RT&A, No 4 (89)
Volume 20, December 2025

STRESS-STRENGTH RELIABILITY UNDER MULTIVARIATE LOG-NORMAL SETUP AND ITS APPLICATION

Anirban Goswami[1], Babulal Seal[2]

•

[1]Department of Statistics, Regional Research Institute of Unani Medicine, Kolkata, India.
[2]Department of Mathematics and Statistics, Aliah University, Kolkata, India.
[1]anirbanstat09@gmail.com
[2]babulal.seal@gmail.com

Abstract

In this article it is mainly focused on discussion about estimation of stress-strength reliability (R) under log-normal multivariate setup. We propose a procedure to compute and estimate stress-strength reliability of weighted geometric function when both stress and strength follow a dependent log-normal multivariate distribution with two weightage vectors. We consider the principal component analysis to estimate these vectors. MVUE, MLE and Confidence Intervals of R are obtained. Through simulation studies, their performances are compared using different measures. Finally, we provide a real data analysis.

Keywords: Stress-strength, Principal component, Maximum Likelihood Estimator (MLE), Minimum Variance Unbiased Estimator (MVUE), Confidence Intervals.

I. Introduction

The log-normal distribution is characterized as the distribution of a random variable whose logarithm is normally distributed. The normal distribution was used in engineering stress-strength models. But, the stress-strength model with log-normal distribution for both stress and strength are widely used in engineering domains [1]. However, it has been gradually replaced by the lognormal distribution model, because of the more realistic features as the positiveness of its values and the positive skewness [2] of its shape. The log-normal distribution has also been widely applied to real life data fitting in an empirical way. Hence, many variables in real life, from the sizes of organisms and the numbers of species in biology to rainfalls in meteorology and sizes of incomes in economics, are positive in nature. Although, it has been derived by Gibrat [3] theoretically using qualitative assumptions, calling it the law of proportionate effect. Kolmogoroff [4] derived the distribution using the asymptotic result of an iterative process of successive breakage of a particle into two randomly sized particles. The frequently use of log-normal distributions in material properties and fatigue lives [5]. Lu et al. [6] used the log-normal probability density function in the robust design of an arbor. This distribution is typically adopted for load variable [7] to the analysis of ship structural reliability, the random variables that always take positive values.

For satellites data, it is seen that data is more log-normally rather than normally distributed [8-10] and for the first time, a description for lognormal errors was given for NWP. In addition, it may show that log-normal variables in NWP is a modern problem. There were data sets in the 1970s [11] that were recognized as log-normally distributed. A score test for testing the equality of two means

DOI: https://doi.org/10.24412/1932-2321-2025-489-449-470

of two independent log-normal distributions presented by Gupta and Li [12] compared the results with those of Zhou et al. [14]. It is showed that, the bivariate log-normal distribution is more appropriate in the model of outpatient to test the equality of means of costs, six months before and after the medicaid policy change in the state of Indiana in a paired design study for a paired design study (Zhou et al., 2001). Schneider and Holst [15] and Cheng [16] demonstrate the uses of log normal distribution in modeling of the size distribution of aerosol particles and airborne fibers.

Log-normal distributions is used in the air pollution concentration data, for predicting the future air quality and it is important aspect to give the best prediction. In addition, the log-normal probability density functions have been used for many years to analyze the air pollution and describe the frequency of high concentration of the pollutant [17]. Norazian et al. [18] found Log-normal distribution as the best fit in the data sets of PM10, for 2006 and 2007, in an industrialized area with high population and traffic density. Sedek et al. [19] analyzed the data of PM_{10} concentrations in Kuala Lumpur from 1998 to 2002 and found that the log-normal distribution as the best-fitted model to prediction. Also, Hamid et al. [20] used log normal to fit the distribution. There are many types of statistical distribution that can be used to fit the air quality data. Georgopoulus and Seinfeld [21] explored and described the statistical distribution of air quality of pollutant concentration.

Lai et al. [22] used the lognormal probability distribution is a function of geometric mean and geometric standard deviation to obtain the distribution properties from observed data. Nguyen and Sevando [23] used the weighted geometric mean function to integrate sub-indices of 8 parameters for Coastal Water Quality and these water quality Parameters sleeted by their weights. Generally, eight pollutants namely particulate matter [24] PM10, PM2.5, nitrogen dioxide (NO2), ammonia (NH3), carbon monoxide (CO), Sulphur dioxide (SO2) Ozone (O3) and Lead (Pb) act as major parameters in deriving the air quality index (AQI) of an area. Generally, each of the AQI parameters follows the log-normal distribution and they jointly follow the multivariate log-normal distribution. Also, multivariate lognormal distribution used in a particular case of the PERT type problems [25], where random vectors for which the component lifetime distributions are lognormal. The multivariate log-normal distribution used in astronomy data of the Hertzprung-Russel [26] diagram of the star clusters CYG OB1, where the first component is the logarithm of the effective temperature at the surface of the star and the second is the logarithm of its light intensity.

Gupta and Gupta [27] first introduced the concept to estimate the reliability under multivariate normal setup. They considered the forms of R when $\left(\mathbf{x}^*_{p_1 x1}, \mathbf{y}^*_{p_2 x1}\right)'$ follows multivariate normal distribution with dependent vector between $\mathbf{x}^*_{p_1 x1}$ and $\mathbf{y}^*_{p_2 x1}$. Then, the reliability as R= Pr($\mathbf{a}'\mathbf{x}^* > \mathbf{b}'\mathbf{y}^*$) , where $\mathbf{a}'$ and $\mathbf{b}'$ are two vectors. This problem arises when a system in the energy is supplied to the system by p_1 sources and is consumed through p_2 sources and the sources of energy supplied and consumed are linearly dependent with two vectors $\mathbf{a}'$ and $\mathbf{b}'$. Under this set up, they obtained and compared the MVUE and MLE estimate of R with some interesting special cases. In this multivariate setup, Reiser and Farragi [28] derived the lower confidence bounds for R=Pr($\mathbf{a}'\mathbf{x}^* > \mathbf{b}'\mathbf{y}^*$) and solved it iteratively and also derived an approximate lower confidence bounds for R. Enis and Geisser [29] demonstrated that, how to obtain the exact confidence bounds for R. Mukherjee and Sharan [30] obtained the UMVUE for R under the bivariate normal distribution and also obtained their asymptotic variance when parameters other than the means are known and they proposed an estimator R based on maximum likelihood estimators when all the five parameters are unknown. Hor and Seal [31] derived an alternative estimator viz. UMVUE of R under the same case of bivariate normal distribution.

In the multivariate log-normal distribution setup, we obtain the stress-strength reliability of weighted geometric function R=Pr$\left(\prod_{i=1}^{p_1} x_i^{a_i} > \prod_{j=1}^{p_2} y_j^{b_j}\right)$, where $\left(\mathbf{x}_{p_1 x1}, \mathbf{y}_{p_2 x1}\right)'$ follows multivariate log-normal distribution with dependence vector between $\mathbf{x}_{p_1 x1}$ and $\mathbf{y}_{p_2 x1}$ with $\mathbf{a}'$ and $\mathbf{b}'$ are two vectors. We take the log-transformation of the weighted geometric function and consider the principal component analysis to estimate the $\mathbf{a}'$ and $\mathbf{b}'$, where Gupta and Gupta [27] considered only spatial cases of $\mathbf{a}'$ and $\mathbf{b}'$. The estimators of R obtained using MVUE and MLE of the parameters. We do simulation studies to compare their performance in teams of variance (VAR), mean square error (MSE) and mean absolute error (MAE). We take different choices of parameter to obtain the L_1 distance, same form used by Hor and Seal [31]. In this connection, the distributions function of the two estimators MVUE and MLE are derived in section 3. L_1 distance between two functions to compare two estimators is given in section 4. The graphically representation of variance of MVUE and MSE of MLE shows in figure 2 and 3, also it shows that their performances are better even in values neighborhood 0 of $\sqrt{n}\delta^*$. Similarly, relationship between ratio of these two measures and $\sqrt{n}\delta^*$ is given in table 4. Two-sided confidence limits and lower bounds of R defined in section 5. Based on MVUE and MLE, we compare the performance between bootstrap and empirically interval estimator in terms of coverage and accuracy using simulation study and real data set.

II. Methods

I. Maximum Likelihood Estimator of R

Let us consider, a vector $\mathbf{w}=\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}_{p_1+p_2}$ of positive random variables such that $\mathbf{z}=\begin{pmatrix}\ln(\mathbf{x})\\ \ln(\mathbf{y})\end{pmatrix}_{p_1+p_2}$ has an (p_1+p_2)-dimensional normal distribution with mean vector $\mu=\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}_{(p_1+p_2)x1}$ and variance-covariance matrix $\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}_{(p_1+p_2)x(p_1+p_2)}$ and denoted by $\Lambda_{p1+p2}(\mu,\Sigma)$. The corresponding (p_1+p_2)-dimensional normal distribution is denoted by $\begin{pmatrix}\ln(\mathbf{x})\\ \ln(\mathbf{y})\end{pmatrix}_{p_1+p_2}=\begin{pmatrix}\mathbf{x}^*\\ \mathbf{y}^*\end{pmatrix}_{p_1+p_2}\sim N_{p1+p2}(\mu,\Sigma)$.

The probability density function of $\mathbf{w}=\begin{pmatrix}\mathbf{x}\\ \mathbf{y}\end{pmatrix}_{p_1+p_2}$ having $\Lambda_{p1+p2}(\mu,\Sigma)$ is

$$f(\mathbf{w})=\frac{1}{(2\pi)^{(p_1+p_2)/2}\sqrt{|\Sigma|}\prod_{i=1}^{p_1+p_2}w_i}\exp\left\{-\frac{1}{2}(\ln\mathbf{w}-\boldsymbol{\mu})'\Sigma^{-1}(\ln\mathbf{w}-\boldsymbol{\mu})\right\}$$

Let, the two vectors are $\mathbf{a}'$ and $\mathbf{b}'$. Then, we want to find the probability of two weighted geometric function $\mathbf{a}'\mathbf{x}$ and $\mathbf{b}'\mathbf{y}$ as

$$\begin{aligned}R&=\Pr\left(\prod_{i=1}^{p_1}x_i^{a_i}>\prod_{j=1}^{p_2}y_j^{b_j}\right)=\Pr\left(\ln\prod_{i=1}^{p_1}x_i^{a_i}>\ln\prod_{j=1}^{p_2}y_j^{b_j}\right)\\&=\Pr\left(\sum_{i=1}^{p_1}a_i\ln(x_i)>\sum_{j=1}^{p_2}b_j\ln(y_j)\right)\\&=\Pr(\mathbf{a}'\mathbf{x}^*-\mathbf{b}'\mathbf{y}^*>0)\end{aligned}$$

Now, the distribution of $v=\mathbf{a}'\mathbf{x}^*-\mathbf{b}'\mathbf{y}^*$ follows $N(\mu_v,\sigma_v^2)$,
where, $\mu_v=E(\mathbf{a}'\mathbf{x}^*-\mathbf{b}'\mathbf{y}^*)=\mathbf{a}'\boldsymbol{\mu}_1-\mathbf{b}'\boldsymbol{\mu}_2$ and $\sigma_v^2=Var(\mathbf{a}'\mathbf{x}^*-\mathbf{b}'\mathbf{y}^*)=\mathbf{a}'\Sigma_{11}\mathbf{a}-2\mathbf{a}'\Sigma_{12}\mathbf{b}+\mathbf{b}'\Sigma_{22}\mathbf{b}$.

Now, $R=\Pr(\mathbf{a}'\mathbf{x}^*-\mathbf{b}'\mathbf{y}^*>0)=\Pr(v>0)$

$$=\int_0^\infty\frac{1}{\sqrt{2\pi}\sigma_v}\exp\left\{-\frac{1}{2}\left(\frac{v-\mu_v}{\sigma_v}\right)^2\right\}dv=\int_{-\frac{\mu_v}{\sigma_v}}^\infty\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}z^2\right\}dz=\Phi\left(\frac{\mu_v}{\sigma_v}\right)$$

The maximum likelihood estimator of $\mu=\begin{pmatrix}\boldsymbol{\mu}_1\\ \boldsymbol{\mu}_2\end{pmatrix}_{(p_1xp_2)}$ and $\Sigma=\begin{pmatrix}\Sigma_{11} & \Sigma_{12}\\ \Sigma_{21} & \Sigma_{22}\end{pmatrix}_{(p_1+p_2)x(p_1+p_2)}$

are $\begin{pmatrix}\overline{\mathbf{x}^*}\\ \overline{\mathbf{y}^*}\end{pmatrix}$ and $\mathbf{S} = \begin{pmatrix}\mathbf{S_{11}} & \mathbf{S_{12}}\\ \mathbf{S_{21}} & \mathbf{S_{22}}\end{pmatrix}$ respectively.

We have, $\widehat{\mu_v} = \mathbf{a}'\overline{\mathbf{x}^*} - \mathbf{b}'\overline{\mathbf{y}^*}$ and $\widehat{\sigma_v^2} = \mathbf{a}'\mathbf{S_{11}}\mathbf{a} - 2\mathbf{a}'\mathbf{S_{12}}\,\mathbf{b} + \mathbf{b}'\mathbf{S_{22}}\,\mathbf{b}$

So, the maximum likelihood estimate of R is define as R*= $\Phi\left(\frac{\mu_v}{\sigma_v}\right)$ (1)

where Φ = Distribution function of univariate standard normal distribution.

II. Estimation by Principal Component method

The overall representation of the two sets or vectors are related to vectors **a** and **b**, such that they are approximated by $\mathbf{a}'\mathbf{x}^*$ and $\mathbf{b}'\mathbf{y}^*$ through principal component analysis. Principal component analysis explaining the variance-covariance structure $\sum_{11}$ & $\sum_{22}$ of a set of variables $\mathbf{x}^*$ and $\mathbf{y}^*$ through a linear function of these variables, i.e, explain maximum variability. Traditional way is to take first principal component to estimate $\mathbf{a}'$ by $\mathbf{e}_1'$ normalized eigenvector of $\sum_{11}$ corresponding to eigen value λ_1 and $\mathbf{b}'$ by $\mathbf{l}_1'$ normalized eigenvector of $\sum_{22}$ corresponding to eigen value λ_1 [32]. We take the maximum likelihood estimate of $\sum_{11}$ as $\mathbf{S}_{11}^*$ and take estimate of $\mathbf{a}'$ by $\mathbf{e}_1'$ normalized eigenvectors of $\mathbf{S}_{11}^*$ corresponding to eigen value λ_1. Similarly, we have estimate of $\sum_{22}$ as $\mathbf{S}_{22}^*$ and take estimate of $\mathbf{b}'$ by $\mathbf{l}_1'$ normalized eigenvectors of $\mathbf{S}_{22}^*$ corresponding to λ_1 eigen value. Then from equation (1), the estimate of R becomes

$$R^* = \Phi\left(\frac{\widehat{\mu_v}}{\widehat{\sigma_v}}\right), \text{ where } \widehat{\mu_v} = \mathbf{e}_1'\overline{\mathbf{x}^*} - \mathbf{l}_1'\overline{\mathbf{y}^*} \text{ and } \widehat{\sigma_v^2} = \mathbf{e}_1'\mathbf{S_{11}}\mathbf{e_1} - 2\mathbf{e}_1'\mathbf{S_{12}}\,\mathbf{l_1} + \mathbf{l}_1'\mathbf{S_{22}}\,\mathbf{l_1}\,. \quad (2)$$

III. Minimum Variance Unbiased Estimator (MVUE) of R

We determine the Minimum Variance Unbiased Estimator (MVUE) of R=Pr($\mathbf{a}'\mathbf{x}^* - \mathbf{b}'\mathbf{y}^* > 0$). Here, it is assumed that the random sample $\begin{pmatrix}\mathbf{x}_\alpha^*\\ \mathbf{y}_\alpha^*\end{pmatrix}$, $\alpha = 1,2,\ldots,n$ are from multivariate normal distribution i.e. $\begin{pmatrix}\mathbf{x}_\alpha^*\\ \mathbf{y}_\alpha^*\end{pmatrix} \sim N_{p1+p2}(\mu, \Sigma)$.

Then, $v_\alpha = (\mathbf{b}'\mathbf{y}_\alpha^* - \mathbf{a}'\mathbf{x}_\alpha^*) \sim N(\mu_v, \sigma_v^2)$, α=1,2,….,n be the random sample of size n. Now, ($\bar{v}$, S_v^2) is a complete sufficient statistic for (μ_v, σ_v^2), where $\bar{v} = \frac{1}{n}\sum_{\alpha=1}^{n} v_i$ and $s_v^2 = \frac{1}{n}\sum_{\alpha=1}^{n}(v_i - \bar{v})^2$, the MVUE of R [33] is $\hat{R} = \int_c^1 \frac{\Gamma(\frac{n-1}{2})}{\Gamma(\frac{1}{2})\Gamma(\frac{n-2}{2})}(1-z^2)^{\frac{n-2}{2}-1}dz$, where $c = \frac{\bar{v}}{(\sqrt{(n-1)}s_v}$

Then the MVUE of R=Pr($\mathbf{a}'\mathbf{x}^* - \mathbf{b}'\mathbf{y}^* > 0$) is [27]

$$\hat{R} = \begin{pmatrix} 0 & \text{if} & c > 0 \\ \frac{1}{2}\left(1 - B\left(c^2; \frac{1}{2}, \frac{n-2}{2}\right)\right) & \text{if} & 0 < c \leq 1 \\ \frac{1}{2}\left(1 + B\left((-c)^2; \frac{1}{2}, \frac{n-2}{2}\right)\right) & \text{if} & -1 < c \leq 0 \\ 1 & \text{if} & c \leq -1 \end{pmatrix} \quad (3)$$

where, $c = \frac{(\boldsymbol{b}'\overline{\boldsymbol{y}^*} - \boldsymbol{a}'\overline{\boldsymbol{x}^*})}{(\sqrt{(n-1)}\ (\boldsymbol{a}'S_{11}\boldsymbol{a} - 2\boldsymbol{a}'S_{12}\,\boldsymbol{b} + \boldsymbol{b}'S_{22}\,\boldsymbol{b})^{1/2}}$ and $B(k;\alpha,\beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^k x^{\alpha-1}(1-x)^{\beta-1}\,dx$

An estimator $\hat{\hat{R}}$ may be taken as above expression where 'c' is replaced by

$$c' = \frac{(l_1'\overline{y^*} - e_1'\overline{x^*})}{(\sqrt{(n-1)}\ (e_1'S_{11}e_1 - 2e_1'S_{12}\,l_1 + l_1'S_{22}\,l_1)^{1/2}}\,.$$

IV. Distribution Function of $\hat{R}$ and R^*

In this section, we use results of the previous section by taking of log-transformation between vectors and derive the distributions function of $\hat{R}$ and R^*. Consider the log-transformed vectors as $\mathbf{x}^*$ and $\mathbf{y}^*$. Then, from section 2.3, the distribution function of $\hat{R}$ is given by

Let the distribution function of $\hat{R}$ be $F_{\hat{R}}(x)$, then

$$F_{\hat{R}}(x) = 0 \text{ if } x < 0 \text{ and } F_{\hat{R}}(x) = 1 \text{ if } x \geqslant 1$$

If $0 \leqslant x \leqslant \frac{1}{2}$, then the distribution function of $\hat{R}$ is given by

$$F_{\hat{R}}(x) = P(\hat{R} \leqslant x) = P[(\hat{R} \leqslant x)|(c > 0)]P[c > 0] + P[(\hat{R} \leqslant x)|(c \leqslant 0)]P[c \leqslant 0]$$

$$= P[\frac{1}{2} - \frac{1}{2}B(c^2; \frac{1}{2}, \frac{n-2}{2}) \leqslant x]P[\frac{(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}^*)}{\sqrt{(n-1)}(\mathbf{a}'\mathbf{S_{11}a} - 2\mathbf{a}'\mathbf{S_{12}b} + \mathbf{b}'\mathbf{S_{22}b})^{\frac{1}{2}}} > 0]$$

$$= P[B(c^2; \tfrac{1}{2}, \tfrac{n-2}{2}) \geqslant 1 - 2x]P[(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}^*) > 0]$$

$$\text{,where } (\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}^*) \sim N_1((\mathbf{b}'\boldsymbol{\mu}_2 - \mathbf{a}'\boldsymbol{\mu}_1), \tfrac{1}{n}(\mathbf{a}'\boldsymbol{\Sigma}_{11}\mathbf{a} - 2\mathbf{a}'\boldsymbol{\Sigma}_{12}\mathbf{b} + \mathbf{b}'\boldsymbol{\Sigma}_{22}\mathbf{b}))$$

$$= P[c^2 \geqslant B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x)]\Phi(\frac{\sqrt{n}(\mathbf{b}'\boldsymbol{\mu}_2 - \mathbf{a}'\boldsymbol{\mu}_1)}{(\mathbf{a}'\boldsymbol{\Sigma}_{11}\mathbf{a} - 2\mathbf{a}'\boldsymbol{\Sigma}_{12}\mathbf{b} + \mathbf{b}'\boldsymbol{\Sigma}_{22}\mathbf{b})^{\frac{1}{2}}})$$

$$= \{P[c \geqslant \left(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x)\right)^{\frac{1}{2}}] + P[c \leqslant -\left(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x)\right)^{\frac{1}{2}}]\}\Phi(-\sqrt{n}\delta)$$

$$\text{,where } \delta^* = \frac{-(\mathbf{b}'\boldsymbol{\mu}_2 - \mathbf{a}'\boldsymbol{\mu}_1)}{(\mathbf{a}'\boldsymbol{\Sigma}_{11}\mathbf{a} - 2\mathbf{a}'\boldsymbol{\Sigma}_{12}\mathbf{b} + \mathbf{b}'\boldsymbol{\Sigma}_{22}\mathbf{b})^{\frac{1}{2}}}$$

$$= \{P[\frac{(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}^*)}{\sqrt{(n-1)}(\mathbf{a}'\mathbf{S}^*_{11}\mathbf{a} - 2\mathbf{a}'\mathbf{S}^*_{12}\mathbf{b} + \mathbf{b}'\mathbf{S}^*_{22}\mathbf{b})^{\frac{1}{2}}} \geqslant (B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}] +$$

$$P[\frac{(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}^*)}{\sqrt{(n-1)}(\mathbf{a}'\mathbf{S}^*_{11}\mathbf{a} - 2\mathbf{a}'\mathbf{S}^*_{12}\mathbf{b} + \mathbf{b}'\mathbf{S}^*_{22}\mathbf{b})^{\frac{1}{2}}} \leqslant -(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}]\}\ \Phi(-\sqrt{n}\delta^*)$$

$$= \{P[-\sqrt{n}\hat{\delta}^* \geqslant (n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}] + P[-\sqrt{n}\hat{\delta}^* \leqslant -(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}]\}\Phi(-\sqrt{n}\delta^*)$$

$$\text{,where } \hat{\delta}^* = \frac{-\sqrt{(n-1)}(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}^*)}{\sqrt{n}(\mathbf{a}'\mathbf{S}^*_{11}\mathbf{a} - 2\mathbf{a}'\mathbf{S}^*_{12}\mathbf{b} + \mathbf{b}'\mathbf{S}^*_{22}\mathbf{b})^{\frac{1}{2}}}$$

$$= \{P[\sqrt{n}\hat{\delta}^* \leqslant -(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}] + 1 - P[\sqrt{n}\hat{\delta}^* \leqslant (n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}]\}\Phi(-\sqrt{n}\delta^*)$$

$$= \{F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}) + [1 - F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}})]\}\Phi(-\sqrt{n}\delta^*)$$

Using the standard distribution theory, if $\mathbf{x}^* \sim N_p(\mu, \Sigma)$ then $\mathbf{a}'\mathbf{x}^* \sim N_1(\mathbf{a}'\mu, \mathbf{a}'\Sigma\mathbf{a})$ [34]. Let, $\bar{\mathbf{x}}^*$ and $\mathbf{S}^*$ be the unbiased estimator of μ and Σ respectively, then $\mathbf{a}'\bar{\mathbf{x}}^* \sim N_1(\mathbf{a}'\mu, \frac{1}{n}(\mathbf{a}'\Sigma\mathbf{a}))$ and

$\mathbf{S}^* \sim W_p(n-1,\Sigma)$. Thus, we can write, $\frac{\mathbf{a'S^*a}}{\mathbf{a'\Sigma a}} \sim \boldsymbol{\chi}^2_{n-1}$, hence $\sqrt{n-1}\hat{\delta}^* \sim t'_{(n-1),\sqrt{n}\delta^*}$ where $t'_{(n-1),\sqrt{n}\delta^*}$ denotes the non - central t - distribution with (n - 1) d.f. We use the unbiased estimator of Σ instead of MLE, then $\sqrt{n}\hat{\delta}^* \sim t'_{(n-1),\sqrt{n}\delta^*}$ with non-centrality parameter $\sqrt{n}\delta^*$ and $F_{t'_{(n-1),\sqrt{n}\delta^*}}(.)$ be the cdf of non - central t - distribution.

If $\frac{1}{2} < x < 1$, then the distribution function of $\hat{R}$ is given by

$$F_{\hat{R}}(x) = P(\hat{R} \leqslant x) = P[(\hat{R} \leqslant x)|(c>0)]P[c>0] + P[(\hat{R} \leqslant x)|(c \leqslant 0)]P[c \leqslant 0]$$

$$= P[\frac{1}{2} - \frac{1}{2}B(c^2; \frac{1}{2}, \frac{n-2}{2}) \leqslant x]P[\frac{(\mathbf{b'}\bar{\mathbf{y}}^* - \mathbf{a'}\bar{\mathbf{x}}^*)}{\sqrt{(n-1)}(\mathbf{a'S^*_{11}a} - 2\mathbf{a'S^*_{12}b} + \mathbf{b'S^*_{22}b})^{\frac{1}{2}}} > 0] +$$

$$P[\frac{1}{2} + \frac{1}{2}B((-c)^2; \frac{1}{2}, \frac{n-2}{2}) \leqslant x]\Phi(\sqrt{n}\delta^*)$$

$$= \Phi(-\sqrt{n}\delta^*) + P[B(c^2; \frac{1}{2}, \frac{n-2}{2}) \leqslant 2x-1]\Phi(\sqrt{n}\delta^*)$$

$$= \Phi(-\sqrt{n}\delta^*) + P[-(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}} \leqslant c \leqslant (B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}}]\Phi(\sqrt{n}\delta^*)$$

$$= \Phi(-\sqrt{n}\delta^*) + P[(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}} \geqslant \sqrt{n}\hat{\delta}^* \geqslant -(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}}]\Phi(\sqrt{n}\delta^*)$$

$$= \Phi(-\sqrt{n}\delta^*) + [F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}}) - F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}})]\Phi(\sqrt{n}\delta^*)$$

Thus, the distribution function of $\hat{R}$ is given by

$$F_{\hat{R}}(x) = 0 \text{ if } x < 0 \text{ and } F_{\hat{R}}(x) = 1 \text{ if } x \geqslant 1$$

If $0 \leqslant x \leqslant \frac{1}{2}$,

$$F_{\hat{R}}(x) = \{F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}) + [1 - F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}})]\}\Phi(-\sqrt{n}\delta^*)$$

If $\frac{1}{2} < x < 1$,

$$F_{\hat{R}}(x) = \Phi(-\sqrt{n}\delta^*) + [F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}}) - F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}})]\Phi(\sqrt{n}\delta^*) \quad (4)$$

The MLE estimate of R, $R^* = \Phi(\frac{-(\mathbf{b'}\bar{\mathbf{y}}^* - \mathbf{a'}\bar{\mathbf{x}}^*)}{(\mathbf{a'S^*_{11}a} - 2\mathbf{a'S^*_{12}b} + \mathbf{b'S^*_{22}b})^{\frac{1}{2}}})$

Let the distribution function of R^* be $F_{R^*}(x)$, then $F_{R^*}(x) = 0$ if $x < 0$ and $F_{R^*}(x) = 1$ if $x \geqslant 1$,

$$F_{R^*}(x) = P(R^* \leqslant x) = P[\Phi(\frac{-(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}*)}{(\mathbf{a}'\mathbf{S}_{11}^*\mathbf{a} - 2\mathbf{a}'\mathbf{S}_{12}^*\mathbf{b} + \mathbf{b}'\mathbf{S}_{22}^*\mathbf{b})^{\frac{1}{2}}}) \leqslant x]$$

$$= P[\frac{-(\mathbf{b}'\bar{\mathbf{y}}^* - \mathbf{a}'\bar{\mathbf{x}}*)}{(\mathbf{a}'\mathbf{S}_{11}^*\mathbf{a} - 2\mathbf{a}'\mathbf{S}_{12}^*\mathbf{b} + \mathbf{b}'\mathbf{S}_{22}^*\mathbf{b})^{\frac{1}{2}}} \leqslant \Phi^{-1}(x)]$$

$$= P[\sqrt{n}\hat{\delta}^* \leqslant \sqrt{(n-1)}\Phi^{-1}(x)]$$

$$= F_{t'_{(n-1),\sqrt{n}\delta^*}}(\sqrt{(n-1)}\Phi^{-1}(x)) \qquad (5)$$

V. Distance Between $F_{\hat{R}}(.)$ and $F_{R^*}(.)$

Let us calculate the distance between two function (L_1) $F_{\hat{R}}(x)$ and $F_{R^*}(x)$ [31], $U(n, \delta^*) = \int_0^1 |F_{R^*}(x) - F_{\hat{R}}(x)|dx$, for different values of n, $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{12}$, $\boldsymbol{\Sigma}_{22}$. According to this method, this can be taken as a measure of deviation or equal deviation between $\hat{R}$ and R^*.

Let, $\bar{R}$ be any other estimator of R, then the maximum deviation between distribution of $\bar{R}$ and $\hat{R}$ as

$$M(n, \delta^*) = \sup_{\bar{R}} \int_0^1 |F_{\bar{R}}(x) - F_{\hat{R}}(x)|\, dx$$
$$= \int_0^1 F_{\hat{R}}(x)dx\,, \text{if} \int_0^1 F_{\hat{R}}(x)dx > \frac{1}{2}$$
$$= 1 - \int_0^1 F_{\hat{R}}(x)dx\,, \text{if} \int_0^1 F_{\hat{R}}(x)dx \leqslant \frac{1}{2}$$

The ration $R(n, \delta^*) = \frac{U(n,\delta^*)}{M(n,\delta^*)}$ has to be taken as a relative measure of deviation between $\hat{R}$ and R^*, the maximum deviation between any other estimator of R, i.e. $\bar{R}$ and $\hat{R}$.

VI. Derivation of Var($\hat{R}$) and MSE(R^*)

Since, $0 < \hat{R} < 1$, then $E(\hat{R}) = \int_0^1 \{1 - F_{\hat{R}}(x)\}dx$ and using equation (4)

If $0 \leqslant x \leqslant \frac{1}{2}$, then we have

$$E_1(\hat{R}) = \int_0^1 [1 - \{F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}) + [1 - F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}})]\}\Phi(-\sqrt{n}\delta^*)]\, dx$$

If $\frac{1}{2} < x < 1$, then we have

$$E_2(\hat{R}) = \int_0^1 [1 - \{\Phi(-\sqrt{n}\delta^*) + [F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}}) - F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}})]\Phi(\sqrt{n}\delta^*)\}]\,dx$$

So,

$$Var(\hat{R}) = \int_0^1 2x\{1 - F_{\hat{R}}(x)\}dx - \{E(\hat{R})\}^2$$

$$= \int_0^1 2x[1 - \{F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}}) + [1 - F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(1-2x))^{\frac{1}{2}})]\}\Phi(-\sqrt{n}\delta^*)]\,dx$$

$$+ \int_0^1 2x[1 - \{\Phi(-\sqrt{n}\delta^*) + [F_{t'_{(n-1),\sqrt{n}\delta^*}}((n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}}) - F_{t'_{(n-1),\sqrt{n}\delta^*}}(-(n-1)(B^{-1}_{(\frac{1}{2},\frac{n-2}{2})}(2x-1))^{\frac{1}{2}})]\Phi(\sqrt{n}\delta^*)\}]\,dx$$

$$- \{E_1(\hat{R}) + E_2(\hat{R})\}^2$$

Using equation (5), we can determine $MSE(R^*) = \int_0^1 2x\left\{1 - F_{t'_{(n-1),\sqrt{n}\delta^*}}\left(\sqrt{(n-1)}\Phi^{-1}(x)\right)\right\}dx - [\int_0^1 \{1 - F_{t'_{(n-1),\sqrt{n}\delta^*}}(\sqrt{(n-1)}\Phi^{-1}(x))\}\,dx]^2$

VII. Exact Two Sided Confidence Intervals for R

Here, we have $\sqrt{n}\hat{\delta}^* \sim t'_{(n-1),\sqrt{n}\delta^*}$. The tradition approach for finding the lower limit of R , we use the probability $p_{\delta^*_L}$ that $t'_{(n-1),\sqrt{n}\delta^*}$ exceeds the value of $\sqrt{n}\hat{\delta}^*$ as

$$p_{\delta^*_L} = Pr\,(t'_{(n-1),\sqrt{n}\delta^*} > \sqrt{n}\hat{\delta}^*) = \alpha/2$$

or,

$$p_{\delta^*_L} = Pr\,(t'_{(n-1),\sqrt{n}\delta^*} < \sqrt{n}\hat{\delta}^*) = 1 - \alpha/2 \quad (6)$$

Similarly, we get the upper limit of δ^* as

$$p_{\delta^*_U} = Pr\,(t'_{(n-1),\sqrt{n}\delta^*} < \sqrt{n}\hat{\delta}^*) = \alpha/2 \quad (7)$$

Equation (6) and (7) can be solved numerically. Finally, we get the (1-α) level confidence Intervals for δ^* as (δ^*_L, δ^*_U). Then, the (1-α) level confidence Intervals for R as $(\Phi(\delta^*_L), \Phi(\delta^*_U))$.

VIII. Exact Lower Confidence bound for R

In order to determine the lower bound of the lower bound of R, we use the probability $p_{\delta^*_{LB}}$ that $t'_{(n-1),\sqrt{n}\delta^*}$ exceeds the value of $\sqrt{n}\hat{\delta}^*$ as

$$p_{\delta^*_{LB}} = Pr\,(t'_{(n-1),\sqrt{n}\delta^*} > \sqrt{n}\hat{\delta}^*) = \alpha$$

or, $$p_{\delta^*_{LB}} = \Pr\left(t'_{(n-1),\sqrt{n}\delta^*} < \sqrt{n}\hat{\delta}^*\right) = 1-\alpha \tag{8}$$

Thus, the (1-α) level confidence lower bound for δ^* can be obtained by solving equation (8). Then, the (1-α) level confidence lower bound for R is $\Phi(\delta^*_{LB})$.

IX. Approximate Two Sided Confidence Intervals for R

From section 3, we have $\Pr(\mathbf{a}'\mathbf{x}^* > \mathbf{b}'\mathbf{y}^*) = \Phi\left[\frac{-(\mathbf{b}'\boldsymbol{\mu}_2 - \mathbf{a}'\boldsymbol{\mu}_1)}{(\mathbf{a}'\boldsymbol{\Sigma}_{11}\mathbf{a} - 2\mathbf{a}'\boldsymbol{\Sigma}_{12}\mathbf{b} + \mathbf{b}'\boldsymbol{\Sigma}_{22}\mathbf{b})^{\frac{1}{2}}}\right] = \Phi(\delta^*)$,

where $\sqrt{n}\hat{\delta}^* \sim t'_{(n-1),\sqrt{n}\delta^*}$ with non-centrality parameter $\sqrt{n}\delta^*$.

In order to determine the two sided confidence Intervals, we use following well known approximation for large n [35] as

$$Z = \frac{\left[t'_{(n-1),\sqrt{n}\delta^*} - \sqrt{n}\delta^*\right]}{\left[1 + \frac{(t'_{(n-1),\sqrt{n}\delta^*})^2}{2(n-1)}\right]^{\frac{1}{2}}} \sim N(0,1)$$

Using this,

$$\Pr\left[-z_{\alpha/2} \leqslant \frac{\left[\sqrt{n}\hat{\delta}^* - \sqrt{n}\delta^*\right]}{\left[1 + \frac{(\sqrt{n}\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}} \leqslant z_{\alpha/2}\right] = 1\text{-}\alpha$$

or, $$\Pr\left[\hat{\delta}^* - z_{\alpha/2}\left[\frac{1}{n} + \frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}} \leqslant \delta^* \leqslant \hat{\delta}^* + z_{\alpha/2}\left[\frac{1}{n} + \frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}\right] = 1\text{-}\alpha$$

Thus, an approximate (1-α) level confidence Intervals for δ^* is given by

$$(\delta^*_L, \delta^*_U) = \left\{\hat{\delta}^* - z_{\alpha/2}\left[\frac{1}{n} + \frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}, \ \hat{\delta}^* + \left[\frac{1}{n} + \frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}} z_{\alpha/2}\right\}$$

Then, an approximate (1-α) level confidence Intervals for R is represented by

$$(\Phi(\delta^*_L), \Phi(\delta^*_U)) = \left\{\Phi\left(\hat{\delta}^* - z_{\alpha/2}\left[\frac{1}{n} + \frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}\right), \ \Phi\left(\hat{\delta}^* + z_{\alpha/2}\left[\frac{1}{n} + \frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}\right)\right\} \tag{9}$$

where, $z_{\alpha/2}$ upper critical value for the standard normal distribution

X. Approximate Lower Confidence bound for R

The lower bounds based on approximate results is given by

$$\Pr(\delta^*_{LB} \leqslant \delta^*) = 1-\alpha$$

or, $$\Pr\left(\frac{\left[\delta^*_{LB}-\hat{\delta}^*\right]}{\left[\frac{1}{n}+\frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}} \leqslant \frac{\left[\delta^*-\hat{\delta}^*\right]}{\left[\frac{1}{n}+\frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}}\right) = 1-\alpha$$

or, $$\Pr\left(z \leqslant \frac{\left[\hat{\delta}^*-\delta^*_{LB}\right]}{\left[\frac{1}{n}+\frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}}\right) = 1-\alpha$$

or, $$\delta^*_{LB} = \hat{\delta}^* - z_{1-\alpha}\left[\frac{1}{n}+\frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}$$

So, an approximate (1-α) confidence lower bound for R as

$$\Phi(\delta^*_{LB}) = \Phi\left(\hat{\delta}^* - z_{1-\alpha}\left[\frac{1}{n}+\frac{(\hat{\delta}^*)^2}{2(n-1)}\right]^{\frac{1}{2}}\right) \tag{10}$$

XI. Bootstrap confidence Intervals for R

Now, we consider the confidence intervals based on percentile bootstrap method. Efron suggests [36] the procedure to find out the confidence intervals for a parameter. It works as follows

(1) Draw random sample $\begin{pmatrix}\mathbf{X}_\alpha\\ \mathbf{Y}_\alpha\end{pmatrix}$, $\alpha = 1,2,\dots,n$ from multivariate log-normal distribution, where $\begin{pmatrix}\mathbf{X}\\ \mathbf{Y}\end{pmatrix} \sim \Lambda_{p_1+p_2}(\mu,\Sigma)$.

(2) Take the log-transformation as $\begin{pmatrix}\ln(\mathbf{X}_\alpha)\\ \ln(\mathbf{Y}_\alpha)\end{pmatrix} = \begin{pmatrix}\mathbf{X}^*_\alpha\\ \mathbf{Y}^*_\alpha\end{pmatrix} \sim N_{p_1+p_2}(\mu,\Sigma), \alpha = 1,2,\dots,n.$

(3) Generate bootstrap samples $\begin{pmatrix}\mathbf{X}^B_\alpha\\ \mathbf{Y}^B_\alpha\end{pmatrix}$, $\alpha = 1,2,\dots,n$, by using random sample of $\begin{pmatrix}\mathbf{X}^*_\alpha\\ \mathbf{Y}^*_\alpha\end{pmatrix}$, $\alpha = 1,2,\dots,n$.

(4) Compute the bootstrap estimates of linear dependent vectors $\mathbf{a}'$ and $\mathbf{b}'$ using PC1, say $\mathbf{e}'^*$ and $\mathbf{l}'^*$ respectively. Also, Compute the bootstrap MLE estimates of $\boldsymbol{\mu}_1$, $\boldsymbol{\mu}_2$, $\boldsymbol{\Sigma}_{11}$, $\boldsymbol{\Sigma}_{12}$, $\boldsymbol{\Sigma}_{22}$ by $\overline{\mathbf{x}^*}, \overline{\mathbf{y}^*}, \mathbf{S}^*_{11}, \mathbf{S}^*_{12}, \mathbf{S}^*_{12}$. Using these estimates compute the bootstrap estimate of R, say R^*_B.

(5) Repeat steps 3 and 4, number of boot time B (B sufficiently large, i.e. 1000), thus we obtain the bootstrap distribution of $\{R^*_B\}$.

(6) Estimate $(1-\alpha)$ bootstrap percentile confidence intervals for R from $\{R^*_B\}$ by taking the $\left(\frac{\alpha}{2}\right)$ and $\left(1-\frac{\alpha}{2}\right)$ quantiles as $\left(R^*_{B,\frac{\alpha}{2}}, R^*_{B,\left(1-\frac{\alpha}{2}\right)}\right)$.

or, $(1-\alpha)$ bootstrap percentile lower bound for R as $R^*_{B,\alpha}$.

In order to corrects for narrowness bias of percentile confidence intervals for R, we use the Better Bootstrap Confidence Intervals proposed by Efron [37].

III. Simulation Study

I. Calculation of mean, variance, MSE and MAE of $\hat{\hat{R}}$ and R*

The simulation study we perform here is to compare the behaviours of two estimators $\hat{\hat{R}}$ and R^*

based on estimator of $\mu = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix}_{(p_1 x p_2)}$ and $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}_{(p_1+p_2)x(p_1+p_2)}$. For this purpose, we compute the following measures:

(i) Mean of $\hat{\hat{R}}$ and R^*

(ii) Variance of $\hat{\hat{R}}$ and R^* : $E(\hat{\hat{R}} - R)^2$ and $E(R^* - R)^2$

(iii) Mean square error of $\hat{\hat{R}}$ and R^* : $Var(\hat{\hat{R}})+Bias(\hat{R},R)^2$ and $Var(R^*)+Bias(R^*,R)^2$

(iv) Mean absolute error of $\hat{\hat{R}}$ and R^* : $E(|\hat{\hat{R}} - R|)$ and $E(|R^* - R|)$

It is difficult to obtain the analytical form of the above expressions for different values of 'R'. So, we obtain these by using simulation study. Hence, we generate random samples of size n from $\begin{pmatrix} \mathbf{x} \\ \mathbf{y} \end{pmatrix} \sim \Lambda_{p1+p2}(\mu, \Sigma)$. Then take the log-transformation of whole random sample of size n. For each of sample drawn of size n, we compute the above measures by taking 500 replications each time.

For this purpose, here, R programming language is used.

$$\text{Suppose,} \quad \begin{pmatrix} x_1 \\ x_2 \\ y_1 \\ y_2 \\ y_3 \end{pmatrix} \sim \Lambda_5(\mu, \Sigma) \text{, then} \begin{pmatrix} \ln(x_1) \\ \ln(x_2) \\ \ln(y_1) \\ \ln(y_2) \\ \ln(y_3) \end{pmatrix} = \begin{pmatrix} x_1^* \\ x_2^* \\ y_1^* \\ y_2^* \\ y_3^* \end{pmatrix} \sim N_5(\mu, \Sigma)$$

$$\text{where ,} \quad \mu' = (2,4,2,1,2) \;; \Sigma = \begin{pmatrix} 3.61 & 2.23 & -0.10 & 0.16 & 2.32 \\ 2.23 & 4.74 & 3.32 & -0.69 & 1.76 \\ -0.10 & 3.32 & 5.68 & -2.34 & -1.23 \\ 0.16 & -0.69 & -2.34 & 3.05 & 1.53 \\ 2.32 & 1.76 & -1.23 & 1.53 & 4.45 \end{pmatrix}$$

Therefore, the estimated values of vectors based on first principal component are $\mathbf{a}'$ = (0.614, 0.789) and $\mathbf{b}'$= (0.737, -0.491, -0.465). Then the exact value of stress strength reliability R= $Pr(\mathbf{a}'\mathbf{x}^* > \mathbf{b}'\mathbf{y}^*)$ is 0.8861. We take the sample size of n up to 200 in order to achieved exact value for the reliability. Calculated mean, variance, MSE and MAE of $\hat{\hat{R}}$ and R* based on 500 repetitions are reported in Table 1. It can be observed that Variance of $\hat{\hat{R}}$ is lesser than the Variance of R* in each sample size. Also, it is noted that the MSE's and MAE's of $\hat{\hat{R}}$ is lesser than MSE's and MAE's of R*. However, the sample mean of $\hat{\hat{R}}$ is less than the R* in each case. But, $\hat{\hat{R}}$ and R* are under-estimates the true value of R, when sample sizes are small. It is also interesting to observe that, the Variance, MSE and MAE of $\hat{\hat{R}}$ and R* are reduces as the sample size increases: when n=200 and they almost achieved the true value of R.

Table 1: *Sample Mean, Variance, MSE and MAE of $\hat{\hat{R}}$ and R^**

Sample	Sample Mean		Variance		MSE		MAE	
Size	$\hat{\hat{R}}$	R^*	$\hat{\hat{R}}$	R^*	$\hat{\hat{R}}$	R^*	$\hat{\hat{R}}$	R^*
10	0.6723	0.6751	0.1054	0.1097	0.1509	0.1540	0.2475	0.2508
20	0.7302	0.7322	0.0949	0.0969	0.1190	0.1203	0.1903	0.1924
30	0.7550	0.7566	0.0823	0.0836	0.0993	0.1002	0.1524	0.1633
40	0.7935	0.7951	0.0591	0.0598	0.0675	0.0680	0.1103	0.1173
50	0.8050	0.8063	0.0544	0.0549	0.0609	0.0612	0.0970	0.1046
60	0.8126	0.8138	0.0481	0.0485	0.0534	0.0536	0.9097	0.0927
70	0.8382	0.8392	0.0326	0.0328	0.0348	0.0349	0.0667	0.0681

80	0.8449	0.8459	0.0305	0.0307	0.0322	0.0323	0.0630	0.0631
90	0.8487	0.8496	0.0264	0.0265	0.0277	0.0278	0.0553	0.0558
100	0.8560	0.8568	0.0211	0.0212	0.0219	0.0220	0.0476	0.0486
200	0.8810	0.8815	0.0037	0.0038	0.0038	0.0038	0.0182	0.0183

II. Relative measure between $\hat{R}$ and R^*

It' s difficult to get the exact values of relative measure of deviation between $\hat{R}$ and R^*, i.e., $R(n, \delta^*)$. So, we compute these measure values numerically using R - programming. Using different choices of parameters in Table 2 and variance-covariance matrix, results of these two measures and ratio are reported in Table 3. The results shows that the overall output of MVUE and MLE of R are not too distant and the values for these differences are show in columns U(.), M(.) and R(.) of this tables. From this tables, it is seen that empirical values of the parameters and the performance of MVUE of R is better than MLE and Figure 1 shows that, MVUE estimator of R is better than the other estimators, i.e. R^* also L_1 distance and graphical impression show this.

Table 2: *Parameters for the simulation study*

Parameters and values													
Scenario	μ_1			μ_2			Scenario	μ_1			μ_2		
	$\mu_{p_{11}}$	$\mu_{p_{12}}$	$\mu_{p_{13}}$	$\mu_{p_{21}}$	$\mu_{p_{22}}$	$\mu_{p_{23}}$		$\mu_{p_{11}}$	$\mu_{p_{12}}$	$\mu_{p_{13}}$	$\mu_{p_{21}}$	$\mu_{p_{22}}$	$\mu_{p_{23}}$
1	1.639	1.477	1.036	2.802	2.114	2.197	48	1.161	2.647	2.508	1.114	2.918	2.077
2	1.956	2.427	1.019	2.612	2.908	2.286	49	2.112	2.288	2.621	2.439	2.949	1.363
3	1.747	1.103	1.545	1.382	2.837	2.796	50	1.599	2.294	2.771	1.81	2.973	1.726
4	1.052	1.972	1.014	1.938	2.271	2.616	51	2.002	1.414	1.431	1.751	1.288	1.724
5	1.204	1.233	1.927	2.122	2.581	2.732	52	2.728	1.52	2.06	1.633	2.111	2.192
6	1.765	1.765	1.233	1.723	2.927	2.312	53	2.86	1.224	1.351	2.287	1.528	1.1
7	2.309	1.705	1.445	2.915	2.683	2.203	54	1.484	1.694	2.933	2.965	2.23	1.067
8	1.109	1.096	1.503	1.893	2.91	1.011	55	1.954	1.083	2.862	2.939	1.038	2.878
9	2.581	1.069	1.407	2.761	2.031	2.805	56	1.527	1.242	2.898	1.421	2.408	1.707
10	1.668	1.757	1.574	2.553	2.331	2.603	57	2.381	1.492	2.392	2.253	1.725	2.298
11	1.626	1.027	2.062	2.543	2.185	2.471	58	1.412	1.049	2.362	1.095	1.226	2.716
12	1.363	2.054	1.407	2.969	1.907	2.101	59	2.242	2.174	2.343	2.327	1.556	2.921
13	2.354	1.169	2.77	2.79	2.86	2.754	60	2.301	2.947	2.319	2.915	2.65	1.167
14	1.455	2.164	1.552	2.953	2.34	1.635	61	1.615	2.318	1.856	1.358	1.825	2.031
15	2.499	2.846	1.217	2.784	2.31	3	62	1.465	1.733	1.341	1.327	1.119	1.583
16	1.163	2.517	1.353	2.661	1.724	2.465	63	2.88	1.068	1.288	1.136	1.223	2.074
17	1.321	1.85	2.579	2.177	2.764	2.424	64	2.93	2.501	2.164	2.105	2.625	1.769
18	2.192	1.462	1.447	1.047	2.458	2.48	65	2.283	2.939	1.882	2.732	1.257	2.994
19	1.787	1.767	1.045	1.44	2.134	1.938	66	2.48	1.709	1.753	2.115	1.713	1.262
20	1.038	1.531	1.905	1.026	2.66	1.885	67	2.848	1.375	2.943	2.55	2.689	1.076
21	2.296	1.292	1.875	2.673	2.658	1.108	68	2.291	1.737	1.526	1.188	1.378	2.215
22	2.321	1.997	1.648	1.416	2.765	2.606	69	1.633	1.614	2.93	2.544	1.64	2.054
23	1.149	1.965	1.995	2.545	2.5	1.396	70	2.281	2.899	1.503	2.123	2.186	1.006
24	1.388	2.776	1.139	2.336	2.515	1.18	71	1.799	2.802	1.392	1.163	1.72	1.93
25	1.985	2.072	1.754	2.5	2.921	1.034	72	2.621	1.027	2.746	1.542	2.282	1.567
26	2.776	1.44	1.543	1.804	2.287	2.518	73	1.084	2.624	2.256	1.19	1.819	2.237
27	2.925	1.498	1.107	1.556	2.407	1.982	74	2.978	1.269	2.428	1.349	1.725	2.747
28	2.874	2.704	2.074	2.64	2.985	2.804	75	2.291	2.088	1.718	1.525	1.883	1.099
29	2.817	2.141	1.408	2.425	2.45	2.083	76	1.269	2.765	2.325	1.596	1.928	1.799
30	1.989	2.875	1.918	1.898	2.757	2.941	77	1.718	1.935	2.987	1.314	2.037	2.471
31	1.718	1.69	1.809	2.842	2.077	1.35	78	1.629	2.716	2.517	2.634	1.104	2.737
32	1.138	2.103	2.945	2.398	2.475	2.96	79	2.451	2.782	2.616	1.883	2.402	1.915
33	2.037	2.076	1.99	1.936	2.375	2.637	80	2.292	1.616	2.521	2.446	1.182	1.848
34	2.274	1.688	1.447	1.729	2.127	2.109	81	1.224	2.875	2.904	1.137	1.989	2.739
35	2.58	1.507	1.01	2.67	1.789	1.086	82	2.357	1.374	2.61	2.75	1.099	1.501
36	2.586	2.478	2.217	2.404	2.719	2.889	83	2.386	2.173	1.502	1.14	1.121	2.063
37	1.633	1.338	2.475	2.433	2.152	2.109	84	2.754	2.519	2.483	1.889	2.52	1.055
38	1.123	2.964	2.605	1.937	2.841	2.832	85	2.524	1.71	2.428	1.518	1.982	1.141

39	1.133	2.364	2.303	1.328	2.643	2.561	86	1.882	2.371	2.43	1.926	1.061	2.558
40	1.632	1.531	2.555	1.319	2.92	1.988	87	2.011	1.434	2.784	1.274	1.017	2.76
41	2.208	2.555	1.623	2.826	2.576	1.053	88	2.413	1.888	2.558	1.201	1.479	2.338
42	2.316	2.249	1.692	2.231	2.229	1.963	89	2.772	1.542	2.714	2.086	1.157	2.04
43	1.616	2.362	1.274	1.946	1.757	1.733	90	2.286	2.884	2.07	1.48	1.605	1.513
44	1.626	1.16	2.207	2.525	1.291	2.176	91	2.558	2.333	2.824	1.094	1.791	2.499
45	2.833	1.197	2.598	2.295	2.687	1.706	92	2.699	1.638	2.683	1.118	1.133	2.746
46	2.104	2.872	1.278	1.494	2.588	1.608	93	2.975	2.718	2.373	1.044	2.075	1.988
47	2.109	2.691	1.471	1.83	2.115	2.243	94	2.999	2.917	2.907	1.395	2.85	1.057

and variance-covariance matrix as $\Sigma = \begin{pmatrix} 3.199 & 0.686 & 0.610 & 0.387 & 0.480 & 0.978 \\ 0.686 & 3.771 & 0.536 & 0.869 & -0.441 & -0.845 \\ 0.610 & 0.536 & 4.463 & 0.161 & 0.853 & -0.403 \\ 0.387 & 0.869 & 0.161 & 2.690 & 0.911 & 0.804 \\ 0.480 & -0.441 & 0.853 & 0.911 & 4.169 & 0.897 \\ 0.978 & -0.845 & -0.403 & 0.804 & 0.897 & 2.706 \end{pmatrix}$

Table 3: *Performance of point estimators:* $\sqrt{n}\delta^*$ *and 100** $\{U(n,\delta^*), M(n,\delta^*), R(n,\delta^*)\}$

Non-negative values of δ^*					Negative values of δ^*				
Scenario	$\sqrt{n}\delta^*$	$U(n,\delta^*)$	$M(n,\delta^*)$	$R(n,\delta^*)$	Scenario	$\sqrt{n}\delta^*$	$U(n,\delta^*)$	$M(n,\delta^*)$	$R(n,\delta^*)$
1	2.43	0.876	70.393	1.245	48	-0.038	0.277	50.217	0.552
2	2.415	0.884	70.265	1.258	49	-0.052	0.277	50.299	0.552
3	2.389	0.898	70.045	1.282	50	-0.054	0.277	50.309	0.551
4	2.345	0.923	69.673	1.324	51	-0.155	0.538	50.892	1.056
5	2.329	0.932	69.535	1.341	52	-0.198	0.684	51.15	1.338
6	2.251	0.982	68.867	1.426	53	-0.224	0.768	51.303	1.498
7	2.131	1.072	67.812	1.58	54	-0.273	0.925	51.601	1.793
8	2.102	1.095	67.555	1.621	55	-0.284	0.957	51.665	1.853
9	2	1.187	66.628	1.781	56	-0.297	0.998	51.747	1.929
10	1.95	1.236	66.163	1.868	57	-0.326	1.086	51.93	2.092
11	1.705	1.503	63.838	2.354	58	-0.43	1.373	52.601	2.611
12	1.55	1.684	62.329	2.701	59	-0.446	1.412	52.704	2.68
13	1.485	1.757	61.695	2.848	60	-0.492	1.522	53.016	2.871
14	1.447	1.799	61.323	2.934	61	-0.496	1.532	53.045	2.888
15	1.406	1.843	60.922	3.025	62	-0.524	1.594	53.24	2.994
16	1.207	2.017	58.998	3.419	63	-0.525	1.596	53.245	2.997
17	1.108	2.07	58.061	3.565	64	-0.536	1.62	53.324	3.037
18	1.088	2.077	57.877	3.589	65	-0.565	1.68	53.536	3.138
19	1.088	2.077	57.876	3.589	66	-0.596	1.738	53.76	3.233
20	1.072	2.082	57.728	3.606	67	-0.609	1.761	53.853	3.27
21	1.069	2.083	57.699	3.609	68	-0.624	1.786	53.963	3.311
22	1.062	2.084	57.639	3.616	69	-0.635	1.805	54.048	3.34
23	1.054	2.086	57.563	3.624	70	-0.682	1.876	54.403	3.449
24	1.013	2.091	57.193	3.657	71	-0.761	1.973	55.027	3.586
25	0.966	2.09	56.764	3.681	72	-0.801	2.011	55.351	3.633
26	0.947	2.087	56.597	3.687	73	-0.833	2.036	55.622	3.66
27	0.942	2.086	56.553	3.688	74	-0.864	2.055	55.882	3.678
28	0.887	2.067	56.076	3.686	75	-1.009	2.092	57.153	3.66
29	0.863	2.055	55.875	3.678	76	-1.023	2.091	57.275	3.65
30	0.822	2.028	55.528	3.652	77	-1.05	2.087	57.524	3.628

31	0.782	1.994	55.197	3.612	78	-1.124	2.063	58.211	3.545
32	0.764	1.976	55.049	3.59	79	-1.276	1.966	59.653	3.295
33	0.686	1.883	54.438	3.459	80	-1.317	1.93	60.055	3.213
34	0.679	1.873	54.382	3.444	81	-1.353	1.897	60.399	3.14
35	0.671	1.861	54.318	3.425	82	-1.402	1.847	60.884	3.033
36	0.66	1.844	54.235	3.401	83	-1.414	1.834	61.002	3.007
37	0.656	1.838	54.206	3.392	84	-1.526	1.711	62.096	2.755
38	0.566	1.68	53.538	3.139	85	-1.536	1.7	62.192	2.733
39	0.549	1.647	53.419	3.083	86	-1.547	1.686	62.304	2.707
40	0.531	1.609	53.288	3.019	87	-1.67	1.543	63.503	2.43
41	0.43	1.373	52.6	2.61	88	-1.776	1.422	64.519	2.204
42	0.306	1.026	51.804	1.98	89	-1.916	1.27	65.846	1.929
43	0.238	0.814	51.389	1.585	90	-2.092	1.104	67.462	1.637
44	0.183	0.634	51.06	1.241	91	-2.097	1.1	67.507	1.629
45	0.164	0.568	50.945	1.115	92	-2.106	1.092	67.593	1.615
46	0.159	0.554	50.92	1.088	93	-2.157	1.05	68.049	1.544
47	0.14	0.49	50.809	0.964	94	-2.384	0.9	70.008	1.286

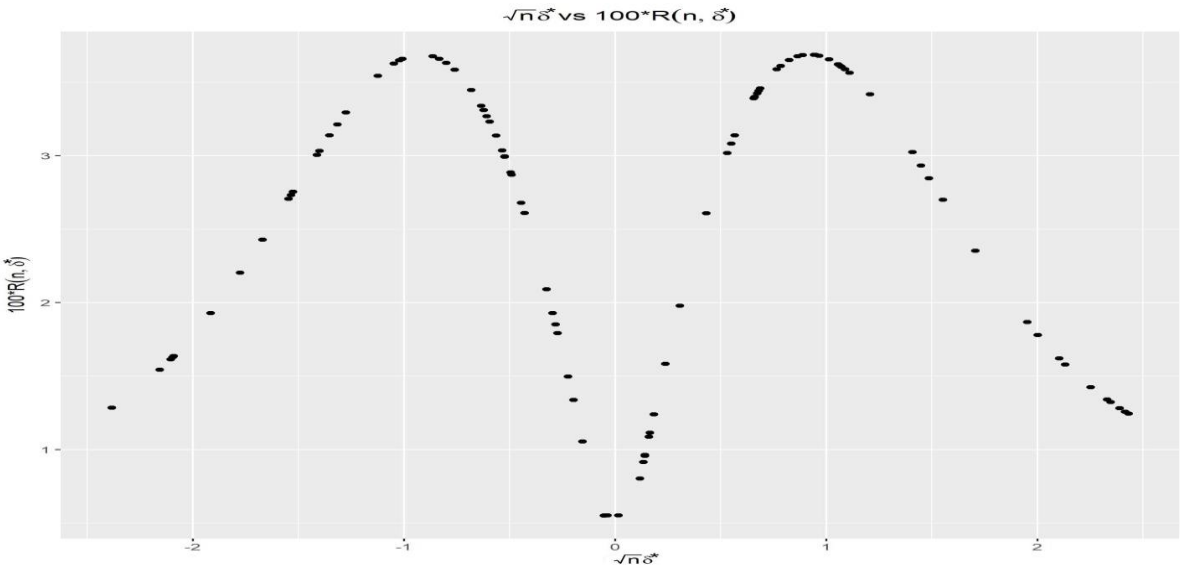

Figure 1: $\sqrt{n}\delta^*$ vs $100 * R(n, \delta^*)$

III. Calculation of $Var(\hat{R})$ and $MSE(R^*)$

From Figure 2 and 3 it is observe that the, values of $Var(\hat{R})$ and $MSE(R^*)$ are almost close to zero of δ^*. Values of $Var(\hat{R})$ are less as compared to other values of $MSE(R^*)$. Thus, the performance of MVUE of R is better than MLE. In addition, for the given data in Table 7, we calculate that the above measure by principal component analysis as $\sqrt{n}\delta^*$ = 5.486, $U(n, \delta^*)$ = 0.13, $M(n, \delta^*)$= 71.027, $R(n, \delta^*)$= 0.182, $Var(\hat{R})$ = 0.00138 and $MSE(R^*)$ = 0.00139.

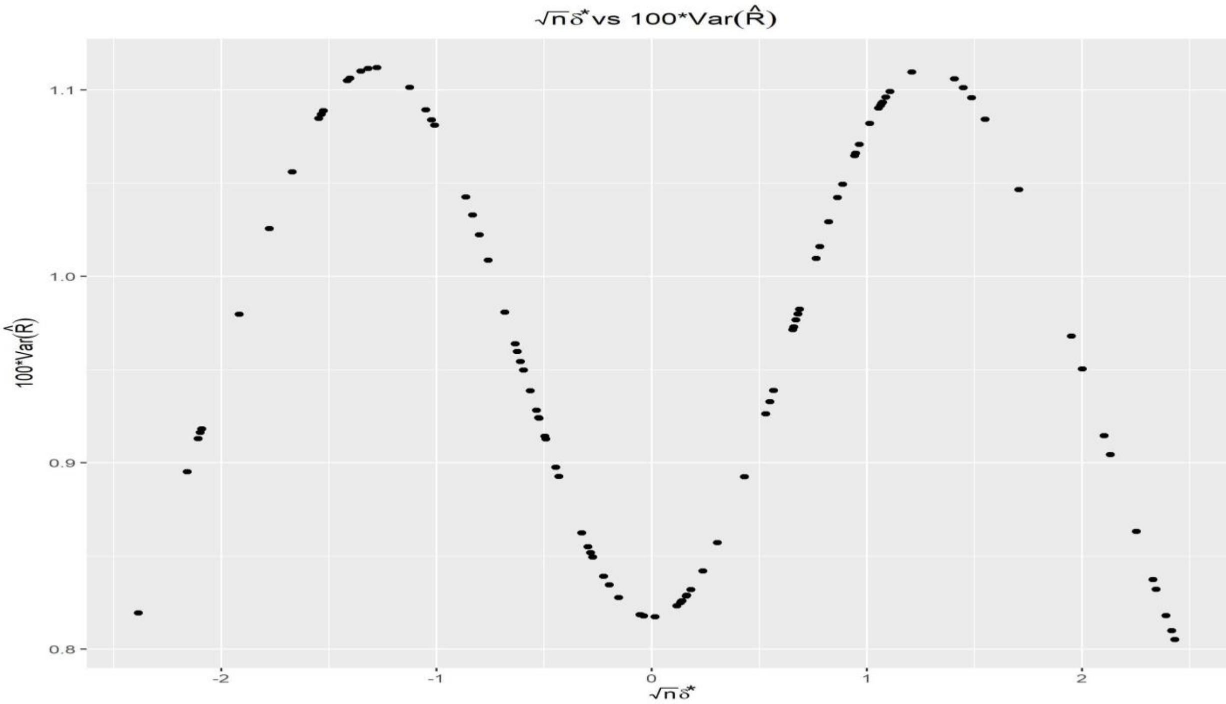

Figure 2: $\sqrt{n}\delta^{*}$ vs $100 * Var(\hat{R})$

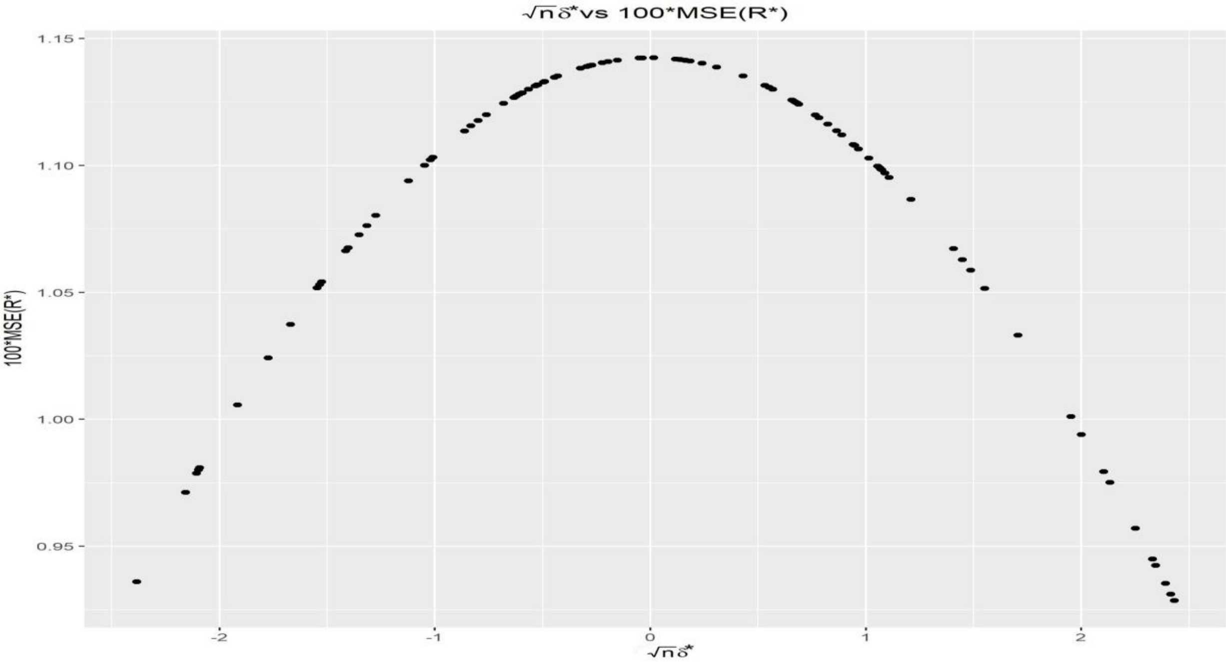

Figure 3: $\sqrt{n}\delta^{*}$ vs $100 * MSE(R^{*})$

IV. Confidence Intervals for R

From the section of VII to XI of methods, we present simulation study to investigate the statistical properties of the interval estimators using the given matrix in section I of simulation study. The simulation study define as follows

(1) Draw the random samples of size n from $\binom{\mathbf{X}}{\mathbf{Y}} \sim \Lambda_{p1+p2}(\mu, \Sigma)$. Then Take the log-transformation as $\binom{\ln(\mathbf{X})}{\ln(\mathbf{Y})} = \binom{\mathbf{X}^*}{\mathbf{Y}^*} \sim N_{p1+p2}(\mu, \Sigma)$. For each of sample drown of size n, considered different sample sizes (n=50, 100, 150,….etc). We compute the above measures by taking 500 replications each time.
(2) Estimate MLE estimate of R using PC1 for different sample size and a Confidence Intervals (two-sided and lower bound) for R.
(3) Compute exact, approximate and bootstrap confidence intervals using step 2, where number of boot time B=1000.

The results of the simulation study are recorded in Table 4-6. Figure 4-6, represent the exact, approximate and bootstrap confidence belt at 90%, 95% and 99% levels. It has been observed that for a small sample size, the estimate of R is getting high and also confidence intervals. The results get better as the sample sizes increase and the reliability R gets closer to true value. The same phenomenon is observed for the exact, approximate and bootstrap confidence intervals. The overall band of exact and approximate confidence intervals is almost same, whereas bootstrap confidence intervals give the large confidence band for small sample size. But, exact, approximate confidence intervals and Bootstrap confidence intervals all are almost same for large sample size at 90%, 95% and 99% levels. All most the same variation found in confidence belt of exact, approximate CIs, but quite irregular variation in bootstrap CIs shows in Figure 4-6.

Table 4: *Exact Confidence Intervals*

Sample size	R*	90%			95%			99%		
		L	U	LB	L	U	LB	L	U	LB
50	0.9092	0.8436	0.9507	0.8598	0.8288	0.9568	0.8436	0.7973	0.9669	0.8104
100	0.9043	0.8599	0.9368	0.8705	0.8503	0.9420	0.8599	0.8302	0.9513	0.8385
150	0.8984	0.8621	0.9267	0.8707	0.8544	0.9314	0.8621	0.8386	0.9401	0.8452
200	0.8953	0.8640	0.9207	0.8713	0.8575	0.9251	0.8640	0.8441	0.9330	0.8496
250	0.8925	0.8644	0.9158	0.8709	0.8586	0.9198	0.8644	0.8468	0.9273	0.8516
300	0.8920	0.8665	0.9135	0.8724	0.8613	0.9173	0.8665	0.8506	0.9243	0.8550
350	0.8913	0.8678	0.9114	0.8732	0.8630	0.9150	0.8678	0.8533	0.9216	0.8573
400	0.8897	0.8677	0.9088	0.8728	0.8632	0.9122	0.8677	0.8542	0.9186	0.8579
450	0.8895	0.8688	0.9076	0.8736	0.8646	0.9108	0.8688	0.8561	0.9169	0.8596
500	0.8893	0.8697	0.9065	0.8742	0.8658	0.9096	0.8697	0.8578	0.9155	0.8611
550	0.8848	0.8659	0.9017	0.8702	0.8620	0.9047	0.8659	0.8544	0.9104	0.8575
600	0.8810	0.8626	0.8974	0.8668	0.8589	0.9004	0.8626	0.8515	0.9061	0.8545

Table 5: *Approximate Confidence Intervals*

Sample size	90%			95%			99%		
	L	U	LB	L	U	LB	L	U	LB
50	0.8447	0.9512	0.8611	0.8296	0.9572	0.8447	0.7973	0.9670	0.8108
100	0.8604	0.9371	0.8711	0.8507	0.9423	0.8604	0.8303	0.9514	0.8387
150	0.8625	0.9269	0.8711	0.8547	0.9316	0.8625	0.8387	0.9401	0.8453
200	0.8643	0.9209	0.8716	0.8577	0.9252	0.8643	0.8441	0.9331	0.8497
250	0.8646	0.9159	0.8711	0.8588	0.9199	0.8646	0.8468	0.9274	0.8517
300	0.8667	0.9136	0.8726	0.8614	0.9174	0.8667	0.8507	0.9243	0.8551
350	0.8679	0.9115	0.8734	0.8631	0.9151	0.8679	0.8533	0.9217	0.8573
400	0.8679	0.9089	0.8729	0.8633	0.9123	0.8679	0.8542	0.9186	0.8580
450	0.8689	0.9077	0.8737	0.8647	0.9109	0.8689	0.8562	0.9169	0.8597
500	0.8698	0.9066	0.8743	0.8659	0.9097	0.8698	0.8578	0.9155	0.8611

550	0.8660	0.9017	0.8703	0.8621	0.9048	0.8660	0.8544	0.9105	0.8575
600	0.8627	0.8975	0.8669	0.8590	0.9005	0.8627	0.8515	0.9061	0.8546

Table 6: *Bootstrap Confidence Intervals*

Sample	90%			95%			99%		
size	L	U	LB	L	U	LB	L	U	LB
50	0.8662	0.9479	0.8752	0.8583	0.9558	0.8662	0.8430	0.9711	0.8492
100	0.8605	0.9474	0.8701	0.8522	0.9557	0.8605	0.8359	0.9720	0.8425
150	0.8686	0.9274	0.8751	0.8630	0.9330	0.8686	0.8520	0.9440	0.8564
200	0.8657	0.9256	0.8723	0.8599	0.9313	0.8657	0.8487	0.9425	0.8533
250	0.8677	0.9191	0.8734	0.8628	0.9240	0.8677	0.8531	0.9336	0.8570
300	0.8675	0.9165	0.8729	0.8628	0.9212	0.8675	0.8536	0.9304	0.8573
350	0.8696	0.9146	0.8746	0.8653	0.9189	0.8696	0.8569	0.9273	0.8603
400	0.8713	0.9087	0.8754	0.8677	0.9122	0.8713	0.8607	0.9192	0.8636
450	0.8701	0.9093	0.8744	0.8663	0.9130	0.8701	0.8590	0.9204	0.8620
500	0.8711	0.9060	0.8750	0.8678	0.9093	0.8711	0.8612	0.9159	0.8639
550	0.8660	0.9045	0.8703	0.8623	0.9081	0.8660	0.8551	0.9153	0.8580
600	0.8653	0.8968	0.8688	0.8623	0.8998	0.8653	0.8564	0.9057	0.8588

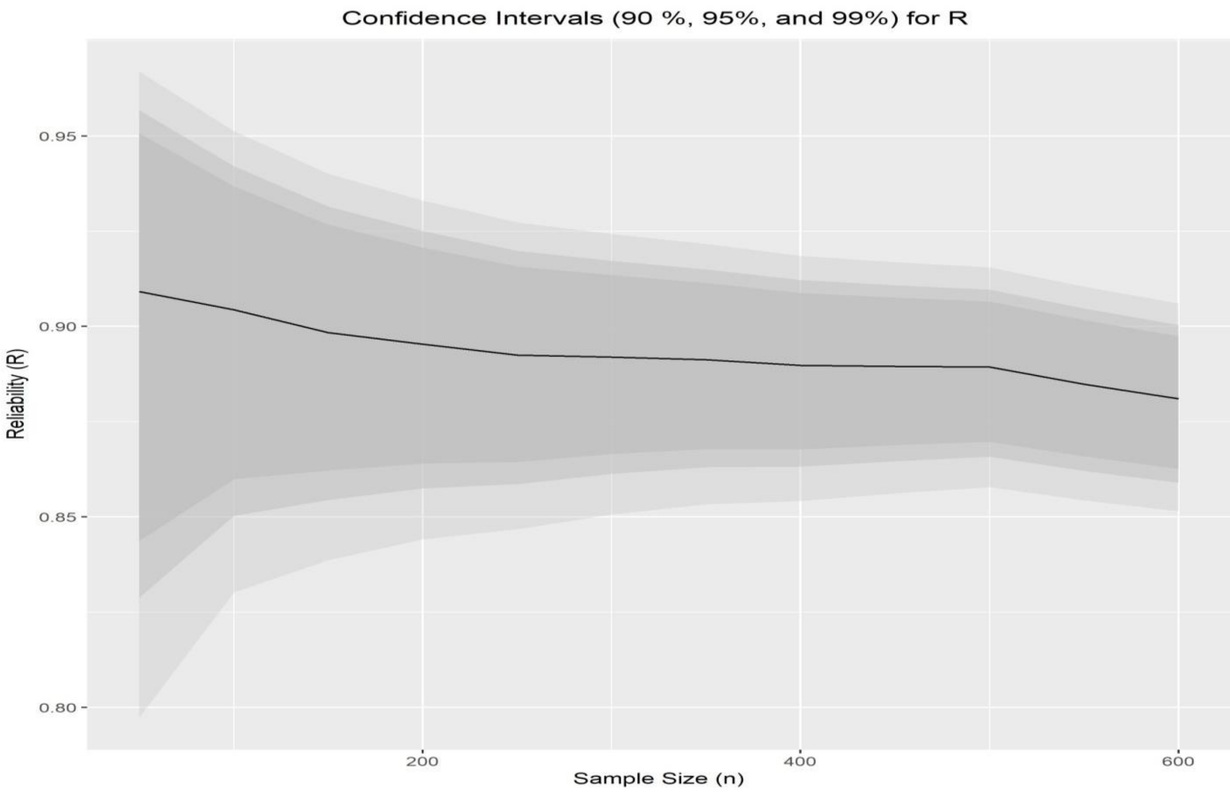

Figure 4: *Exact Confidence Intervals*

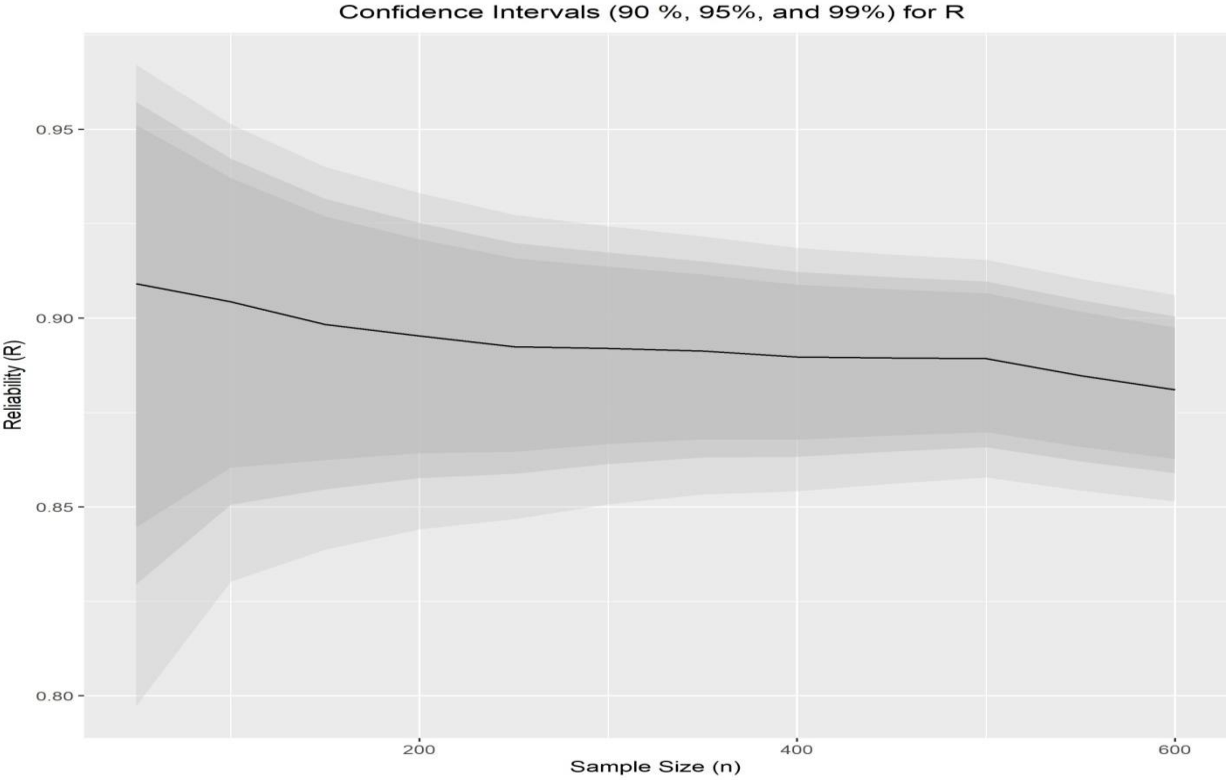

Figure 5: *Approximate Confidence Intervals*

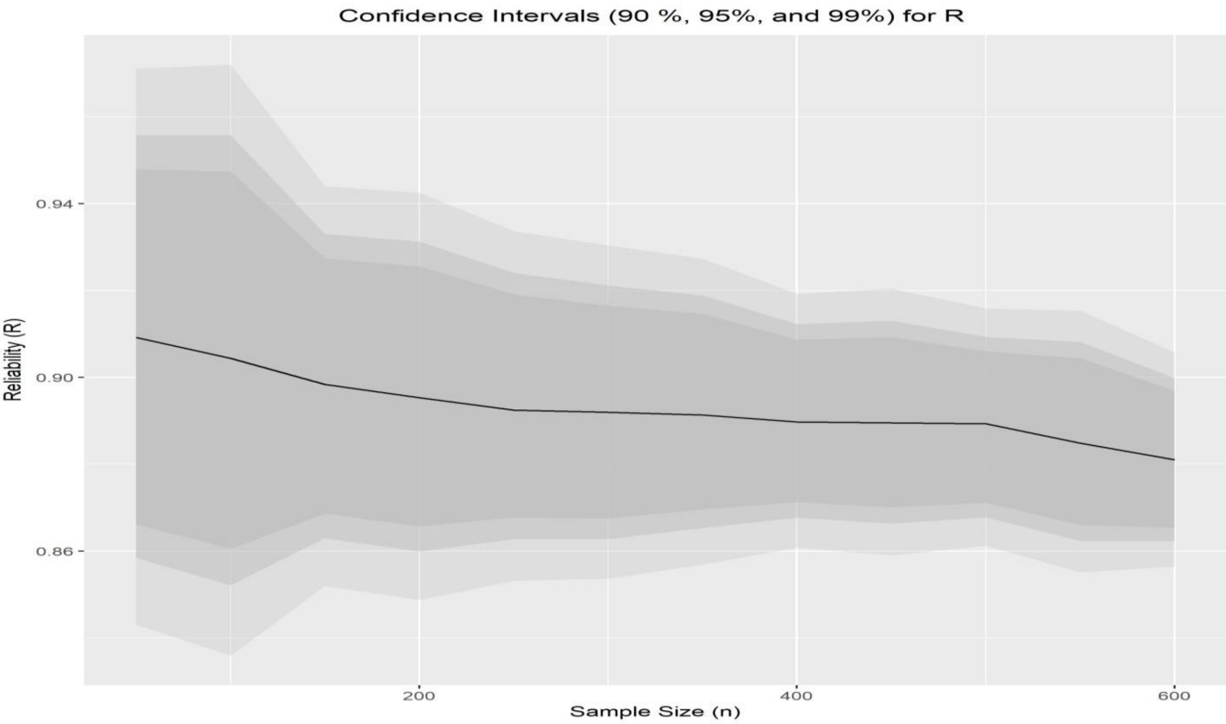

Figure 6: *Bootstrap Confidence Intervals*

IV. Application on Real Data Set

I. Estimation of $\hat{\hat{R}}$ and R^*

In this section, applies the methods developed in the previous sections, using the Air Quality Index Data taken from "https://www.kaggle.com/rohanrao/air-quality-data-in-india?select=station_day.csv". The dataset contains air quality data and AQI (Air Quality Index) at hourly and daily level of various stations across multiple cities in India. Consider the seven pollutants namely particulate matter PM_{10}, $PM_{2.5}$, nitrogen dioxide (NO_2), ammonia (NH_3), carbon monoxide (CO), Sulphur dioxide (SO_2) and Ozone (O_3) act as major parameters in deriving the AQI of an area. Data were recorded from January-2015 to June-2020 (five years and six months) representing the longest freely available recordings of on field deployed air quality. We selected only two cites as Kolkata and Patna because distance between the two cites is much less near about 500 KM and Kolkata is a metropolitan city. Being a Kolkata metropolitan city, its air quality is higher than Patna city. We estimate the stress strength reliability of air quality between the two cities and compare them before and during lockdown period for COVID-19. We have taken the data from Dec-19 to March-20 for before lockdown and data from 25-03-2020 to 30-06-2020 for lockdown period. Table 7 represents the sample data set of air quality from 25-03-2020 to 30-06-2020.

Table 7: *Data set of Air Quality Index (AQI) of Kolkata and Patna*

City	Date	$PM_{2.5}$	PM_{10}	NO_2	NH_3	CO	SO_2	O_3
	25-03-2020	53.01	77.59	15.94	19.44	0.57	12.52	44.48
	26-03-2020	37.02	60.49	14.36	19.23	0.52	6.8	47.17
	27-03-2020	48.1	83.63	14.86	18.61	0.55	11.62	60.57
	28-03-2020	46.57	77.11	15.71	18.62	0.57	11.46	58.18
	29-03-2020	40.85	68.04	14.53	18.28	0.54	9.52	64.88
Kolkata	-	-	-	-	-	-	-	-
	26-06-2020	9.59	28.44	7.91	10.35	0.34	8.2	22.66
	27-06-2020	7.89	24.73	9.25	8.39	0.35	5.73	23.14
	28-06-2020	10.58	25.56	12.63	6.92	0.38	5.92	27.76
	29-06-2020	14.53	32.4	15.98	7.59	0.45	7.01	30.64
	30-06-2020	14	35.85	12.29	9.64	0.37	5.66	24.59
	25-03-2020	73.42	121.38	47.99	14.13	1.74	13.57	21.77
	26-03-2020	58.01	118.17	46.63	14.24	1.71	13.26	20.18
	27-03-2020	58.25	129.36	45.25	13.88	1.68	12.52	20.63
	28-03-2020	31.87	87.38	43.29	12.38	1.49	11.73	15.88
	29-03-2020	29.1	81.98	50.82	13.37	0.62	11.6	11.82
Patna	-	-	-	-	-	-	-	-
	26-06-2020	23.16	39.8	25.68	3.13	0.79	2.92	19.42
	27-06-2020	17.71	63.73	23.01	1.91	0.87	3.63	23.39
	28-06-2020	19.27	57.42	18.13	2.05	0.72	3.92	17.37
	29-06-2020	17.24	42.83	20.51	2.26	0.88	3.6	17.5
	30-06-2020	29.76	60.68	27.5	1.59	0.83	3.91	21.7

Using the log-transformation of given data sets, the proportions of dispersions of air quality in Kolkata as $\boldsymbol{x}^* = \{\ln(PM_{2.5}), \ln(PM_{10}), \ln(NO_2), \ln(NH_3), \ln(CO), \ln(SO_2), \ln(O_3)\}$ and Patna as $\boldsymbol{y}^* = \{\ln(PM_{2.5}), \ln(PM_{10}), \ln(NO_2), \ln(NH_3), \ln(CO), \ln(SO_2), \ln(O_3)\}$ are 0.8361 and 0.6170 respectively based on their largest eigen values. Then $\hat{\hat{R}} = 0.9724$ based on MVUE and $R^* = 0.9726$ based on

MLE before lockdown. Similarly, during the lockdown, proportion of variance of air quality in Kolkata and Patna for first principal component are 0.7631 and 0.8970 respectively, then the similar things of R are $\hat{\hat{R}} = 0.7089$ and $R^* = 0.7102$. Thus, it is shown that the air quality in the city of Kolkata worse than Patna based on the air quality parameters at before lockdown and during lockdown period. But the air quality has been increased among of these two cites during lockdown.

II. Confidence Intervals

We apply the above methods to find out the exact, approximate and bootstrap confidence intervals, using the sample data set reported in Table 7. The results on air quality data set shows in Table 8 that, the exact and approximate CIs are almost same band, but confidence band of bootstrap CIs is bigger than these.

Table 8 Confidence Intervals for tests of the air quality data

Confidence	90%			95%			99%		
Intervals	L	U	LB	L	U	LB	L	U	LB
Exact	0.6459	0.7677	0.6604	0.6331	0.7781	0.6459	0.6078	0.7975	0.6181
Approx.	0.6464	0.7681	0.6609	0.6336	0.7784	0.6464	0.6082	0.7978	0.6185
Bootstrap	0.4719	0.9155	0.5209	0.4294	0.9580	0.4719	0.3464	0.9899	0.3800

V. Conclusions

In this multivariate log-normal setup, estimator of stress-strength model of reliability R is obtained by log-transformation and using the estimator based on MLE of μ_1, μ_2, Σ_{11}, Σ_{12}, Σ_{22}. Simulation studies illustrate that, the variance and MSE of two estimators reduces as the sample size increases and they almost achieved the true value of R. An application to the given real data set is described and shows that the same result as above. So that, the performance of MVUE based estimator of R is better than MLE in this case. In air quality data it is observed that before the lockdown air pollution in Kolkata was higher than Patna. But during the lockdown, air pollution in Kolkata was much reduced. If the lockdown would continue, then the air quality becomes almost same in between Kolkata and Patna.

In addition, from L_1 distance between distribution functions we can judge the improvement of such estimators. Difference in terms of MSE is much less as the values are given after multiplying by 100, though detailed calculations are required for other parametric values. Therefore, we concluded that our estimator performs better even close to zero. The exact confidence intervals are preferable for marginally short band of confidence intervals than the approximate confidence intervals. The performance of bootstrap CIs is slightly worse than the exact and approximate CIs in terms of confidence band for small sample size. But the performance of bootstrap confidence intervals and other methods of CIs are almost same for large sample. Thus, the overall performance of the confidence interval is quite good for exact confidence intervals.

References

[1] Lipson, C., Sheth, N.J. and Disney, R.L. (1967). Reliability Prediction-Mechanical Stress/Strength Interference. *New York: Rome Air Development Center, Griffiss Air Force Base, Report No. RADC-TR-66-710*.

[2] Feng, C., Wang, H., Lu, N., Chen, T., He, Hua., Lu, Ying. and M.Tu, X. (2014). Log-transformation and its implications for data analysis. *Shanghai Arch Psychiatry*, 26(2):105–109.

[3] Gibrat, R. (1930). Une loi des répartitions économiques: l'effet proportionnel. *Bulletin de la Statistique Générale de la France*, 1929-30:469-514.

[4] Kolmogoroff, A.N. (1941). Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchenbei Zerstückelung. *C. R. (Doklady) Acad. Sci. URSS (N. S.)*, 31:99–101.

[5] Schijve, J. (2005). Statistical distribution functions and fatigue of structures. *Int. J. Fatigue*, 27:1031–1039.

[6] Lu, H., He, Y. and Zhang, Y. (2014). Reliability-Based Robust Design of Mechanical Components with Correlated Failure Modes Based on Moment method. *Advances in Mechanical Engineering*, 1-17.

[7] Paik, J.K. and Frieze, P.A. (2001). Ship Structural Safety and Reliability. *Prog. Struct. Engng Mater*, 3(2):198– 210.

[8] Cohn, S.E. (1997). An introduction to estimation theory. *Journal of the Meteorological Society of Japan*, 75(1B):257– 288.

[9] Wang, Z. and Sassen, K. (2002). Cirrus cloud microphysics property retrieval using lidar and radar measurements. Part II: Midlatitude cirrus microphysical and radiative properties. *Journal of Atmospheric Sciences*, 59(14):2291–2302.

[10] Sengupta, M., Clothiaux, E.E. and Ackerman, T.P. (2004). Climatology of warm boundary layer clouds at the ARM SGP site and their comparison to models. *Journal of Climate*,17(24):4760–4782.

[11] Mielke, P.W.Jr., Williams, J.S. and Wu, S.C. (1977). Covariance analysis techniques based upon bivariate log-Normal distribution with weather modification application. *Journal of Applied Meteorology*, 16(2):183–187.

[12] Gupta, R.C. and Li, X. (2006). Statistical inference for the common mean of two log-normal distributions and some applications in reliability. Comput. *Statist. Data Anal*, 50(11):3141–3164.

[13] Zhou, X.H., Gao, S.J. and Hui, S.L. (1997). Methods for comparing means of two independent log-normal samples. *Biometrics*, 53(3):1129–1135.

[14] Zhou, X.H., Li, C., Gao, S. and Tierney, W.M. (2001). Methods for testing equality of means of health care costs in a paired design study. *Stat. Med*, 20(11):1703–1720.

[15] Schneider, T. and Holst, E. (1983). Man-made mineral fibre size distributions utilizing unbiased and fibre length biased counting methods and bivariate log-normal distribution. *J. Aerosol Sci.*, 14(2):139–146.

[16] Cheng, Y.S. (1986). Bivariate lognormal distribution for characterizing asbestos fibre aerosols. *Aerosol Sci. Technol.*, 5(3):259–368.

[17] Rumberg, B., Richard, A. and Claiborn, C. (2001). Statistical Distribution of particulate matter and the error associated with sampling frequency. *Atmospheric Environment*, 35:2907-2920.

[18] Norazian, M.N., Abdulla, M.M.A., Tan, C.Y., Ramlic, N.A., Yahayac, A.S. and Fitric, N.F.M.Y. (2011). Modelling of PM10 Concentration for Industrialized Area in Malaysia: A case Study in Shah Alam. *Physics Procedia*, 22:318-324.

[19] Sedek, J.N.M., Ramli, N.A. and Yahaya, A.S. (2006). Air quality predictions using log normal distribution functions of particulate matterin Kuala Lumpur. *Malaysian journal of Environmental Management*, 7:33-41.

[20] Hamid, H., Yahaya, A.S., Ramli, N.A. and Ul-saufie, A.Z. (2013). Finding the best statistical distribution model in PM10 concentration modeling by using log normal distribution. *Journal of Applied Sciences*, 12(2):294-300.

[21] Georgopoulus, P.G. and Seinfeld, J.H. (1982). Statistical Distribution of air quality concentrations. *Environmental Science and Technology*, 16(7):401A-416A.

[22] Lai, H.K., Hedley, A.J., Thach, T.Q. and Wong, C.M. (2013). A method to derive the relationship between the annual and short-term air quality limits—Analysis using the WHO Air Quality Guidelines for health protection. *Environment International*, 59: 86–91.

[23] Nguyen, T.T.N. and Sevando, M. (2019). Assessing Coastal Water Quality through an Overall Index. *Pol. J. Environ. Stud*, 28(4):2321-2330.

[24] Wikipedia.org. https://en.wikipedia.org/wiki/Air_quality_index.

[25] Vaduva, I. (1973). Computer generation of random variables and vectors related to PERT problems. *In: Proceedings of the Fourth Conference on Probability Theory, Bucharest*, 381-395.

[26] Rousseeuw P.J. and Leroy A.M. Robust Regression and Outliers Detection, John Wiley & Sons, New York, 1987.

[27] Gupta, R.D. and Gupta, R.C. (1990). Estimation on Pr($\mathbf{a}'\mathbf{x} > \mathbf{b}'\mathbf{y}$) in the Multivariate Normal Case. *Statistics: A Journal of Theoretical and Applied Statistics*, 21(1):91-97.

[28] Reiser, B. and Farragi, D. (1994). Confidence bounds for P(a'x>b'y). *Statistics*, 25(2):107–111.

[29] Enis, P. and Geisser, S. (1971). Estimation of the Probability that Y < X. *J. Amer. Statist. Assoc.*, 66(333):162-168.

[30] Mukherjee, S.P. and Sharan, L.K. (1985). Estimation of failure probability from a bivariate normal stress-strength distribution. *Microelectronics Reliability*, 25(4):699-702.

[31] Hor, P. and Seal, B. (2015). Minimum variance unbiased estimation of stress–strength reliability under bivariate normal and its comparisons. *Communications in Statistics - Simulation and Computation*, 46(3):2447–2456.

[32] Johnson R.A. and Wichern D.W. Applied Multivariate Statistical Analysis, 6th Edition, Pearson, New Jersey, 2008.

[33] Barton, D.E. (1961). Unbiased estimation of a set of probabilities. *Biometrika*, 48(1-2):227–229.

[34] Mukhopadhyay P. Multivariate statistical analysis, World Scientific, Singapore, 2009.

[35] Johnson, N.L. and Welch, B.L. (1939). Applications of the non-central t distribution. *Biometrika*, 31:362-389.

[36] Efron B. The Jackknife, the Bootstrap and Other Resampling Plans, CBMS-NSF Regional Conference Series in Applied Mathematics, Monograph 38, SIAM, Philadelphia, 1982.

[37] Efron, B. (1987). Better bootstrap confidence intervals. *J. Am. Stat. Assoc*, 82(397):171–185.