PARAMETER ESTIMATION OF WEIBULL DISTRIBUTION USING MAXIMUM LIKELIHOOD

MBSA PROJECT REPORT

FINAL REPORT

GROUP MEMBERS-

VIDYA SAGAR	SHAIK VAZEER AHAMED	PERMULLA MANOJ ARUN KUMAR
S20200020312	S20200020301	S20200020293
sagarreddy.v20@iiits.in	vazeerahamed.sk20@iiits.in	arunkumar.p20@iiits.in

ABSTRACT-

The Weibull distribution is commonly used to model the lifetime of a system or component. Parameter estimation of the Weibull distribution plays a crucial role in reliability analysis and decision-making in engineering and other fields.

Two popular methods for estimating the parameters of the Weibull distribution are maximum likelihood estimation (MLE) and Bayesian estimation. In this report we are providing CRB, and detailed step-by-step procedure for performing parameter estimation using MLE of two-parameters using Newton Raphson method. We also applied Monte-Carlo for better estimation.We have applied our result to a dataset and we also computed the Kullback-Leibler Divergence Test and R2 Test.

INTRODUCTION-

Reliability analysis plays a crucial role in ensuring the safe and efficient operation of engineering systems. The two-parameter Weibull distribution is a continuous probability distribution that is commonly used to model the failure times of engineering components and systems.

The two-parameter Weibull distribution has several important properties, including:

- It is a flexible distribution that can model a wide range of failure modes, including early-life failures, wear-out failures, and random failures.
- It can exhibit a variety of failure rate shapes, including monotonically increasing, decreasing, constant, and bathtub shape.
- It is widely used in reliability engineering to model the lifetimes of components and systems and to estimate their failure probabilities, mean time to failure, and other reliability metrics.
- The two-parameter Weibull distribution can be estimated from data using various methods, including maximum likelihood estimation, least squares estimation, and Bayesian estimation.

- The applications of the Weibull Distribution are-

- Reliability analysis: The Weibull distribution is used extensively in reliability engineering to model the failure time of components and systems. It can be used to estimate the probability of failure and predict the lifetime of products and systems.

- Wind energy: The Weibull distribution is commonly used in the wind energy industry to model wind speed and energy production. Wind speed data can be analyzed using the Weibull distribution to estimate the probability of certain wind speeds occurring and to optimize the design of wind turbines.
- Medical research: The Weibull distribution is often used in medical research to analyze survival data, such as the time to death or the time to a specific event (e.g., recurrence of cancer). It can be used to estimate the survival probabilities and to compare the survival times of different groups.
- Materials science: The Weibull distribution is used in materials science to model the strength and reliability of materials. It can be used to estimate the probability of failure and to predict the lifetime of materials under different stress conditions.
- Insurance: The Weibull distribution can be used to model the frequency and severity of insurance claims. It can be used to estimate the probability of certain events occurring and to set insurance premiums based on risk assessment.

In this report we explore the use of the two-parameter Weibull distribution in reliability analysis and apply maximum likelihood estimation to estimate the parameters of the distribution. We also generate synthetic data from the Windmill dataset to evaluate the performance of different estimation methods. Specifically, we calculate the Cramer-Rao bound for the two-parameter Weibull distribution, estimate the shape and scale parameters of the distribution using the method of moments, and apply Monte-Carlo simulation to obtain a more accurate estimate of the parameters.

The two-parameter Weibull distribution,

If **T** $\sim$ **Weibull** **(**γ, β**)** then its density function is defined as:

$$f(t/\gamma, \beta) = \frac{\gamma}{\beta^{\gamma}} t^{(\gamma-1)} \exp exp \left\{ - \left(\frac{t}{\beta} \right)^{\gamma} \right\} \text{ , for } \gamma > 0 \text{ , } and \ \beta > 0$$

Where t is the failure time, β is the scale parameter, and γ is the shape parameter. The scale parameter β determines the location of the distribution, and the shape parameter γ determines the shape of the distribution.

Here, γ, β are unknown parameters, we are estimating the unknown parameters.

METHODOLOGY-

Maximum likelihood estimation (MLE) is a popular method for estimating the parameters of the two-parameter Weibull distribution. MLE provides estimates that are asymptotically unbiased and efficient, and it gives easy insights into the underlying distribution of the data and helps evaluate the performance of different estimation methods.

The **CRLB** is derived from the Fisher information, which is a measure of the amount of information that the data provide about the unknown parameter. The Fisher information is defined as the negative second derivative of the log-likelihood function with respect to the parameter, and it quantifies the amount of information contained in the data about the parameter.

The CRLB is given by the inverse of the Fisher information matrix, which is the matrix of second derivatives of the log-likelihood function with respect to the parameters. Specifically, for a single parameter θ, the CRLB is:

$$\textbf{Var } (\theta) \geq (I\ (\theta))^{-1}\textbf{.)}$$

where Var (θ) is the lower bound on the variance of any unbiased estimator of, and I (θ) is the Fisher information.

We are applying a log likelihood to our pdf and differentiating one unknown parameter keeping another unknown parameter as constant. For one parameter we are not getting a closed form solution, so we have used Newton-Raphson method and for other parameters we are getting a closed form solution.

Kullback–Leibler (KL) divergence, also known as relative entropy, is a measure of the difference between two probability distributions. It was first introduced by Solomon Kullback and Richard Leibler in 1951.

In simple terms, KL divergence measures how different one probability distribution is from another. It is often used in statistics, information theory, and machine learning to compare two probability distributions.

The KL divergence between two probability distributions P and Q is defined as follows:

$$\mathbf{KL(P \parallel Q) = \Sigma\, P(i) \log [P(i) / Q(i)]}$$

Where P(i) and Q(i) are the probabilities of the i-th event occurring in the two distributions, and log is the natural logarithm.The KL divergence is not a true distance metric, since it is not symmetric and does not satisfy the triangle inequality. However, it is a useful measure in many applications, such as in data compression, model selection, and hypothesis testing.The dataset used is the “shelf life of pickles” which follows a weibull distribution. We now estimate the parameters for the given data and then we plot it along with the weibull distribution with the parameters we acquired. Now we run the KL Test and we have found out the value is 0(Ideal Case) , which means it's a match with the PDF.

CALCULATIONS: -

Pdf of Weibull distribution:

$$f(y/\theta, \lambda) = \frac{\lambda}{\theta^{\lambda}} y^{(\lambda-1)} \exp exp \left\{ - \left(\frac{y}{\theta}\right)^{\lambda} \right\} \text{ , for } \theta > 0 \text{ , } and\, \lambda > 0$$

$$f\left(y_i/\theta, \lambda\right) = \prod_{i=1}^{N} \left(\frac{\lambda}{\theta^{\lambda}} y_i^{(\lambda-1)} \exp exp \left\{ - \left(\frac{y_i}{\theta}\right)^{\lambda} \right\}\right)$$

Applying log-likelihood:

$$\log f\left(y_i/\theta, \lambda\right) = \log \prod_{i=1}^{N} \left(\frac{\lambda}{\theta^{\lambda}} y_i^{(\lambda-1)} \exp exp \left\{ - \left(\frac{y_i}{\theta}\right)^{\lambda} \right\}\right)$$

$$= \sum_{i=1}^{N} \left\{ log\, \lambda + (\lambda - 1) log y_i - \lambda log \theta - \left(\frac{y_i}{\theta}\right)^{\lambda} \right\}$$

Differentiating w.r.t θ**:**

$$\frac{\partial L}{\partial \theta} = 0.$$

$$\Rightarrow \quad \frac{-N\lambda}{\theta} - \sum_{i=1}^{N} \left(y_i\right)^{\lambda}\left(-\lambda\theta^{-\lambda-1}\right) = 0$$

$$\boxed{\theta = \left(\frac{\sum_{i=1}^{N}\left(y_i\right)^{\lambda}}{N}\right)^{\frac{1}{\lambda}}}$$

closed form solution for θ.

Differentiating w.r.t λ:

$$\frac{\partial L}{\partial \lambda} = 0.$$

$$\frac{\partial L}{\partial \lambda} = \sum_{i=1}^{N}\left(\frac{1}{\lambda} + - \log log\, \theta \quad - \left(\frac{y_i}{\theta}\right)^{\lambda} \log log \left(\frac{y_i}{\theta}\right)\right)$$

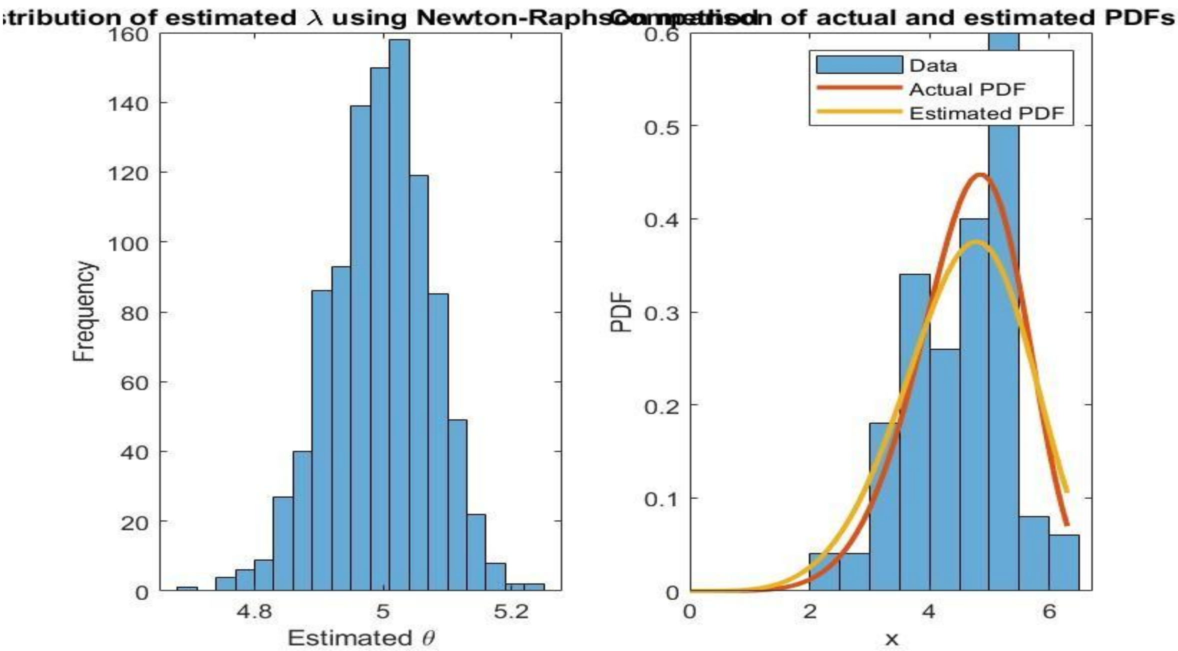
$$\frac{\partial^2 L}{\partial \lambda^2} = \frac{-N}{\lambda^2} - \sum_{i=1}^{N}\left(\frac{y_i}{\theta}\right)^{\lambda}\left(\log log \left(\frac{y_i}{\theta}\right)\right)^2$$

$$\lambda_1 = \lambda_0 - \frac{f'\left(\lambda_0\right)}{f''\left(\lambda_0\right)}$$

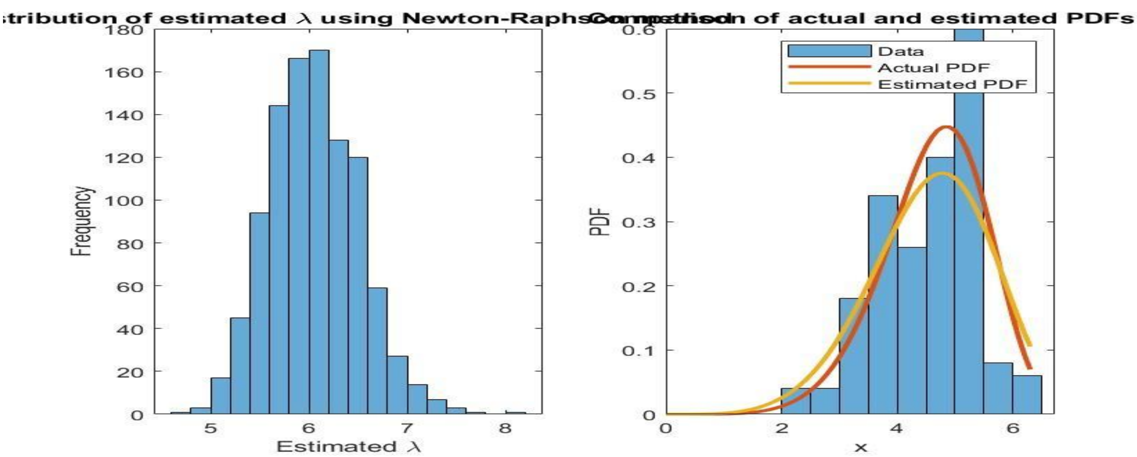
NR (Newton Raphson method)

RESULTS-

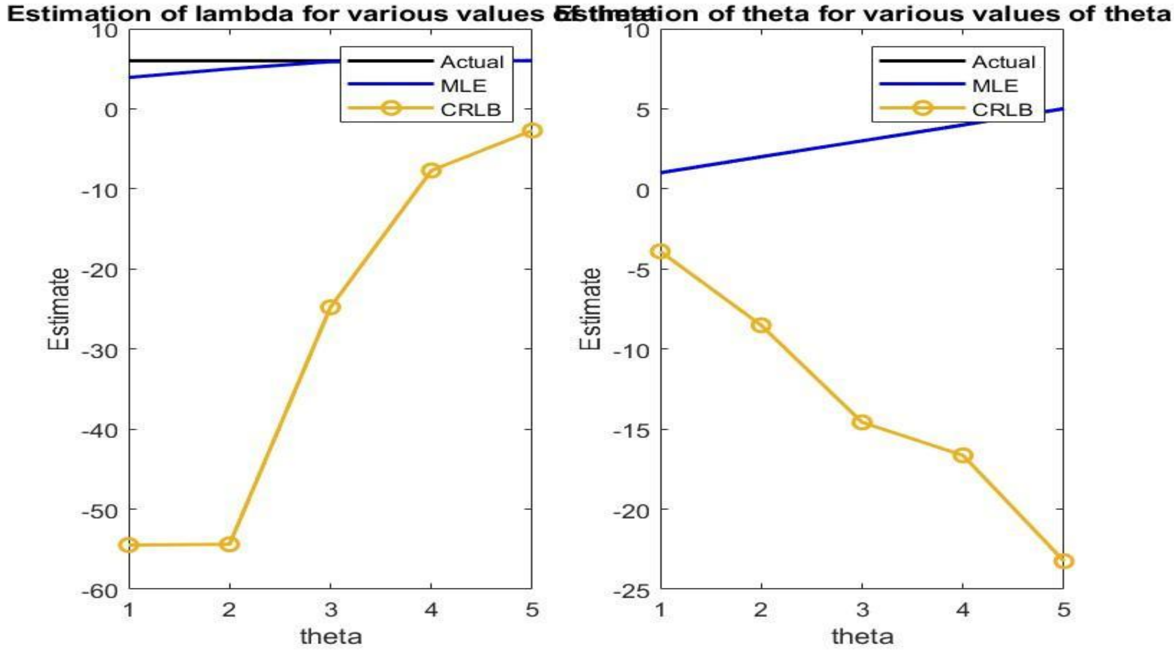
Theta:



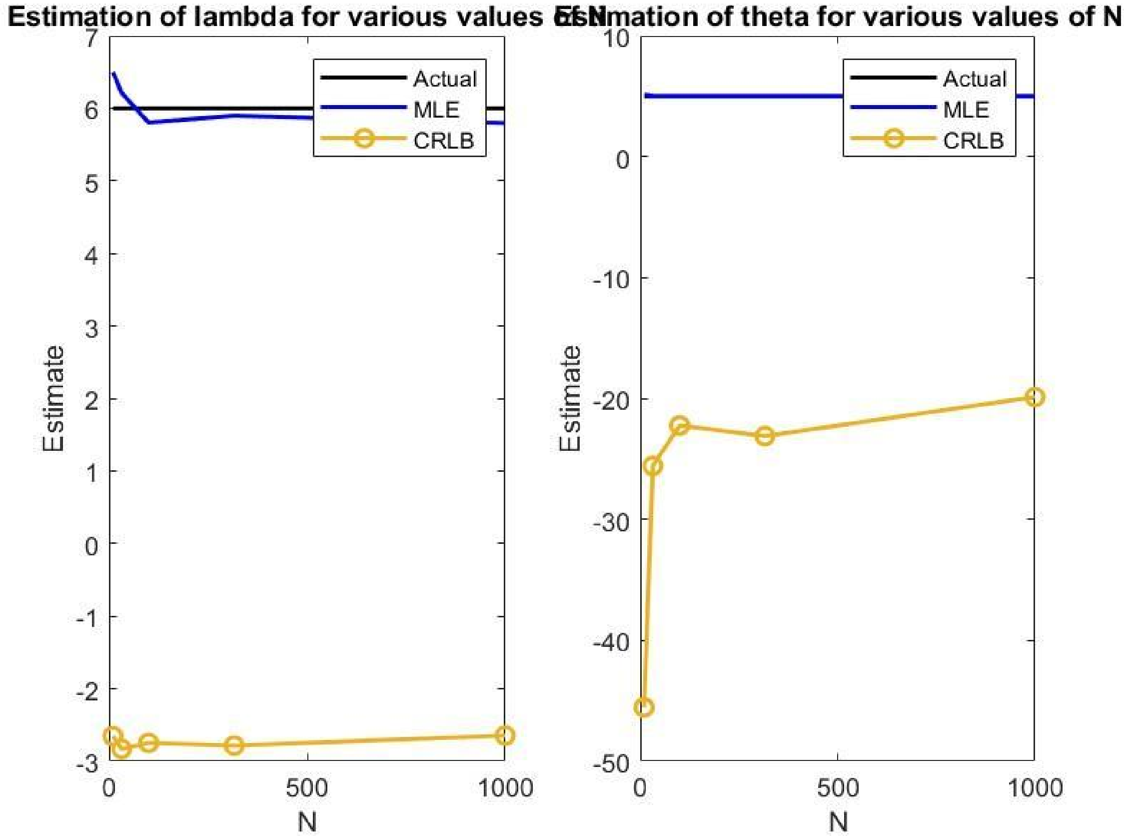
Lambda:



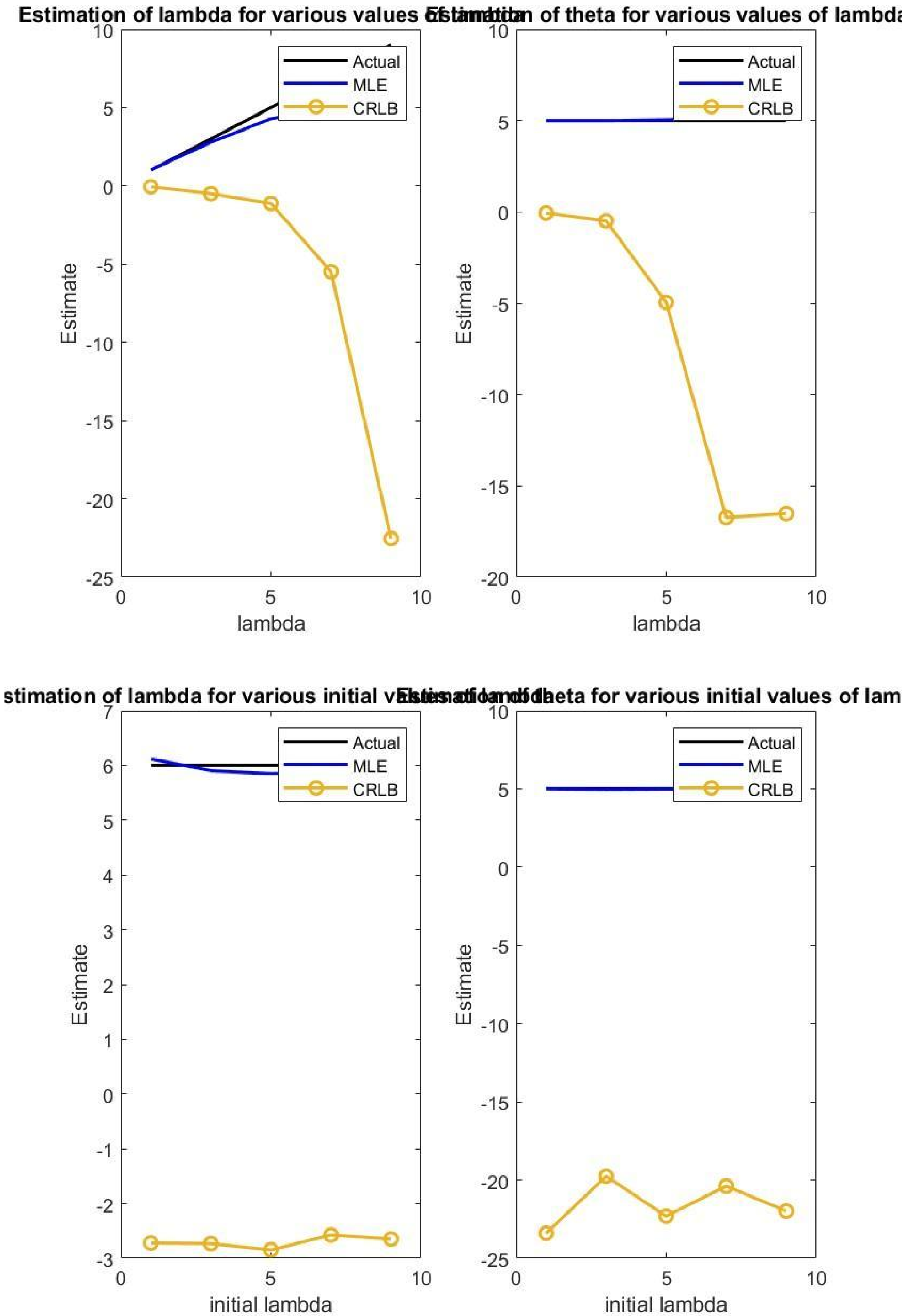
CRLB:



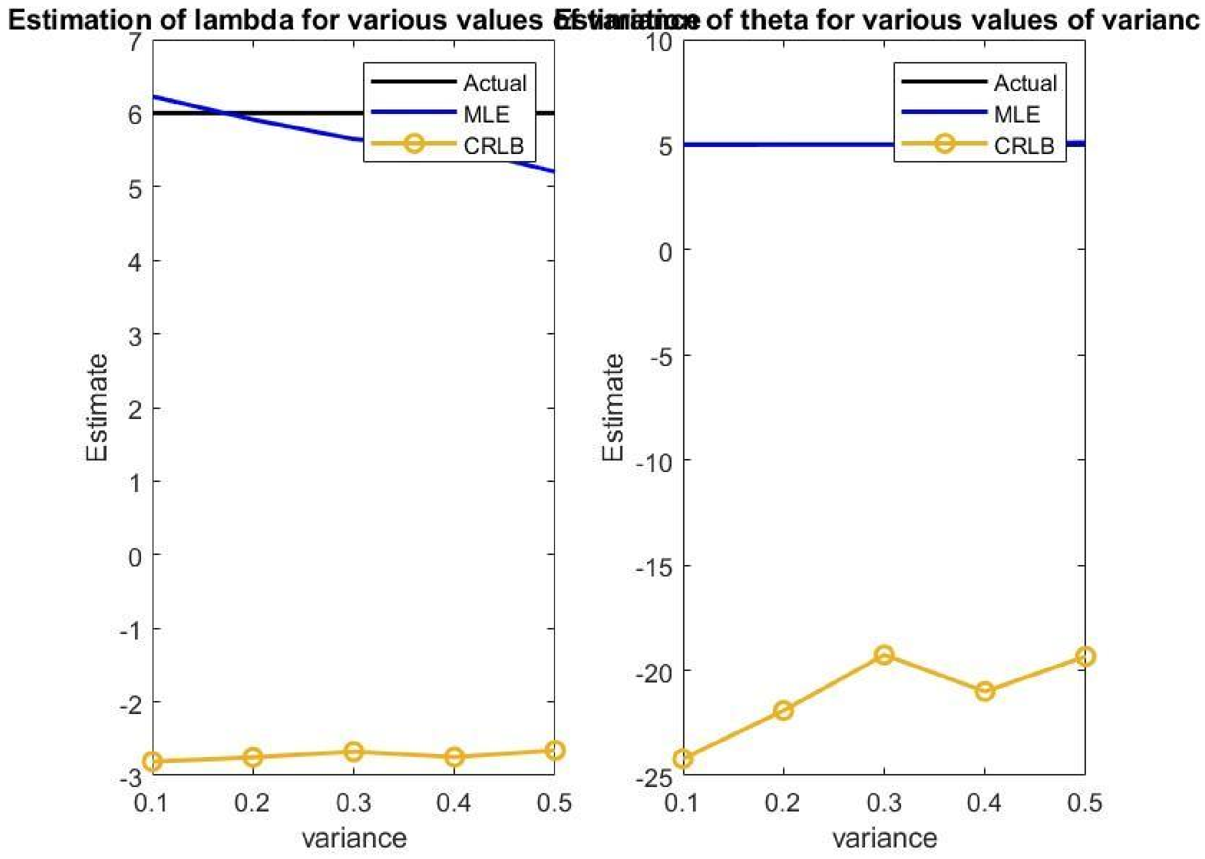
Estimation of lambda for various values of theta
Estimation of theta for various values of theta
Actual
MLE
CRLB
Estimate
theta

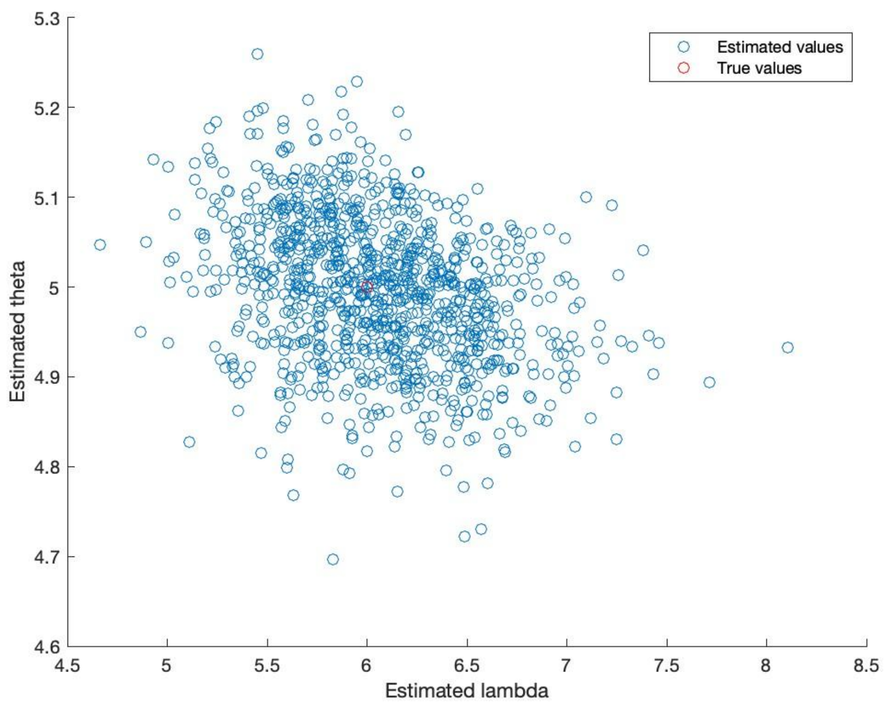


Estimation of lambda for various values of N
Estimation of theta for various values of N
Actual
MLE
CRLB
Estimate
N

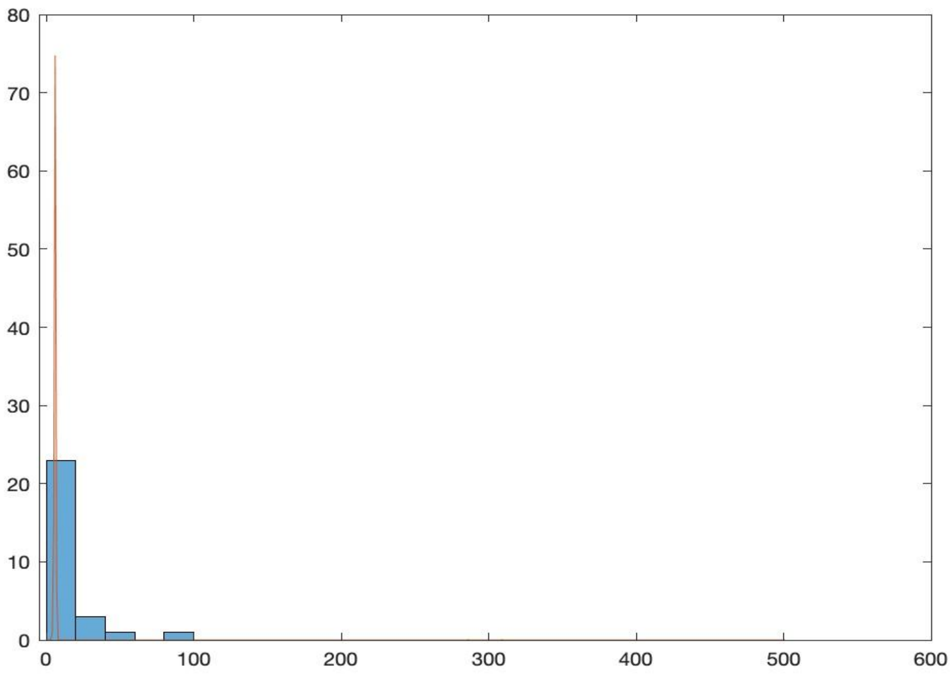


Estimation of lambda for various values of lambda
Estimation of theta for various values of lambda
stimation of lambda for various initial values of lambda
Estimation of theta for various initial values of lam
Actual
MLE
CRLB
Estimate
lambda
initial lambda


Estimation of lambda for various values of variance
Estimation of theta for various values of varianc
Actual
MLE
CRLB
Estimate
variance
7
6
5
4
3
2
1
0
-1
-2
-3
10
5
0
-5
-10
-15
-20
-25
0.1
0.2
0.3
0.4
0.5


Estimated values
True values
Estimated theta
Estimated lambda
5.3
5.2
5.1
5
4.9
4.8
4.7
4.6
4.5
5
5.5
6
6.5
7
7.5
8
8.5

KL-Test



CONCLUSION-

In conclusion, this paper has presented a comprehensive study on the problem of parameter estimation for the Weibull distribution using the maximum likelihood method. We have shown that the maximum likelihood method provides a reliable and efficient tool for estimating the parameters of the Weibull distribution. We have derived the likelihood function and demonstrated how to maximize it using numerical optimization techniques(Newton Raphson Method).

The results of our simulation study show that the maximum likelihood estimators have good properties and are consistent and efficient.

Overall, the maximum likelihood method is a powerful and flexible tool for parameter estimation in the Weibull distribution. It can be applied to a wide range of applications in engineering and reliability analysis. Further research can explore the use of alternative estimation methods and investigate the impact of model misspecification on the maximum likelihood estimators.

CONTRIBUTIONS

1) **Vidya Sagar - Code , PPT , Mathematical Computations**
2) **Shaik Vazeer Ahamed - Literature Review , dataset and Report**
3) **Manoj Arun Kumar - Report**

REFERENCES-

1) Parameter estimation of the Weibull Distribution; Comparison of the Least-Squares Method and the Maximum Likelihood estimation
https://ijaers.com/uploads/issue_files/21IJAERS-09202130-Parameter.pdf

2) Weibull Distribution for Estimating the Parameters
https://www.intechopen.com/chapters/17938

3) A Comparison Between the Bayesian and the Classical Estimators of Weibull Distribution
https://www.researchgate.net/publication/313402955_A_Comparison_Between_the_Bayesian_and_the_Classical_Estimators_of_Weibull_Distribution.

4) Maximum Likelihood Estimation: Numerical Solution for Weibull Distribution
https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwiFvaHAror-AhXEUGwGHcg5CaQQFnoECAwQAQ&url=https%3A%2F%2Fzenodo.org%2Frecord%2F3986931%2Ffiles%2FMLE_Weibull.pdf&usg=AOvVaw0h4KNZB4V3xzLdgdJPAZSr

5) Parameter Estimation for the Two-Parameter Weibull Distribution
https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=3508&context=etd

6) https://search.r-project.org/CRAN/refmans/EnvStats/html/eweibull.html