

# Modelling Job Satisfaction and Age Using a Composite Function of Gamma and Power Transformation Function

Soyinka Taiwo<sup>1,\*</sup>, Olosunde Akin<sup>2</sup>, Oyetola Shola<sup>1</sup>, Akinhanmi Akin<sup>1</sup>

<sup>1</sup>Department of Research and Training, Neuropsychiatric Hospital Aro, Abeokuta, Nigeria

<sup>2</sup>Department of Mathematics, Obafemi Awolowo University, Ile-Ife, Nigeria

**Abstract** In this paper, we derive a joint distribution model to enhance a bi-variate relationship among variables with quantitative responses. Obviously, most real life responses are non-normal in nature, so the use of Box-Cox transformation was employed to improve normality. We assumed a gamma distribution for the non normal data and estimated the degree of shrinkage or spread required for normality as the numerical distance of the location parameter. A proper Jacobean density transformation was used to ease interpretation and avoid distortion in the original data unit of measurement. All unknown parameters, except location parameter which is fixed were determined by the method of moment estimation and maximum likelihood estimation using codes in R. The interactive dependence of job satisfaction and age was established in the study with substantial mathematical and statistical evidence. A scatter plot of job satisfaction against age, shows a non-linear relationship which resembles a U-shape; parabolic in nature. The scatter plot shows a significance difference between younger staffs (20 – 33 years of age) and the older staffs (with at least 50 years of age) in terms of their level of job satisfaction and their job satisfaction growth rate. The limitation of the study is the use of transformed data.

**Keywords** Job satisfaction, Age, Gamma distribution, Box-Cox transformation, Kolmogorov-smirnov test (KS)

## 1. Introduction

Employee attitude to work and their general job performance is dependent on the psychosocial characteristics of their working environment. These psychosocial characteristics which include status, promotion, decision making, and personal growth are variables that significantly affect the sense of empowerment felt by workers and determine the level of employees' commitment and satisfaction to the work [1].

Previous studies by [2-4] have shown that job satisfaction is the perception and evaluation of an individual's job over years and is dependent on demographic factors like age and gender.

Hence the aim of this study was however to model the relationship between age and job satisfaction using gamma distribution to ascertain the claim of non linearity relationship by past researchers with substantial mathematical and statistical evidence.

In statistical techniques, data transformation has been found very helpful as it improves the quality of model approximation and also makes inferences from such model

to be robust [5, 6]. However, despite the importance of transformation, some imperfection in the parameter values of the approximate model compared to the raw data parameters, often times has always put a limit to the degree of asymmetry achieved [7]. This imperfection is such that if the original data were negatively skewed, then the transformation index shifts the transformed data mean position to the right to improve the degree of asymmetry and if the original mean position is skewed to the right, the transformed mean position is shifted inward to achieve approximate normality [5, 7]. Hence the choice of gamma (3P) distribution as assumed model is based firstly on the non normal nature of the variables and that normally distributed variables can be achieved by manipulating its parameter values. Secondly on the assumption that the imperfection noticed as a result of shift in the mean (location) position due to transformation can be estimated from the location parameter and finally that the transformation has no effect on the variance of the distribution since the Box-Cox transformation engaged is a variance stabilizing transformation. A proper Jacobean transformation between the power transformation function and the assumed model was ensured to correct the distortion in the original measurement scale and ease interpretation of result. See [5, 6, 8] on Box-Cox transformation and [9] for the properties of gamma distribution.

\* Corresponding author:

soyinkataiwo@yahoo.ie (Soyinka Taiwo)

Published online at <http://journal.sapub.org/statistics>

Copyright © 2017 Scientific & Academic Publishing. All Rights Reserved

## 2. Literature Review

The power transformation of data to achieve approximate normality and enhance statistically based inference has been shown to be of great importance in the last five decades [6, 8]. Of all these power transformations, Box-Cox transformation

defined by  $x_j^\lambda = \frac{x_j^\lambda - 1}{\lambda}$ ,  $\lambda \neq 0$ , or  $x_j^\lambda = \ln x$ ,  $\lambda = 0$

is the most popular. However, most statistician prefers to obtain the value of  $\lambda$  at which the expression

$$B(x_j^\lambda) = \frac{x_j^\lambda - 1}{\lambda \left[ \left( \prod_{i=1}^n x_i \right)^{\frac{1}{n}} \right]^{\lambda-1}}$$
 has the minimum variance.

For each  $j = 1, 2, \dots, n$ , the variance value of the function  $B(x_j^\lambda)$  is obtained for each  $\lambda$ . The minimum of all the variances and its associated  $\lambda$  is desired [5-8].

Note that in gamma (3P);  $f(x) = \frac{(x-c)^{a-1}}{\Gamma(a)b^a} e^{-\left(\frac{x-c}{b}\right)}$ ,

$x > c$ ,  $a > 0$ ,  $b > 0$

1. The location parameter  $c$  could be positive ( $x > c$ ) or negative ( $x > -c$ ) depending on the shift direction leading to a transformation variable  $(x-c)$  or  $(x+c)$  respectively.

2. The value of parameter  $a$  and  $b$  will determine if the variable can be fitted with gamma 3P, gamma 2P or the standard normal density.

In situation when the original data  $x$  is approximately normally distributed, there is no need for transformation and so  $c = 0$ . The distribution thus reduces to a gamma (2P)

$$f(x) = \frac{x^{a-1}}{\Gamma(a)b^a} e^{-\left(\frac{x}{b}\right)}, \quad x > 0, a > 0, b > 0.$$
 Then if the

shape parameter is large ( $a \geq \text{mean}$ ) and the scale parameter  $b = 1$ , the variable is approximately distributed as standard normal random variable  $\frac{x-a}{\sqrt{a}}$  [9].

## 3. Material and Data Collection

Two self administered questionnaires was used to capture the required information for the study. A self developed questionnaire was used to obtain the demographic variables

such as age, sex, marital status and so on. The job satisfaction of respondent was measured using the Minnesota Satisfaction Questionnaire (MSQ). MSQ was developed by the Vocational Psychological Research unit of the University of Minnesota in 1967 and it has a reliability coefficient of 0.90. A total of 393 teachers from different secondary schools within Abeokuta the Ogun state capital in Nigeria were used in the study. Necessary ethical approvals were obtained before the commencement of the study.

## 4. Methodology

- Deriving the composite gamma-power function for bi-variate relationship using Jacobean transformation.
- Transformation of both job satisfaction and age to achieve normality via Box-Cox techniques.
- Estimating the parameters of the gamma distribution using moment estimation method.
- Hypothesis testing of the fitted models using kolmogorov-smirnov test (KS).
- Plotting Job satisfaction against Age using codes written with R.

## 5. Result

### Deriving the composite gamma-power function and its joint probability density function

The proper Jacobean transformation for a transformed

variable can be derived as  $f(y) = \frac{y^{\frac{a}{\lambda}-1}}{\lambda \Gamma(a) b^a} e^{-\left(\frac{y^{\frac{1}{\lambda}}}{b}\right)}$  where

$y = (x \pm c)^\lambda$ .  $f(y)$  can be shown to be a probability density function. The parameters  $a$  and  $b$  can be obtained from the moment estimates of the model.  $\lambda$  is the transformation index from Box-Cox.

Suppose we have two successfully transformed variables  $y_1$  and  $y_2$  then the joint density model of the two random variables is given as

$$f(y_1, y_2) = \frac{y_1^{\left(\frac{a_1}{\lambda_1}-1\right)} y_2^{\left(\frac{a_2}{\lambda_2}-1\right)}}{(\lambda_1 \lambda_2) (\Gamma(a_1) \Gamma(a_2)) (b_1^{a_1} b_2^{a_2})} e^{-\left[\frac{y_1^{\frac{1}{\lambda_1}}}{b_1} + \frac{y_2^{\frac{1}{\lambda_2}}}{b_2}\right]}.$$

Lets define a ratio bi-variate relationship between  $y_1$  and

$y_2$  then if  $k = \frac{y_1}{y_2}$  and  $u = y_2$  then

$$J = \left| \frac{\partial(y_1, y_2)}{\partial(k, u)} \right| = u \quad \text{and} \quad f(k) = \frac{k^{\frac{a_1}{\lambda_1}-1}}{(\lambda_1 \lambda_2) (\Gamma(a_1) \Gamma(a_2)) (b_1^{a_1} b_2^{a_2})} \left[ \int_0^\infty u^{\frac{a_1+a_2}{\lambda_1+\lambda_2}-1} e^{-\left[\frac{k^{\frac{1}{\lambda_1}} u^{\frac{1}{\lambda_1}}}{b_1} + \frac{u^{\frac{1}{\lambda_2}}}{b_2}\right]} du \right]$$

Note that the function can be simplified further via parameter manipulations, or else the integral part will be solved by numerical integration. However given that  $y_1$

and  $y_2$  are job satisfaction scores and age respectively then  $k$  is measure of job satisfaction growth rate with age.

**Table 1.** Summary of the transformation index value and the distribution parameter values

Variables	Original data mean	Box-Cox index ( $\lambda$ )	Transformed data mean (P)	Transformed data Variance (Q)	Gamma distribution parameters estimate
Job satisfaction	62.67	0.9	62.43	176.33	a=22.1; b=2.82; c= 0.24
Age	39.5	0.7	38.98	17.96	a=84.6; b=0.461; c= 0.52

R codes to calculate the maximum likelihood value of  $\lambda$  is given in APPENDIX A.

Note: For age, the shape parameter  $a$  is greater than the mean, so according to [9] the distribution reduces to gamma (2P) since  $c=0$ .

**Table 2a.** Kolmogorov-Smirnov test for goodness of fit on Job Satisfaction groupings

Job satisfaction groupings	Frequency Observed	Relative Frequency Observed	Gamma (3P) Expected	Maximum Deviation $D_{\max}$
30 – 39	5	0.088	0.0228	0.0652
40 – 49	10	0.175	0.119	0.1212
50 – 59	10	0.175	0.287	0.0092
60 – 69	10	0.175	0.297	0.1128
70 – 79	10	0.175	0.186	0.1238**
80 – 89	10	0.175	0.08	0.029
90 - 99	2	0.037	0.008	0
Total	57	1	1	

$$D_{tab,\alpha=0.05} = \frac{1.36}{\sqrt{N}} = \frac{1.36}{\sqrt{57}} = 0.18. \text{ Since } D_{\max} = 0.1238 < D_{tab,\alpha=0.05}; \text{ the model is well fitted.}$$

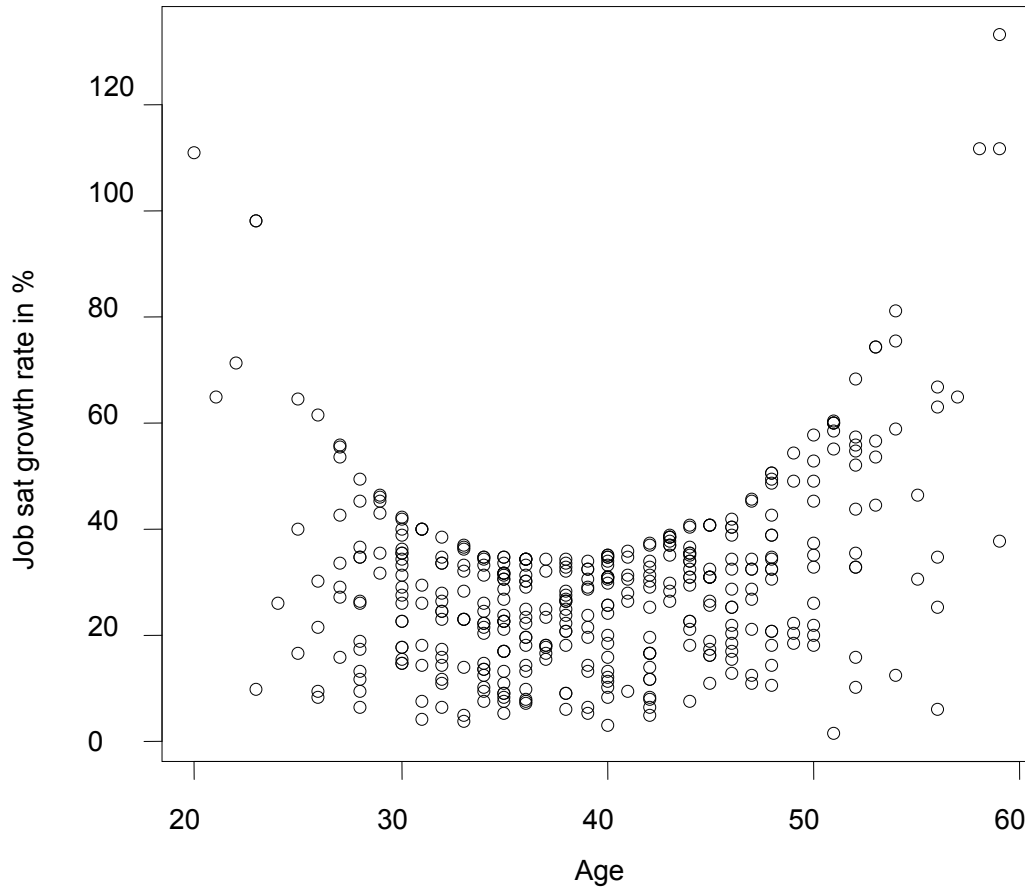
**Table 2b.** Kolmogorov-Smirnov test for goodness of fit on Age groupings

Age Groups	Frequency	Relative frequency Observed	Gamma (2P) Expected	Cummulative Observed	Cummulative Expected	Maximum Deviation $D_{\max}$
20 - 29	10	0.25	0.1944	0.25	0.1944	0.0556
30 – 39	10	0.25	0.3674	0.5	0.5618	0.0618
40 – 49	10	0.25	0.297	0.75	0.8588	0.1088**
50 – 59	10	0.25	0.1412	1	1	0
Total	40	1	1			

$$D_{tab,\alpha=0.05} = \frac{1.36}{\sqrt{N}} = \frac{1.36}{\sqrt{40}} = 0.215. \text{ Since } D_{\max}=0.1088 < D_{tab,\alpha=0.05}; \text{ the model is well fitted.}$$

See Appendix B for programme on estimated probability values.

## SCATTER PLOT



**Figure 1.** The graphical relationship between Job satisfaction growth rate and Ages of teachers in secondary schools across Abeokuta. (See Appendix C for the R codes used for the scatter plot.)

## 6. Discussion of Result

The condition that the shape parameter  $a$  is greater than the mean and the scale parameter  $b$  is within the range  $(0,1]$ ; determines our assumption of assumed normal or non normal distribution for any of the variables and the need for transformation technique to improve normality. In case of non normality, the numerical value of the location parameter  $C$  was used as the measure of shrinkage or spread required to shift the data towards the mean. Table 1 gives the summary of the transformation index and the parameter values of the distribution fitted for each variable. The evidence of model fitness for non repeated responses was ascertained in Table 2 using kolmogorov smirnov test (KS).

The composite gamma-power probability density function for job satisfaction growth rate with age ( $k$ ) was obtained as

$$f(k) = \frac{1}{\lambda \text{Beta}(a, a)} k^{\frac{a}{\lambda}-1} \left( k^{\frac{1}{\lambda}} + 1 \right)^{-2a}, \quad a > 0, \quad \lambda > 0.$$

Codes were written in R to evaluate  $f(k)$  (Appendix C).

A scatter plot of the job satisfaction growth rate to the ages of respondents gave a U shape (figure 1). Freshly employed staffs, within the age of 20-25 years has a high level of job

satisfaction within 80-120 percent; possibly because they just got a new job upon which they can develop their career in the mist of high unemployment rate. A reason why careful selection of career is of great important. As the respondents age moves from 25-30 years, the level of job satisfaction drops to 40% and continues to drop till it becomes stable at 10% between the age brackets of 33-44 years. A job satisfaction of 10% is very low. However beyond 45 years of age, the respondents' job satisfaction begins to pick up at a very slow rate. The continues decrease in job satisfaction from ages 20-30 years till it becomes stable between age brackets 33-44, may be due to the fact that as people grow older with this career, personality and situational factors set into play and these may determine the desire to go on with this career or not. Personality factors include risk taking tendency and sense of control over one's destiny. While, situational factor include increasing disenchantment with one's present career, discovery of other occupation that promises a greater satisfaction and pivotal events (divorce, death of a loved one) that lead one to shift life goals and priorities. These factors may determine the satisfaction one gets from this career and the desire to stay on with the career. The level of job satisfaction growth rate increases

noticeable as respondents attained the age of 50 years and beyond. The respondents within the age bracket of 50-60 years of age, has a high level of job satisfaction between 60-120 percent with sharp job satisfaction growth rate. This may result from the fact that as one gets older above the average age; the rise in job satisfaction could come from reduced aspirations due to the recognition that there are few alternative jobs available once their careers are established [11].

## 7. Conclusions

The assumption of normality and the need for power transformation was checked using the parameter values of gamma distribution. The composite function for job satisfaction growth rate was plotted against the age in a scatter plot (Figure 1). We observed a non-linear relationship between age and job satisfaction. The non-linearity shows a U-shaped relationship, with those in the very young and old age groups being the most satisfied. This illustrates that as age increases from 33 years to 40 years, the job satisfaction levels decreases. This reduction in the job satisfaction level continues until around the age of 40-44 years, after which any further increase in age beyond 44 years shows a corresponding increase in the job satisfaction level.

## Appendix

### APPENDIX A:- Codes to obtain Box-Cox transformation index $\lambda$

```
> z <- c(31, 34, 35, 36, 37, 38, 39, 40, 42, 43, 45, 46, 47, 48,
49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64,
65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80,
81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 92)
> K <- seq (-1, 1.5, by=0.1)
> A <- (prod (z))^(1/(length (z)))
> a <- (((z^1)-1)/(-1*(A^-2)))
> b <- (((z^-.9)-1)/(-.9*(A^-1.9)))
> c <- (((z^-.8)-1)/(-.8*(A^-1.8)))
> d <- (((z^-.7)-1)/(-.7*(A^-1.7)))
> e <- (((z^-.6)-1)/(-.6*(A^-1.6)))
> f <- (((z^-.5)-1)/(-.5*(A^-1.5)))
> g <- (((z^-.4)-1)/(-.4*(A^-1.4)))
> h <- (((z^-.3)-1)/(-.3*(A^-1.3)))
> i <- (((z^-.2)-1)/(-.2*(A^-1.2)))
> j <- (((z^-.1)-1)/(-.1*(A^-1.1)))
> k <- (A*log (z))
> l <- (((z^1)-1)/(1*(A^-0.9)))
> m <- (((z^2)-1)/(2*(A^-0.8)))
> n <- (((z^3)-1)/(3*(A^-0.7)))
> o <- (((z^4)-1)/(4*(A^-0.6)))
> p <- (((z^5)-1)/(5*(A^-0.5)))
> q <- (((z^6)-1)/(6*(A^-0.4)))
> r <- (((z^7)-1)/(7*(A^-0.3)))
```

```
> s <- (((z^8)-1)/(8*(A^-0.2)))
> t <- (((z^9)-1)/(9*(A^-0.1)))
> u <- (((z^10)-1)/(10*(A^-0)))
> v <- (((z^11)-1)/(11*(A^0.1)))
> w <- (((z^12)-1)/(12*(A^0.2)))
> x <- (((z^13)-1)/(13*(A^0.3)))
> y <- (((z^14)-1)/(14*(A^0.4)))
> z <- (((z^15)-1)/(15*(A^0.5)))
> K1 <- c(var (a), var (b), var (c), var (d), var (e), var (f),
var (g), var (h), var (i), var (j), var (k), var (l), var (m), var (n),
var (o), var (p), var (q), var (r), var (s), var (t), var (u), var (v),
var (w), var (x), var (y), var (z)).
```

### APPENDIX B:- Code to estimate the probability values of job satisfaction and age

```
> Job <- c(31, 34, 35, 36, 37, 38, 39, 40, 42, 43, 45, 46, 47,
48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63,
64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79,
80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 92)
> Age <- c(20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,
33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,
49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59)
> x <- c(31, 34, 35, 36, 37, 38, 39, 40, 42, 43, 45, 46, 47,
48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63,
64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79,
80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 92)
> b <- 2.824
> a <- 22.10
> c <- 0.24
> f <- ((x-c)^(a-1))*(exp (-((x-c)/b)))/((gamma (a))*(b^a))
> i <- f/sum (f)
x non repeated response of job satisfaction
> y <- c(20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32,
33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48,
49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59)
> e <- 3.46
> d <- 11.42
> h <- 0
> g <- ((y-h)^(d-1))*(exp (-((y-h)/e)))/((gamma
(d))*(e^d))
> j <- (g/sum (g))
y non repeated response of respondents age.
```

### APPENDIX C: Code to obtain the scatter plot

```
Job > z1 <- c(74, 78, 74, 52, 82, 75, 79, 69, 66, 65, 66, 67,
72, 61, 69, 69, 69, 69, 65, 69, 66, 69, 62, 69, 71, 68, 73, 68,
73, 71, 59, 72, 67, 58, 79, 61, 74, 65, 70)
Age > z2 <- c(20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31,
32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47,
48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59)
> a1 <- 22.10
> b1 <- 2.824
> k1 <- 0.9
> a2 <- 11.42
> b2 <- 3.46
> k2 <- 1
> x <- (z1^k1)
> y <- (z2^k2)
```

```

> job <- (1/k1)*(Job^((a1/k1)-1))*(exp
(-(Job^(1/k1))/b1)))/((gamma (a1))*(b1^a1))
> Age1 <- (1/k2)*(Age^((a2/k2)-1))*(exp
(-(Age^(1/k2))/b2)))/((gamma (a2))*(b2^a2))
> jobage <- (job/Age1)*100
> plot (Age, jobage)

```

## REFERENCES

- [1] Stansfeld, S. and Bridget, C. Psychosocial Work Environment and Mental Health-A Meta-Analytic Review. *Scand J Work Environ Health* 2006, 32(6, special issue): 443-462.
- [2] Saziye, G. and Tansel, A. Job Satisfaction in Britain: individual and Job related factors. ERC working papers in Economics. Available from: [www.erc.meta.edu.tr](http://www.erc.meta.edu.tr). (Accessed 10 April 2011).
- [3] Buitendach, J. H. and Rothmann, S. The validation of the Minnesota Job Satisfaction Questionnaire in selected organizations in South Africa 2009. *SA Journal of Human Resource Management*, 7: 183-191.
- [4] Paul, E., Pakivathy, P. E. and Phua, S. K. 'Lecturers' job satisfaction in a public tertiary institution in Singapore: ambivalent and non-ambivalent relationships between job satisfaction and demographic variables', *Journal of Higher Education Policy and Management* 2011, 33(2): 141-151.
- [5] Johnson, R.A. and Wichern, D.W. *Applied Multivariate Statistical Analysis*. Englewood Cliffs, NJ: Prentice-Hall, Inc., 2006.
- [6] Mohammad Zakir Hossain. The Use of Box-Cox Transformation Technique in Economic and Statistical Analyses. *Journal of Emerging Trends in Economics and Management Sciences (JETEMS)* 2011. 2(1):32-39.
- [7] Steven M. La Londe. Transforming Variables for Normality and Linearity. *SAS Global Forum* 2012, Rochester Institute of Technology, Rochester, NY.
- [8] Sakia R.M. Review of Box-Cox transformation techniques. *The statistician* volume 41, pp. 169-178, 1992.
- [9] Krishnamoorthy K. *Handbooks of statistical distributions with applications* 2006. Taylor and Francis Group. Chapman and Hall / CRC New York.
- [10] Jeffrey N. Jonkman, Patrick D. Gerard and William H. Swallow (2009); *Journal Computational Statistics and Data Analysis* archive Volume 53 Issue 4, Pages 1099-1109.
- [11] Clifford, T.M., Richard, A.K., John, R.W and John, S. *Introduction to Psychology* 7<sup>th</sup> edition, New Delhi. Tata McGraw-Hill publishing company limited 1993.