

Inference on Stress-Strength Reliability for Log-Normal Distribution based on Lower Record Values

Nkemnole, E. Bridget, Samiyu, M. Abiodun

Department of Mathematics, University of Lagos, Lagos, Nigeria.

(enkemnole@unilag.edu.ng, talk2embee@yahoo.com¹)

Abstract

Research has suggested that there are components or devices which survive due to their strength. Although, these devices survive under a certain level of stress but when a higher level of stress is applied on them, they failed because they can't sustain it. The likelihood that these components are functional during a certain level of stress under a stated condition and a specified operational environment is regarded as its reliability, which in reliability engineering studies can be used to control, evaluate and estimate the capability and lifetime of a device.

This study aims to further contribute to the estimation of the stress-strength reliability parameter $R = P(Y < X)$, where X and Y are independent lognormal distributions based only on the first-observed lower record values. The Maximum Likelihood Estimator (MLE) of R and its asymptotic distribution are obtained as well as the confidence interval. Different parametric bootstrap confidence intervals are also proposed. Simulation and real data set representing Block-Moulding Machine experiment data (of Tola Block industry, Lagos, Nigeria) are fitted using the lognormal distribution and used to estimate the stress-strength parameters and reliability. Empirical analysis shows that the proposed model helps to establish a proficient structure for stress-strength reliability models.

Key Words: Lognormal Distribution, Reliability, Interval Estimation, Lower record values, Stress-Strength Reliability.

1. Introduction

The estimation of stress-strength reliability function based on record values has attracted the attention of authors because of its important role in industrial tests. Record values have a great important role in real life problems involving data relating to several fields such as weather, economics and sports data. The statistical study of record values began with Chandler [8] who introduced the main idea of record values, record times, inter record times and formulated the theory of record values as a model for successive extremes in a sequence of independently and identically distributed random variables.

Some applications used the inverted Weibull distribution as a model of a variety of failure characteristics such as infant mortality, useful life and wear-out periods Khan [14].

Recently, the growing interest about the estimation of stress-strength reliability, R associated with record values have been raised in many fields such as industrial test. The estimation of stress-strength R based on lower records is considered by Hassan et al [13] and a further research was done by Amal S. et al [2] and Hassan with the title “estimation of Stress-Strength Reliability for exponentiated inverted Weibull distribution”. Also, the estimation of stress-strength, R based on record values was considered by Baklizi [5] for generalized exponential distribution. Subsequent papers extended this work for some lifetime models. For instance, Baklizi [7] for one and two parameters exponential distribution, Essam [12] for type I generalized logistic distribution, Baklizi [6] for two parameter Weibull distribution, Bahman and Hossein [20] for inverse Rayleigh distribution, Al-Gashgari, Shawky [4] for exponentiated Weibull distribution.

However, this study deals with the estimation of stress-strength reliability of a log-normal distribution based on lower record values. The log-normal distribution can serve as an alternative to the exponentiated inverted Weibull distribution based on lower record values and in some situations, it is an improvement of it in a simpler form.

2. Related Literature

A great deal of attention has been paid on the estimation of Stress-Strength reliability, $R = P(Y < X)$ using various lifetime distributions and lots of concepts that relate to it. The estimation of stress-strength R based on lower records is considered by Hassan et al [13]. His work deals with the estimation of $R = P(Y < X)$, where X and Y are two independently exponentiated inverted Weibull distribution random variables based on lower record values. It

was assumed that the scale parameter is known and both the maximum likelihood estimate and the exact confidence interval of R were derived. In addition, the Bayes estimate of R based on independent gamma prior for the unknown parameters were obtained under squared error and LINEX loss functions. Also, analysis of simulated data was performed to compare the different estimators and to investigate the coverage probabilities of confidence interval. It was concluded in the paper that the coverage percentage of MLE is better than the coverage percentage of the Bayes estimator at $R = 0.25$. Also, from the result of the study, it was realized that the average confidence interval lengths of the MLE are shorter than the corresponding average confidence interval lengths of the Bayes estimators. Furthermore, the MSEs of the Bayesian estimator under Linear Exponential (LINEX) loss function are less than the Mean Square Errors (MSE) of MLE at different exact values of stress-strength reliability function R . Generally, based on the study, it was concluded that the MLE is better than the Bayes estimates (under squared error and LINEX loss functions).

Also, the estimation of stress-strength, R based on record values was considered by Baklizi [5] for generalized exponential distribution. This paper studied inference for the stress–strength reliability based on lower records data, where the stress and the strength variables are modelled by two independent but not identically distributed random variables from distributions belonging to the proportional reversed hazard family. Likelihood and Bayesian estimators were derived, then the confidence intervals and credible sets were obtained. Moreover, the paper considered the Topp–Leone distribution as a particular case of distribution belonging to the aforementioned family and derived some numerical results in order to show the performance of the proposed procedures. Finally, two applications to real data were reported.

Subsequent papers extended this work for some lifetime models. For instance, Baklizi [6] for two parameter Weibull distribution, Baklizi [7] for one and two parameters exponential distribution, Essam [12] for type I generalized logistic distribution, Bahman and Hossein [20] for inverse Rayleigh distribution, Al-Gashgari and Shawky [4] for exponentiated Weibull distribution. The prominent theoretical contributions and inference issues of the record values have been proposed by Ahsanullah [1].

3. Methodology

3.1 Likelihood Inference

In this section, we obtained the reliability function, R , the maximum likelihood estimate (MLE) of R , six different methods for confidence interval estimation of R . We also formulated a hypothesis for comparing the means of the two lognormal distributions.

3.2 Reliability Function R

Let X be the strength of a component which is subjected to the stress, Y . Assuming that $X \sim LN(\mu_1, \sigma_1^2)$ and $Y \sim LN(\mu_2, \sigma_2^2)$, the reliability function $R = P(Y < X)$ such that $X \sim LN(\mu_1, \sigma_1^2)$ and $Y \sim LN(\mu_2, \sigma_2^2)$ was considered by Nguimkeu et al [18]. However, this research is interested in a case where the reliability function is such that:

$R = P(\log Y < \log X)$. Suppose the joint pdf, is $f(x, y)$, then the reliability of the component is defined as;

$$R = P(Y < X) = \int_{-\infty}^{\infty} \int_{-\infty}^x f(x, y) dy dx \quad (3.1)$$

Since the random variables are statistically independent, then; $f(x, y) = f(x)g(y)$ so that;

$$R = P(Y < X) = \int_{-\infty}^{\infty} \int_{-\infty}^x f(x, y) g(y) dy dx \quad (3.2)$$

Suppose $G_y(x) = \int_{-\infty}^x g(y) dy$, we considered a case where $G_y(x) = \int_0^x g(y) dy$ since the considered distribution is non-negative.

$$\text{Then } R = P(Y < X) = \int_0^{\infty} G_y(x) f(x, y) dx \quad (3.3)$$

In this case, with loss of generality, we can take $G_y(x)$ to be the standard normal distribution, Φ_y .

$$\text{Thus, } R = P(Y < X) = P(Y - X < 0) = \Phi(Y) \quad (3.4)$$

Where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution and it is known that $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$ which defines the standard Gaussian cdf of Φ .

Suppose we consider a lognormal random variable Z such that $Z = Y - X$ then,

$$\mu_z = \mu_2 - \mu_1 \text{ and } \sigma_z^2 = \sigma_1^2 + \sigma_2^2 \Rightarrow \sqrt{\sigma_1^2 + \sigma_2^2} \quad (3.5)$$

$$\text{So, } R = P(Z > 0) = 1 - \Phi\left(-\frac{\mu_z}{\sigma_z}\right) = \Phi\left(\frac{\mu_z}{\sigma_z}\right)$$

by (3.4) and (3.5), it implies that;
$$Y = \frac{\mu_2}{\sigma_2^2} = \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} \quad (3.6)$$

Therefore,
$$R = \Phi(Y) = \Phi\left(\frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}\right) \quad (3.7)$$

3.3 MLE of the Reliability Function R

This section considers estimating R based on lower record values on both variables X and Y . Let $\underline{r} = r_1, r_2, r_3, \dots, r_n$ be a set of the first observed lower record values of size $(n + 1)$ from a lognormal (LN) distribution with parameters (μ_1, σ_1^2) and $\underline{s} = s_1, s_2, s_3, \dots, s_m$ be a set of the first observed lower record values of size $(m + 1)$ from LN with parameter (μ_2, σ_2^2) . Then the likelihood function as given by Ahsanullah (2015) is defined as;

$$L(\mu_1, \sigma_1^2 | \underline{r}) = f(r_n) \prod_{i=1}^{n-1} \frac{f(r_i)}{F(r_i)} \quad (3.8)$$

where, $0 < r_n < r_{n-1} < \dots < r_0 < \infty$

$$L(\mu_2, \sigma_2^2 | \underline{s}) = g(s_m) \prod_{i=1}^{m-1} \frac{g(s_i)}{G(s_i)}, \quad 0 < s_m < s_{m-1} < \dots < s_0 < \infty \quad (3.9)$$

where f and F are the pdf and cdf of $X \sim LN(\mu_1, \sigma_1^2)$ g and G are the pdf and cdf of $Y \sim LN(\mu_2, \sigma_2^2)$ respectively.

The likelihood functions;

$$L(\mu_1, \sigma_1^2 | x) = \prod_{i=1}^{n-1} \left(\frac{1}{x_i}\right) N(\mu_1, \sigma_1^2 | \ln x_i) = \prod_{i=1}^{n-1} \left(\frac{1}{x_i}\right) \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(\ln x_i - \mu_1)^2} \quad (3.10)$$

Similarly,

$$L(\mu_2, \sigma_2^2 | y) = \prod_{i=1}^{m-1} \left(\frac{1}{y_i}\right) N(\mu_2, \sigma_2^2 | \ln y_i) = \prod_{i=1}^{m-1} \left(\frac{1}{y_i}\right) \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2\sigma_2^2}(\ln y_i - \mu_2)^2} \quad (3.11)$$

By substituting f and F and using (3.10), we have:

$$L(\mu_1, \sigma_1^2 | \underline{r}) = f(r_n) \prod_{i=1}^n \left[\left(\frac{1}{r_i}\right) \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(\ln r_i - \mu_1)^2} \right] = \left[\prod_{i=1}^n \left(\frac{1}{r_i}\right) \right] \frac{1}{(2\pi\sigma_1^2)^{\frac{n}{2}}} e^{-\frac{1}{2\sigma_1^2} \sum_{i=1}^n (\ln r_i - \mu_1)^2}$$

Obtaining the log-likelihood of the above, we have;

$$\begin{aligned}
L(\mu_1, \sigma_1^2 | \underline{r}) &= \ln[\prod_{i=1}^n r_i^{-1}] + \ln(2\pi\sigma_1^2)^{-\frac{n}{2}} - \frac{1}{2\sigma_1^2} \sum_{i=1}^n (\ln r_i - \mu_1)^2 \\
&= -\sum_{i=1}^n \ln(r_i) - \frac{n}{2} \ln(2\pi\sigma_1^2) - \frac{1}{2\sigma_1^2} \sum_{i=1}^n (\ln r_i - \mu_1)^2
\end{aligned} \tag{3.12}$$

By substituting g and G and using (3.11), we have:

$$\begin{aligned}
L(\mu_2, \sigma_2^2 | \underline{s}) &= f(s_m) \prod_{i=1}^m \left[\left(\frac{1}{s_i} \right) \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2\sigma_2^2} (\ln y_i - \mu_2)^2} \right] \\
&= \left[\prod_{i=1}^m \left(\frac{1}{s_i} \right) \right] \frac{1}{(2\pi\sigma_2^2)^{\frac{m}{2}}} e^{-\frac{1}{2\sigma_2^2} \sum_{i=1}^m (\ln s_i - \mu_2)^2}
\end{aligned}$$

Obtaining the log-likelihood of the above, we have:

$$\begin{aligned}
L(\mu_2, \sigma_2^2 | \underline{s}) &= \ln[\prod_{i=1}^m s_i^{-1}] + \ln(2\pi\sigma_2^2)^{-\frac{m}{2}} - \frac{1}{2\sigma_2^2} \sum_{i=1}^m (\ln s_i - \mu_2)^2 \\
&= -\sum_{i=1}^m \ln(s_i) - \frac{m}{2} \ln(2\pi\sigma_2^2) - \frac{1}{2\sigma_2^2} \sum_{i=1}^m (\ln s_i - \mu_2)^2
\end{aligned} \tag{3.13}$$

Therefore, the joint log-likelihood of the observed \underline{r} and \underline{s} denoted by \underline{f} is of the form; $\underline{f} = \underline{f}_1 + \underline{f}_2$, hence,

$$\begin{aligned}
\underline{f} &= -\sum_{i=1}^n \ln(r_i) - \frac{n}{2} \ln(2\pi\sigma_1^2) - \frac{1}{2\sigma_1^2} \sum_{i=1}^n (\ln r_i - \mu_1)^2 - \sum_{i=1}^m \ln(s_i) - \frac{m}{2} \ln(2\pi\sigma_2^2) - \\
&\quad \frac{1}{2\sigma_2^2} \sum_{i=1}^m (\ln s_i - \mu_2)^2
\end{aligned} \tag{3.14}$$

To maximize the likelihood function through solving for the first-order conditions, we obtain the maximum likelihood estimators of μ_1, μ_2, σ_1^2 and σ_2^2 denoted by $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ based on the first observed lower record values by solving the following equations:

$$\frac{\partial \underline{f}}{\partial \mu_1} = 0, \quad \frac{\partial \underline{f}}{\partial \mu_2} = 0, \quad \frac{\partial \underline{f}}{\partial \sigma_1^2} = 0 \text{ and}$$

$$\frac{\partial \underline{f}}{\partial \sigma_2^2} = 0$$

$$\begin{aligned} \frac{\partial f}{\partial \mu_1} &= \frac{\partial}{\partial \mu_1} \left[-\frac{1}{2\sigma_1^2} \sum_{i=1}^n (\ln r_i - \mu_1)^2 \right] = \\ \frac{\partial}{\partial \mu_1} \left[-\frac{1}{2\sigma_1^2} \left[\sum_{i=1}^n (\ln r_i)^2 - 2\mu_1 \sum_{i=1}^n \ln r_i + \sum_{i=1}^n \mu_1^2 \right] \right] &= 0 \end{aligned}$$

$$\Rightarrow -\frac{1}{2\sigma_1^2} \left[-2 \sum_{i=1}^n \ln r_i + 2n\mu_1 \right] = -2 \sum_{i=1}^n \ln r_i + 2n\mu_1 = 0$$

$$\text{Hence, } \hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^n \ln r_i \quad \text{or} \quad \hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^n \log r_i \quad (3.15)$$

Similarly,

$$\frac{\partial f}{\partial \mu_2} = \frac{\partial}{\partial \mu_2} \left[-\frac{1}{2\sigma_2^2} \sum_{i=1}^m (\ln s_i - \mu_2)^2 \right] = \frac{\partial}{\partial \mu_2} \left[-\frac{1}{2\sigma_2^2} \left[\sum_{i=1}^m (\ln s_i)^2 - 2\mu_2 \sum_{i=1}^m \ln s_i + \sum_{i=1}^m \mu_2^2 \right] \right] = 0$$

$$\Rightarrow -\frac{1}{2\sigma_2^2} \left[-2 \sum_{i=1}^m \ln s_i + 2m\mu_2 \right] = -2 \sum_{i=1}^m \ln s_i + 2m\mu_2 = 0$$

$$\text{Hence, } \hat{\mu}_2 = \frac{1}{m} \sum_{i=1}^m \ln s_i \quad \text{or} \quad \hat{\mu}_2 = \frac{1}{m} \sum_{i=1}^m \log s_i \quad (3.16)$$

$$\text{Also, } \frac{\partial f}{\partial \sigma_1^2} = \frac{\partial}{\partial \sigma_1^2} \left[-\frac{n}{2} \ln(2\pi\sigma_1^2) - \sum_{i=1}^n \frac{(\ln r_i - \mu_1)^2}{2\pi\sigma_1^2} \right] = -\frac{n}{2\sigma_1^2} + \frac{\sum_{i=1}^n (\ln r_i - \mu_1)^2}{2\pi\sigma_1^4} = 0$$

$$\Rightarrow \sigma_1^2 = \frac{1}{n} \sum_{i=1}^n (\ln r_i - \mu_1)^2$$

$$\text{Hence, } \hat{\sigma}_1^2 = \frac{1}{n} \sum_{i=1}^n (\ln r_i - \mu_1)^2 \quad \text{or} \quad \hat{\sigma}_1^2 = \frac{1}{n} \sum_{i=1}^n (\log r_i - \mu_1)^2 \quad (3.17)$$

Similarly,

$$\frac{\partial f}{\partial \sigma_2^2} = \frac{\partial}{\partial \sigma_2^2} \left[-\frac{m}{2} \ln(2\pi\sigma_2^2) - \sum_{i=1}^m \frac{(\ln s_i - \mu_2)^2}{2\pi\sigma_2^2} \right] = -\frac{m}{2\sigma_2^2} + \frac{\sum_{i=1}^m (\ln s_i - \mu_2)^2}{2\pi\sigma_2^4} = 0$$

$$\Rightarrow \sigma_2^2 = \frac{1}{m} \sum_{i=1}^m (\ln s_i - \mu_2)^2$$

$$\text{Hence, } \hat{\sigma}_2^2 = \frac{1}{m} \sum_{i=1}^m (\ln s_i - \mu_2)^2 \quad \text{or} \quad \hat{\sigma}_2^2 = \frac{1}{m} \sum_{i=1}^m (\log s_i - \mu_2)^2$$

$$(3.18)$$

Equations (3.15) to (3.18) forms the maximum likelihood estimates of μ_1, μ_2, σ_1^2 and σ_2^2 .

By the invariance principle, the MLE of R denoted by \hat{R} can be obtained from (3.4) by;

$$\hat{R} = \phi(\hat{Y}) = \phi(Y(\hat{\theta})) = \phi\left(\frac{\hat{\mu}_2 - \hat{\mu}_1}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}}\right) \quad (3.19)$$

$$\text{Where } \hat{Y} = \frac{\hat{\mu}_2 - \hat{\mu}_1}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}} \quad \hat{\theta} = \hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2 \quad \text{and} \quad \hat{\sigma}_2^2$$

3.4 Asymptotic Distribution and Confidence Interval

$\phi(\cdot)$ is a monotone-increasing function of Y , by finding a confidence bound for Y , the confidence bound for $R = \phi(Y)$ is obtained Ngumkeu (2014). Since the variances are not assumed equal, there exists no exact distribution for \hat{Y} , only approximate distributions have been proposed in the literature. So we first obtain the asymptotic distribution of \hat{R} .

$$\text{We know that; } \mu_z = \mu_2 - \mu_1 \sim \text{LN}\left(\mu_2 - \mu_1, \frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}\right) \quad (3.20)$$

$$\text{Thus, } n \frac{\hat{\sigma}_1^2}{\sigma_1^2} \sim X_{n-1}^2 \quad \text{and} \quad m \frac{\hat{\sigma}_2^2}{\sigma_2^2} \sim X_{m-1}^2.$$

Since the unbiased estimators of σ_1^2 and σ_2^2 are s_1^2 and s_2^2 respectively, then;

$$(n-1) \frac{s_1^2}{\sigma_1^2} \sim X_{n-1}^2 \quad \text{and} \quad (m-1) \frac{s_2^2}{\sigma_2^2} \sim X_{m-1}^2.$$

3.5 Methods of obtaining confidence interval for Reliability Function R

In this research, we used the parametric method (proposed method) to obtain our confidence interval of \hat{R} . Then, we compare the performance of this proposed method to some other existing methods in literature namely; Normal Bootstrap Method (NBM), Raiser & Guttman approximation (RG), Likelihood-based first order Approximation Method (LBAM), Naïve Method (NM) and the Logit Transformation Method (LTM).

3.5.1 Normal Bootstrap Method (NBM)

The procedure is described as follows:

- (i) Estimate the parameters $(\mu_1, \mu_2, \sigma_1^2 \text{ and } \sigma_2^2)$ of the lower record values \underline{r} and \underline{s} of the random lognormal variables X and Y using the maximum likelihood estimate approach. That is, obtain $(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2 \text{ and } \hat{\sigma}_2^2)$ using (3.15) through (3.18), then estimate R using (3.19).
- (ii) Draw independently a bootstrap sample \underline{r}^* of size n from random variable X^* with the same distribution family of \underline{r} with parameters μ_1^* and σ_1^{2*} . In a similar procedure, we do the same for \underline{s}^* of size m with the same distribution family of \underline{s} with parameters μ_2^* and σ_2^{2*} .
- (iii) Estimate R^* using a similar expression for the estimator of R , as used in (3.7).
- (iv) Repeat steps (ii) and (iii) B-times (let B=2000), thus obtaining the MLE \hat{R}^* .
- (v) Estimate a $(1-\alpha)100\%$ percentile bootstrap (boot-n) confidence interval for R from \hat{R}^* distribution is defined as;

$$\left(\hat{R} - Z_{1-\frac{\alpha}{2}} \hat{S}.E_{boot}, \hat{R} + Z_{1-\frac{\alpha}{2}} \hat{S}.E_{boot} \right) \quad (3.21)$$

Where $\hat{S}.E_{boot}$ is a bootstrap estimate of the standard error based on $\hat{R}^*_1, \hat{R}^*_2, \dots, \hat{R}^*_B$

3.5.2 The Reiser & Guttman Approach (RG)

Reiser and Guttman (1986) examined statistical inference for $P(Y < X)$ where X and Y are independent lognormal variables with unknown mean and variance. They generalized the approach used by Geisser and Enis, 1971 (who considered the case where $m = n$ and $\sigma_1 = \sigma_2$) and also, the approach used by Church and Haris, 1970 (who considered the case where μ_1 and μ_2 are known).

As shown by Reiser and Guttman (RG), the distribution of \hat{Y} is derived as an approximation of Behrens-fisher which take the form of a non-central t-distribution. He stated an approximate $(1-\alpha)100th$ percentile of \hat{Y} which is defined as :

$$\hat{Y} \pm \left(\frac{1}{\hat{M}} + \frac{\hat{Y}}{2\hat{d}} \right) Z_{1-\frac{\alpha}{2}} \quad (3.22)$$

where $\hat{M} = \frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{\left(\frac{\hat{\sigma}_1^2}{n} + \frac{\hat{\sigma}_2^2}{m} \right)}$ or $\frac{s_1^2 + s_2^2}{\left(\frac{s_1^2}{n} + \frac{s_2^2}{m} \right)}$ and $\hat{d} = \frac{(\hat{\sigma}_1^2 + \hat{\sigma}_2^2)^2}{\left(\frac{\hat{\sigma}_1^4}{n-1} + \frac{\hat{\sigma}_2^4}{m-1} \right)}$ or $\frac{(s_1^2 + s_2^2)^2}{\left(\frac{s_1^4}{n-1} + \frac{s_2^4}{m-1} \right)}$

Hence, the $(1-\alpha)100th$ percentile of R is given by $(\phi(L), \phi(U))$

where $\Phi(L) = \hat{Y} - \left(\frac{1}{\hat{M}} + \frac{\hat{Y}}{2\hat{d}}\right) Z_{1-\frac{\alpha}{2}}$ and $\hat{Y} + \left(\frac{1}{\hat{M}} + \frac{\hat{Y}}{2\hat{d}}\right) Z_{1-\frac{\alpha}{2}}$ respectively.

3.5.3 Likelihood-Based First order Approximation Method (LBAM)

Based on the MLE of Y and the log-likelihood function, the confidence interval estimation of the parameter, Y using the standardized maximum likelihood estimate method (also known as the Wald method) which is based on the statistic (q) and defined by;

$$q = q(Y) = (\hat{Y} - Y)[var(\hat{Y})]^{-\frac{1}{2}} \quad (3.23)$$

Applying the Delta method to estimate $var(\hat{Y})$ in the above; we have;

$$var(\hat{Y}) = \frac{\partial Y(\hat{\theta})}{\partial \theta} var(\hat{\theta}) \frac{\partial Y(\hat{\theta})}{\partial \theta'} = \frac{1}{\hat{M}} + \frac{\hat{Y}}{4\hat{d}}$$

$$\text{Where } \hat{M} = \frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{\left(\frac{\hat{\sigma}_1^2}{n} + \frac{\hat{\sigma}_2^2}{m}\right)} \text{ or } \frac{s_1^2 + s_2^2}{\left(\frac{s_1^2}{n} + \frac{s_2^2}{m}\right)} \text{ and } \hat{d} = \frac{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}{\left(\frac{\hat{\sigma}_1^4}{n-1} + \frac{\hat{\sigma}_2^4}{m-1}\right)} \text{ or } \frac{s_1^2 + s_2^2}{\left(\frac{s_1^4}{n-1} + \frac{s_2^4}{m-1}\right)}$$

Since q is asymptotically distributed as standard normal, a $(1-\alpha)$ 100th percentile of Y can be approximated by;

$$\left(\hat{Y} - Z_{\frac{\alpha}{2}} \sqrt{var(\hat{Y})}, \hat{Y} + Z_{\frac{\alpha}{2}} \sqrt{var(\hat{Y})} \right) \quad (3.24)$$

where $Z_{\frac{\alpha}{2}}$ is the $(1-\alpha)$ 100th percentile of the standard normal. It should be noted that the Wald method is invariant to parameterization.

3.5.4 Naïve Method (NM)

Having known that $Y = \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}}$, the distribution of Y is asymptotically normal and consequently, \hat{Y} is a consistent estimator of Y [Church and Haris (2012)]. Thus, an approximate confidence interval for Y is obtained by;

$$P \left\{ \hat{Y} - \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sigma_Y < \frac{\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + \sigma_2^2}} < \hat{Y} + \Phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sigma_Y \right\} = 1 - \alpha$$

Replacing σ_Y by $\hat{\sigma}_Y$, where $\hat{\sigma}_Y = \frac{1}{\hat{M}} + \frac{\hat{Y}}{4\hat{d}}$

It can be easily demonstrated that $\hat{Y} - Y = O(n^{-1})$ with probability one; hence, (3.23) is a satisfactory approximate confidence interval for Y . Thus, an approximate confidence interval for R is given by;

$$P \left\{ \phi \left(Y - \phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{\text{var}(\hat{Y})} \right) < R < \phi \left(Y + \phi^{-1} \left(1 - \frac{\alpha}{2} \right) \sqrt{\text{var}(\hat{Y})} \right) \right\} = 1 - \alpha \quad (3.25)$$

3.5.5 Logit Transformation Method (LTM)

In this approach, recalling for the same reason given by Reiser and Guttman, we first estimate the asymptotic variance of the maximum likelihood estimator of Y which can be obtained by;

$$\text{var}(\hat{Y}) = \frac{1}{s^2} \left(\frac{s_1^2}{n} + \frac{s_2^2}{m} + \frac{1}{2} \frac{(\mu_1 - \mu_2)^2}{s^4} \left(\frac{s_1^2}{n} + \frac{s_2^2}{m} \right) \right) \text{ where } s^2 = s_1^2 + s_2^2,$$

Hence, an approximate $(1-\alpha)100\%$ confidence interval for \hat{Y} is given as;

$$(d(L), d(U)) = \hat{Y} \pm Z_{1-\frac{\alpha}{2}} \sqrt{\text{var}(\hat{Y})} \quad \text{where } \hat{Y} = \frac{\hat{\mu}_2 - \hat{\mu}_1}{\sqrt{\hat{\sigma}_1^2 + \hat{\sigma}_2^2}}$$

Thereby, an approximate confidence interval for R is given as;

$$(\phi d(L), \phi d(U)) \quad (3.26)$$

The variance estimate of \hat{R} can be employed to build a confidence interval by stabilizing it through a proper transformation. For normalizing the transformation $g(\hat{R})$, the approximate variance is obtained using the Delta method by; $\text{var}(g(\hat{R})) \cong (g'(\hat{R}))^2 \text{var}(\hat{R})$, Sharan (1985).

Hence the logit transformation as defined by Krishnamoorthy, (2010) is given as; $g(R) = \ln \left(\frac{R}{1-R} \right)$.

Hence, an approximate $(1-\alpha)100\%$ confidence interval is defined as;

$$(e^L(1+e^L)^{-1}, e^U(1+e^U)^{-1}) \quad (3.27)$$

where L and U are respectively the lower and upper bound of the confidence interval defined below;

$$\left(\ln \left(\frac{\hat{R}}{1-\hat{R}} \right) \pm Z_{1-\frac{\alpha}{2}} \frac{\sqrt{\text{var}(\hat{R})}}{\hat{R}(1-\hat{R})} \right) \quad (3.28)$$

3.5.6 Percentile Bootstrap Method (PBM)

This is an extension and application of parametric bootstrap for identical independent distribution case (Efron & Tibshiran, 1994) to a two sample case. It forms the basis of the proposed method considered in this research. The procedure is described as follows:

- (i) Estimate the parameters $(\mu_1, \mu_2, \sigma_1^2 \text{ and } \sigma_2^2)$ of the lower record values \underline{r} and \underline{s} of the random lognormal variables X and Y using the maximum likelihood estimate approach. i.e. Obtain $(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2 \text{ and } \hat{\sigma}_2^2)$ using (3.19), then estimate R using (3.19).
- (ii) Draw independently a bootstrap sample \underline{r}^* of size n from random variable X^* with the same distribution family of \underline{r} with parameters μ_1^* and σ_1^{2*} . In a similar procedure, we do the same for \underline{s}^* of size m with the same distribution family of \underline{s} with parameters μ_2^* and σ_2^{2*} .
- (iii) Estimate R^* using a similar expression for the estimator of R , as used in (3.7).
- (iv) Repeat steps (ii) and (iii) B -times (Let $B=2000$), thus obtaining the MLE \hat{R}^* .
- (v) Estimate a $(1-\alpha)100\%$ percentile bootstrap (boot-p) confidence interval for R from \hat{R}^* distribution taking the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles. This is defined as; $\left(\hat{R}^*_{\frac{\alpha}{2}}, \hat{R}^*_{1-\frac{\alpha}{2}} \right)$

4. Empirical Results and Discussion

4.1 Data

We applied the estimation methods of R as described in the previous section to model the Block-Moulding Machine experiment data (of Tola Block industry, Lagos, Nigeria) extracted from the robust performance of the machine as reported by the control engineer.

We are interested in testing the reliability of the component (machine) at the highest operating temperature 62°C at which the Operating pressure (Y) distribution tends to be closest to the combustion efficiency chamber (strength X). The datasets are given below.

(X): 22.95, 25.65, 24.45, 24.15, 25.20, 24.90, 25.65, 26.25, 24.15, 24.15, 24.00, 25.20, 26.25, 24.75, 24.00, 24.30, 26.25, 25.10, 24.35, 22.65

(Y): 11.6100, 11.6625, 11.5845, 11.6685, 11.9430, 11.1705, 12.1065, 11.8425, 12.1110, 11.2455,
11.3580, 11.6970, 11.8140, 12.2895, 12.0255, 11.9145, 11.8170, 11.5770, 10.9350, 11.6370

We compute the Cullen and Frey graph, Empirical and theoretical density, Empirical and theoretical CDFs, Q-Q plot and the P-P plot of the corresponding strength (combustion efficiency chamber) and the stress (Operating Pressure) data to ensure that the data fits the lognormal distribution. Their corresponding *AIC* values are **60.11262** and **16.63469** respectively. Figure (i)-Figure(vi) shows the graphical output and the inspection of the lognormal fit.

For a more formal test of agreement and analytical purpose with normality (or not), we applied the Shapiro-Wilk (S-W) tests for each data set to fit the normal and lognormal model. It was observed that for the Combustion Chamber Efficiency (CBE) data and the Operating Pressure (OP) data; the S-W Width are **0.94892** and **0.97077** with corresponding *p – values*; **0.3509** and **0.771** respectively. While for the log-transformed; the S-W Width are **0.94712** and **0.96627** with corresponding *p – values*; **0.3255** and **0.675** respectively. Furthermore, for the CBE and the OP data set, the chisquare values are **0.18561** and **0.77704** respectively. Therefore, it is clear that the normal and lognormal fits quite well to both data sets.

Next, we considered the first observed lower record values from the observed data (log-transformed) as follows;

$$\underline{r} = 3.267666, 3.184284, 3.178054, 3.120160$$

$$\underline{s} = 2.451867, 2.449668, 2.413276, 2.391969$$

Based on the data above, we obtained the Maximum Likelihood Estimators (MLE) of μ_1, μ_2, σ_1^2 and σ_2^2 denoted by $\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2$ and $\hat{\sigma}_2^2$ using (3.15), (3.16), (3.17) and (3.18) respectively. Hence we have;

$$\hat{\mu}_1 = 1.159114 \quad \hat{\mu}_2 = 0.886476 \quad \hat{\sigma}_1^2 = 4.114786 \quad \hat{\sigma}_2^2 = 2.372383.$$

We first obtain the MLE of \bar{Y} , denoted by $\hat{\bar{Y}}$. Applying (3.6), $\hat{\bar{Y}} = -0.107043$

Thus, the estimate of the reliability, R as defined in (3.19), is obtained by solving $\hat{R} = \Phi(\hat{\bar{Y}})$ recalling that; $pnorm(x) = \Phi(x)$. Hence, $\hat{R} = 0.4574$

Therefore, $R = P(Y < X)$ is the probability that in the Block-Moulding machine experiment, the combustion chamber efficiency (X) is higher than the Operating Pressure (Y) at the highest operating temperature ($62^{\circ}C$).

Figure 1. A Graph for inspecting the fit for the Lognormal Combustion Chamber (Strength) data

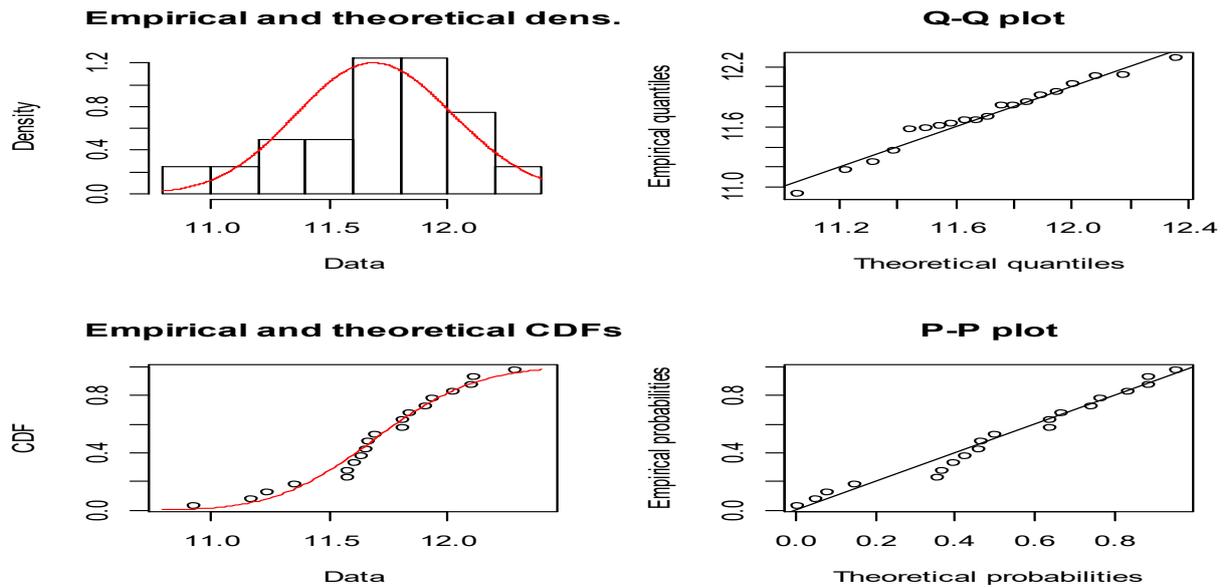
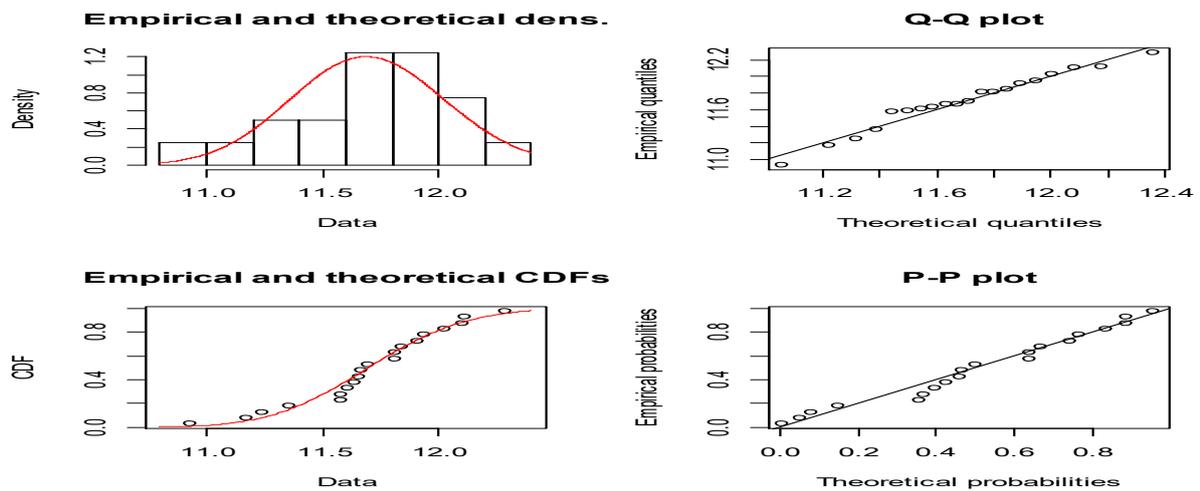


Figure 2. A Graph for inspecting the fit for the Lognormal Operating Pressure (Stress) data



We are interested in establishing that there is less than one in a million chances that the operating pressure, Y may exceed the combustion chamber efficiency strength, X (where we only considered the first observed lower record values). We compute the p -value function of

R (i. e. $p(R^0) = P(R > R^0)$) where R^0 is a specific value of R , for the six methods of estimation discussed.

It is of our interest to test the hypothesis; $H_0 : R = R^0$ against the alternative hypothesis; $H_a : R > R^0$. The higher the value of R^0 , the more accurate and reliable is to establish that there is less chances that the pressure may exceed the strength. In particular, a value of R^0 bigger than 0.4574 would mean that the experiment is reliable more than 50%. The lower and upper confidence bounds of the various methods is given in *Table 1*. The table below shows the result for the reliability values on Block-Moulding Machine Experiment Data at 90%, 95% and 99% confidence intervals (Table 1, below).

Methods	90%		95%		99%	
	Lower Bound	Upper Bound	Lower Bound	Upper Bound	Lower Bound	Upper Bound
RG	0.5274	0.5578	0.5244	0.5607	0.5188	0.5664
LBAM	0.2229	0.8356	0.1757	0.8741	0.1047	0.9291
NM	0.2222	0.8362	0.1759	0.8741	0.1041	0.9296
LTM	0.2321	0.8276	0.1835	0.8651	0.1115	0.9376
NBM	0.5422	0.5430	0.5421	0.5431	0.5420	0.5432
PBM	0.5423	0.5432	0.5424	0.5432	0.5423	0.5433

We can carefully observe from *Table 1* that the six methods give different results for the Block-Moulding Machine Experiment Data. Considering the lower bounds values, it is evident that the proposed method (PBM) yields better result than the other five methods described in the previous section of the study.

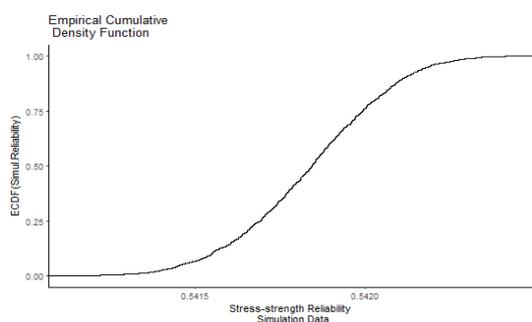
The confidence intervals calculated according to each of the methods described show that the proposed method provides the narrowest interval (this shall be further verified in our simulation study). The RG and NBM yields a better result showing a closely approximate result with the proposed method. However, the NM and the LTM values are very close to each other but far different from the values of other methods applied in the study.

More so, from the table, we can establish the validity of our hypothesis test. The NBM and the PBM (proposed method) provides values of R^0 which are all greater than the estimated R of the machine. However, the PMB yields greater values of the two; hence, we can accurately say

that there is less chances that the pressure may exceed the strength. In particular, since all the PMB values are bigger than 0.4574, it means that the experiment is reliable more than 50%.

4.2 Simulation Study

In order to study the accuracy of the six methods discussed in our study, a simulation study was conducted. We performed a bootstrap based on interval simulation. It is important to note that all inferences procedures in this study depend only on the first observed lower record values, r_n and s_m .



cdf plot of the Stress-Strength reliability estimates.

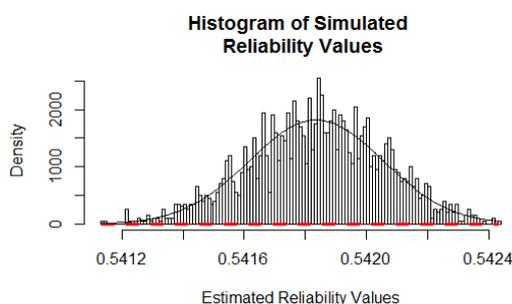


Fig 4: Histogram Plot for the simulated reliability estimates

Fig 3:

In the simulation design, we used all combinations of sample sizes from; $n=5,10,15$ & 20 and $m=5,10,15$ & 20 and the 90%, 95% and 99% confidence intervals for reliability were obtained for each combination of samples. For each combination of the simulation, we generated 2000 samples of the lower record values from $X \sim LN(1.159114, 4.114786)$ and $Y \sim LN(0.886476, 2.372383)$.

The confidence intervals are empirically investigated by simulation using the six methods discussed.

We considered different cases, each corresponding to a different combination of distributional parameters, thus we obtained different reliability values using the different sample sizes as shown in *Table 2*. Without any loss in generality, we varied the reliability values (three values). This was done in order to get a high value for the reliability, since in real practice; we usually look for a higher reliability in the study of component or system.

The reliability values were obtained using the different samples from combinations of $n=5,10,15$ & 20 and $m=5,10,15$ & 20 which is given below:

$\{(n,m)= (5,5), (5,10), (5,15), (5,20), (10,5), (10,10), (10,15), (10,20), (15,5), (15,10), (15,15), (15,20), (20,5), (20,10), (20,15), (20,20)\}$

Table 2. Coverage Rate and Expected length of the 95% Confidence Interval of n (Samples from Strength, X) & m (Samples from Stress, Y)

n	m	METHOD	EXPECTED LENGTH	COVERAGE RATE
5	5	PBM (Proposed Method)	0.0024	0.943
		NBM	0.0036	0.931
		RG	0.6790	0.889
		LBAM	0.6717	0.897
		LTM	0.6784	0.833
		NM	0.6794	0.930
5	10	PBM (Proposed Method)	0.0045	0.937
		NBM	0.0086	0.935
		RG	0.6790	0.913
		LBAM	0.6717	0.902
		LTM	0.6784	0.899
		NM	0.6794	0.910
5	15	PBM (Proposed Method)	0.0043	0.953
		NBM	0.0041	0.949
		RG	0.6792	0.896
		LBAM	0.6720	0.893
		LTM	0.6784	0.809
		NM	0.6794	0.915
5	20	PBM (Proposed Method)	0.0041	0.950
		NBM	0.0038	0.928
		RG	0.6791	0.897
		LBAM	0.6717	0.903
		LTM	0.6784	0.913
		NM	0.6794	0.937
10	5	PBM (Proposed Method)	0.0035	0.947
		NBM	0.0085	0.940
		RG	0.6792	0.879
		LBAM	0.6721	0.910
		LTM	0.6784	0.908
		NM	0.6794	0.931
10	10	PBM (Proposed Method)	0.0053	0.952
		NBM	0.0056	0.941
		RG	0.6791	0.911
		LBAM	0.6723	0.921
		LTM	0.6784	0.891
		NM	0.6794	0.908
10	15	PBM (Proposed Method)	0.0018	0.928
		NBM	0.0031	0.942
		RG	0.6791	0.933
		LBAM	0.6716	0.919
		LTM	0.6784	0.915
		NM	0.6794	0.928
10	20	PBM (Proposed Method)	0.0037	0.949
		NBM	0.0031	0.911
		RG	0.6790	0.930
		LBAM	0.6721	0.918
		LTM	0.6784	0.921
		NM	0.6794	0.925

15	5	PBM (Proposed Method)	0.0044	0.950
		NBM	0.0065	0.948
		RG	0.6791	0.927
		LBAM	0.6724	0.923
		LTM	0.6784	0.932
		NM	0.6794	0.933
15	10	PBM (Proposed Method)	0.0031	0.955
		NBM	0.0046	0.946
		RG	0.6792	0.940
		LBAM	0.6722	0.922
		LTM	0.6784	0.921
		NM	0.6794	0.924
15	15	PBM (Proposed Method)	0.0024	0.961
		NBM	0.0057	0.952
		RG	0.6791	0.940
		LBAM	0.6718	0.934
		LTM	0.6783	0.931
		NM	0.6794	0.935
15	20	PBM (Proposed Method)	0.0028	0.948
		NBM	0.0042	0.883
		RG	0.6791	0.929
		LBAM	0.6718	0.911
		LTM	0.6784	0.898
		NM	0.6794	0.912
20	5	PBM (Proposed Method)	0.0053	0.951
		NBM	0.0029	0.943
		RG	0.6792	0.938
		LBAM	0.6717	0.916
		LTM	0.6784	0.920
		NM	0.6793	0.917
20	10	PBM (Proposed Method)	0.0017	0.962
		NBM	0.0068	0.945
		RG	0.6791	0.916
		LBAM	0.6716	0.924
		LTM	0.6784	0.921
		NM	0.6793	0.934
20	15	PBM (Proposed Method)	0.0027	0.949
		NBM	0.0032	0.938
		RG	0.6791	0.922
		LBAM	0.6720	0.900
		LTM	0.6783	0.913
		NM	0.6794	0.926
20	20	PBM (Proposed Method)	0.0041	0.945
		NBM	0.0052	0.943
		RG	0.6791	0.926
		LBAM	0.6719	0.919
		LTM	0.6784	0.908
		NM	0.6794	0.928

From the table above, which represents the expected lengths and the coverage rate of the simulation results for the 95% confidence interval for the reliability parameter, it is evidently

shown that our proposed method yields better results than the other five methods described in this study.

The expected length of the NBM and the proposed method are very close, however that of the proposed method has a shorter length when carefully observed. The lengths of the LBAM is shorter when compared to the other three methods which are all far longer than the PMB and the NBM. Hence, in terms of performance, especially when the sample size combinations are very small and different, the proposed method (PMB), which possesses the shortest expected lengths, is the best.

In terms of coverage rate, the proposed method also yields the best coverage rates than the others which are all within the standard errors of the nominal values.

5. Conclusion

This work presented the stress-strength reliability of a component or device to assess whether its estimation procedure yields an accurate and robust result. Some previous studies have worked on different methods of estimating the reliability, $R = P(Y < X)$.

More recently, Nguimkeu *et al* (2014), proposed that the modified signed log-likelihood ratio test yields an accurate result in terms of coverage probability and error rate when they considered two independent normal distribution. Also, Tarvirdizade and Kazemzadeh (2016) also proposed that the percentile bootstrap method yields a better result in terms of coverage probability and expected length of confidence intervals when they considered two independent but not identically inverse Rayleigh distributed random variables based on lower record values. Meanwhile, in our study, we considered two independent lognormal distributions based on lower record values. We use a case study of Block-Moulding Machine Experiment whose stress-strength data was collected from Tola Block Industry, Lagos, Nigeria.

This study most important strength is that out of the six methods of estimation described in the study, the Percentile Bootstrap Method provides the best performance in terms of expected length of confidence interval and good coverage probability. It was found that we can accurately say that there is less chances that the machine operating pressure may exceed its strength at the highest operating temperature. In particular, all the proposed method values are bigger than our estimated reliability value, thus, we realized that the experiment is reliable more than **50%**.

This study has obtained both point and interval estimation procedure for the stress-strength reliability of a component which realistically is over **50%** reliable and robust. It also backs it up with numerical simulations at various significance levels. Thereby, we can conclude that among

the methods discussed in the study, the percentile bootstrap method yields the best and accurate results having the shortest expected length and a very good coverage rate.

References

- [1] M. Ahsanullah “Extreme Value Distributions”, Atlantis Studies in Probability and Statistics; Volume 8; ISBN: 978-94-6239-222-9, 2016.
- [2] S. Amal, Hiba Z., Mohammed S. “Estimation of Stress-Strength Reliability for Exponentiated Inverted Weibull Distribution Based on Lower Record Values”, British Journal of Mathematics & Computer Science, 2015; 11(2):1-14.
- [3] B. Alessandro Confidence Intervals for Reliability of Stress-Strength Models in the Normal Case. Communications in Statistics, 2011; 40(6):907-925.
- [4] F. Al-Gashgari , Shawky A. Estimation of $P(Y < X)$ Using Lower Record Data from the Exponentiated Weibull Distributions: Classical and MCMC Approaches. Life Science Journal, 2014; 11(7):768-777.
- [5] A. Baklizi Estimation of $\Pr(X < Y)$ using Record Values in the One and Two-Parameter
- [6] Exponential Distributions. Communications in Statistics-Theory and methods, 2008; 37:692-698.
- [7] Baklizi Inference on $\Pr(X < Y)$ in the Two-Parameter Weibull Model Based on Records. ISRN Probability and Statistics, 2012; Article ID: 263612: 1-11.
- [8] Baklizi A. Interval Estimation of Stress-strength reliability in the two-parameter exponential distribution based on records. Journal of Statistical Computation and Simulations, 2013; 84: 2670-2679.
- [9] Baklizi Bayesian Inference for $P(Y < X)$ in the Exponential Distribution Based on Records. Applied Mathematical Modelling, 2014; 38: 1698-1710.
- [10] Basu The estimation of $P(X < Y)$ for distributions useful in life testing. Naval Research Logistics Quarterly, 1981; 28:383-392.
- [11] K. Chandler The Distribution and Frequency of record Values. Journal of Royal Society, 1952. 14:220-228.
- [12] J. Church The Estimation of Reliability from Stress-Strength Relationships. Technometrics, 1970; 12: 49-54.
- [13] A. Essam The Sampling Distribution of the Maximum Likelihood Estimators from type I Generalized Logistic Distribution Based On Lower Record Values. International Journal of Contemporary Math. Sciences, 2012; 24:1205-1212.
- [14] Hassan et al. Estimation of Stress-Strength Reliability function Based on Lower Record
- [15] Values. British Journal of Mathematics and Computer Science, 2015; 11(2):1-14.
- [16] M. Khan Modified Inverse Rayleigh Distribution. Journal of statistics Applications and Probability, 2012; 2: 115-132.
- [17] S. Kotz , Lumelskii Y., Pensky M. The Stress-Strength Model and Its Generalizations: Theory and Applications, London: Imperial College Press, 2003; ISBN: 978-981-238-057-9.

- [18] Krishnamoorthy K. et al . New Inferential Methods for the Reliability Parameter in a Stress-
Strength Model; The Normal Case. *Communication in Statistics Theory and Methods*, 2014; 33:1715-1731.
- [19] M. Mazumdar Some Estimates of Reliability using Inference Theory. *Naval Research Logistics Quarterly*,1970; 17:159-165.
- [20] Nguimkeu P., Marie R., Augustine W. Interval Estimation of the Stress-Strength Reliability with Independent Normal Random Variables. *Communication in Statistics Theory and Methods*, 2014; 15:23-67.
- [21] R.G. Srinivasa Estimation of Stress-Strength Reliability from Inverse Rayleigh Distribution based on lower record values. *Journal of Quality and Reliability Engineering*, 2013; 30(4):256-263.
- [22] Bahman T., Hossein K. Interval Estimation of Stress-Strength Reliability Based on Lower Record Values from Inverse Rayleigh Distribution. *Journal of Quality and Reliability Engineering*, 2014; 2014: Article ID 192072.