	F.S.C. Score	AU Score
Mean	.92	2.05	18.18	74.30	64.98
Std. Error of Mean	.012	.031	.047	.285	.195
Median	1.00	2.00	19.00	74.55	63.76
Mode	1	2	19	83	61[a]
Std. Deviation	.274	.685	.984	6.393	4.301
Variance	.075	.470	.969	40.872	18.497
Skewness	-3.064	.160	-.371	-.247	1.030
Std. Error of Skewness	.109	.109	.116	.109	.111
Kurtosis	7.418	-.272	-1.870	-.397	.822
Std. Error of Kurtosis	.218	.217	.231	.217	.221

Majority of students have their study background from federal, Rwp and Lahore board. 56% students belongs to urban domiciles whereas 44% have rural domicile Most of the enrolled student have more than 70% in their previous degree score The ratio of male candidates are 92% in engineering discipline as compare to female students. 58% enrollment is in Electrical engineering, 17% in Mechanical engineering and 25% is in MTS. As shown in the following table

Table 2
Domicile, Feeder School, Poverty Level and Graduation

Descriptive	Domicile	Feeder School	Geographic Identities	Poverty Level	Graduation
Mean	1.17	1.32	.45	.75	1.69
Std. Error of Mean	.029	.021	.022	.019	.048
Median	1.00	1.00	.00	1.00	1.00
Mode	1	1	0	1	1[a]
Std. Deviation	.646	.466	.497	.435	1.076
Variance	.417	.217	.248	.189	1.157
Skewness	4.949	.784	.221	-1.143	1.848
Kurtosis	26.571	-1.392	-1.959	-.697	4.031

Table 3
Probation and Dismissal

Descriptive	On Probation after 1st Semester	On Probation after 2nd Semester	Ever on Probation	Dismissal
Mean	.29	.44	.76	.47
Std. Error of Mean	.044	.052	.042	.052
Median	.00	.00	1.00	.00
Mode	0	0	1	0
Std. Deviation	.454	.499	.429	.502
Variance	.206	.249	.184	.252
Skewness	.962	.242	-1.233	.130
Kurtosis	-1.095	-1.985	-.490	-2.027

From the following figured it is observed that dismissal rate is 9% and graduation rate is 91%. Moreover, after 1st semester 6% and 9% are on probation after their complication of 2nd semester and 16% candidates are ever on probation.

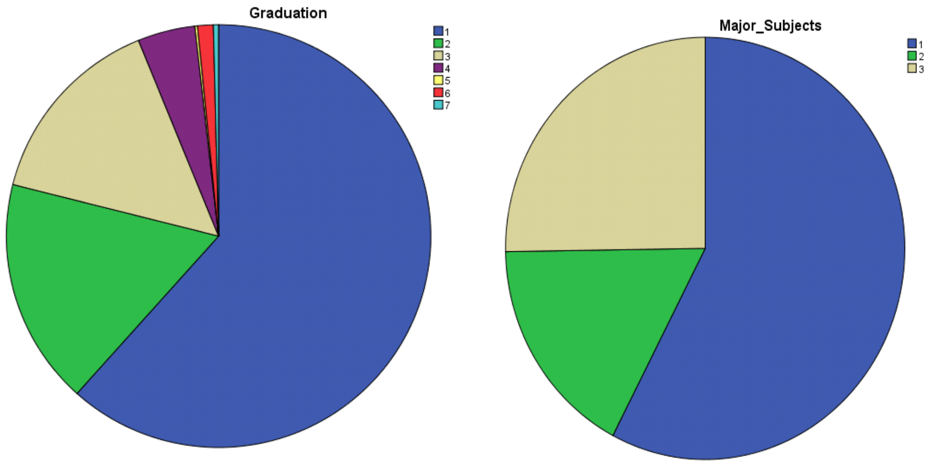


Figure 1-2: Graduation Rate and Major Subjects

4. COMMENTS AND CONCLUSION

- Statistical relation towards AD have been formulated
- Basic statistical analysis is carried out to check the effectiveness of said dataset
- To judge the relative performance of graduate rate, 62% successfully completed their degree, 17% have refunded, 15% is the dismissal rate, and 5% are migrated to other discipline of study
- Poverty level indicates 75% of students can easily afford their education expenses where as 25% belongs to low income class
- Majority of students have their study background from federal, Rwp and Lahore board.
- 56% students belongs to urban domiciles whereas 44% have rural domicile
- Most of the enrolled student have more than 70% in their previous degree score
- The ratio of male candidates are 92% in engineering discipline as compare to female students
- 58% enrollment is in Electrical engineering, 17% in Mechanical engineering and 25% is in MTS
- Dismissal rate is 9% and graduation rate is 91%
- After 1st semester 6% and 9% are on probation after their complication of 2nd semester and 16% candidates are ever on probation.

REFERENCES

1. Calcagno, J.C. and Long, B.T. (2008). *The impact of postsecondary remediation using a regression discontinuity approach: Addressing endogenous sorting and noncompliance* (No. w14194). National Bureau of Economic Research.
2. Fletcher, J.M. and Tokmouline, M. (2017). *The effects of academic probation on college success: Regression discontinuity evidence from four Texas universities*. IZA Discussion Paper No. 11232. Available at SSRN: https://ssrn.com/abstract=3097330.

3. Lee, D.S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2), 281-355.
4. Lindo, J.M., Sanders, N.J. and Oreopoulos, P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal: Applied Economics*, 2(2), 95-117.
5. Martorell, P. and McFarlin Jr, I. (2011). Help or hindrance? The effects of college remediation on academic and labor market outcomes. *The Review of Economics and Statistics*, 93(2), 436-454.
6. Scalise, A., Besterfield-Sacre, M., Shuman, L. and Wolfe, H. (2000). First term probation: Models for identifying high risk students. In *30th Annual Frontiers in Education Conference. Building on A Century of Progress in Engineering Education*. Conference Proceedings (IEEE Cat. No. 00CH37135) (Vol. 1, pp. F1F-11). IEEE.
7. Bandura, A. and National Inst. of Mental Health (1986). *Prentice-Hall series in social learning theory. Social foundations of thought and action: A social cognitive theory*. Prentice-Hall, Inc.
8. Bradley, G.W. (1978). Self-serving biases in the attribution process: A reexamination of the fact or fiction question. *Journal of Personality and Social Psychology*, 36(1), 56-71.
9. Brinckerhoff, L. (1989). Diagnostic considerations for students with learning disabilities at selective colleges. In *Proceedings of the next step: An invitational symposium on learning disabilities in selective colleges* (pp. 114-136).
10. Heider, F. (1958). *The psychology of interpersonal relations*. New York: John Wiley.
11. Katz, I. and Hass, R.G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, 55, 893-905.
12. Kelley, K.N. (1996). Causes, reactions, and consequences of academic probation: A theoretical model. *NACADA Journal*, 16(1), 28-34.
13. Multon, K.D., Brown, S.D. and Lent, R.W. (1991). Relation of self-efficacy to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38, 30-38.
14. Schwartz, S.H. and Howard, J. (1982). A Normative decision making and altruism. In V.J. Derlega, & J. Grzelak (Eds), *Cooperation and helping behavior* (pp 189-211) Hillsdale, NJ: Erlbaum.