NOVEL ESTIMATORS OF SOFTWARE RELIABILITY FOR FINITE FAILURES CATEGORY MODELS

Thesis

Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

B. ROOPASHRI TANTRI

DEPARTMENT OF MATHEMATICAL & COMPUTATIONAL SCIENCES

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA

SURATHKAL, MANGALORE - 575025

February 2020

Dedicated to

My beloved parents, all my family members
&
my teachers

DECLARATION

I hereby *declare* that the Research Thesis entitled **NOVEL ESTIMATORS OF SOFTWARE RELIABILTY FOR FINITE FAILURES CATEGORY MODELS** which is being submitted to the **National Institute of Technology Karnataka, Surathkal** in partial fulfillment of the requirements for the award of the Degree of **Doctor of Philosophy** in **Mathematical and Computational Sciences** is a *bonafide report of the research work carried out by me*. The material contained in this Research Thesis has not been submitted to any University or Institution for the award of any degree.

B. Roopashri Tantri

Reg. No.: 112022 MA11P01

Department of Mathematical and Computational Sciences

Place: NITK, Surathkal.

Date: February , 2020

CERTIFICATE

This is to *certify* that the Research Thesis entitled **NOVEL ESTIMATORS OF SOFTWARE RELIABILITY FOR FINITE FAILURES CATEGORY MODELS** submitted by **B. ROOPASHRI TANTRI**, (Reg. No.: 112022 MA11P01) as the record of the research work carried out by her, is *accepted as the Research Thesis submission* in partial fulfillment of the requirements for the award of degree of **Doctor of Philosophy**.

(Dr. Murulidhar N. N.)
Research Supervisor

Chairman - DRPC

ACKNOWLEDGEMENT

I would like to take this opportunity to thank all those people who supported me throughout my research work. With a few words, I try to express my gratitude and heartfelt thanks to these people.

One of the most important persons in the journey of my research work is, my research guide, Dr. Murulidhar N. N. A kind hearted person, he would always encourage, guide and help me in every stage of my research work. He would always motivate and cheer me whenever I lost spirit during some delays in the progress of my work. I remain grateful to my guide for all his support and guidance.

I extend my thanks to RPAC members, Prof. G. Ram Mohana Reddy, Dept. of I.T. and Dr. Sam Johnson, Dept. of M.A.C.S., for their valuable suggestions, which helped me to improve my presentations.

I also remain grateful to NITK Surathkal for providing me an opportunity to carryout this work.

I am grateful to the Head, Department of MACS for providing me a very nice working environment and all the facilities in the department. I also extend my thanks to all the teaching and non-teaching staff members of the Department of M.A.C.S., who helped and guided me in one or the other way. I also thank all the research scholars of the department for helping me during my research work. I must thank all my colleagues of NCET, for their support and encouragement throughout my research work.

My family stood by me, throughout the ups and downs in my research tenure. Without the support of my mother, Mrs. Varija Tantri and my father Shri. B. V. Tantri, this work would not have been possible. I ever remain grateful to them throughout my life, for the care they extended to me and my family during this journey. I also remain grateful to my brother and sister for their support. I extend my heartfelt thanks to another important support during my entire research work, my uncle Shri. P. B. K. Murthy. He always made my stay very comfortable when I stayed with him especially during the last stages of my research work.

This acknowledgement will remain incomplete if I don't mention the names of my beloved husband, Dr. Sathya Prasanna and my sweet son Sudhanva Bhat. My husband would always support me, encourage me and guide me at every stage of my research work. My son would always bring back smile on my face, whenever I lost it due to the work pressure. Without these two persons, I would not have done my research work. I thank them a lot.

Place: NITK, Surathkal B. Roopashri Tantri

Date: February , 2020

ABSTRACT

The most important characteristic of the software product is its quality. One such important measure of the quality of the software is its reliability, which is the probability of failure-free operation of a computer program in a specified environment for a specified period of time. Estimating this software reliability enables the software developers to decide whether or not the user requirements are met. It also enables the users of the software to decide whether or not to accept the software. Thus, there is a strong need for estimating the reliability of the software. Software reliability models, with certain failure time distributions are used to estimate this reliability. Software reliability models are classified based on many attributes. One such classification is based on the number of failures. Depending on the number of failures, the software reliability models have been classified into two categories: (i) finite failures category models, where the number of failures is assumed to be finite and (ii) infinite failures category models, where the number of failures is assumed to be infinite. Finite failures category models are further classified into four classes, depending on the distribution of the failure times, namely, (i) Exponential class models, (ii) Weibull class models, (iii) Gamma class models and (iv) Pareto class models. Herein, the finite failures category models are considered and the reliability are estimated for the above four classes of models using the methods of Maximum Likelihood Estimation and Minimum Variance Unbiased Estimation. Further, the bias if any, present in the Maximum Likelihood Estimators (MLEs) are found using the Minimum Variance Unbiased Estimators (MVUEs). The MLEs are then improved by removing the bias present in them, thus getting the Improved Estimators of reliability. Several sample failure time data have been used to obtain these estimators, namely, MLE, MVUE and the Improved Estimators. The three estimators are then compared through the properties satisfied by these estimators. It is found that the Improved Estimator possesses most of the desirable properties of good estimators for all finite failures category models, which indicates that the Improved Estimator is most efficient and accurate as compared to MLE and MVUE. Hence, it is concluded that the software reliability can be estimated more accurately using the Improved Estimator, for any finite failures category software reliability model.

Keywords : Bias, Blackwellization, Coefficient of variation, Estimation, Exponential class models, Gamma class models, Improved Estimator, Method of Maximum Likelihood Estimation, Method of Minimum Variance Unbiased Estimation, Pareto class models, Software reliability, Software reliability models, Weibull class models.

Contents

List of Figures

List of Tables

Nomenclature

$R(\cdot)$	Reliability function
$f(\cdot), k(\cdot), h(\cdot)$	Probability density function (pdf)
L	Likelihood function
$E(\cdot)$	Expectation
$g(\cdot, \cdot)$	Joint pdf
$f(\cdot\|\cdot)$	Conditional pdf
$E(\cdot\|\cdot)$	Conditional expectation
$CV(\cdot)$	Coefficient of Variation
$\sim$	Distributed as / Follows
$\mathscr{E}(\alpha)$	Exponential distribution with parameter α
$W(\alpha, \beta)$	Weibull distribution with parameters α and β
$G(\alpha, \beta)$	Gamma distribution with parameters α and β
$P(\alpha, \beta)$	Pareto distribution with parameters α and β
$\Gamma(\cdot)$	Gamma function
$\hat{R}(\cdot)$	Maximum Likelihood Estimator of $R(\cdot)$
$\tilde{R}(\cdot)$	Minimum Variance Unbiased Estimator of $R(\cdot)$
$\breve{R}(\cdot)$	Improved Estimator of $R(\cdot)$

Chapter 1

INTRODUCTION

With the increasing role of software in every field, concern has grown over the quality of the software products. Any failure in the software product may not only result in economic damage, but also loss of life. Thus, software quality is a major concern of all software developers. The most important quality characteristic of the software product is its reliability. Software reliability is the probability of failure free operation of a computer program in a specified environment for a specified period of time (Trivedi (2012)). The failures are random in nature and their behavior with time are described by software reliability models (Musa et al. (1991)). Over 225 models have been developed since early 1970s. Models are classified based on many attributes. Two basic types of software reliability models are prediction models and estimation models. Prediction models are derived from actual historical data from real software projects. These models predict reliability at some future time. Estimation models use failure data from testing to forecast the failure rate or Mean Time Between Failures. Depending on the types of data, models fall into two basic classes: (i) Failures per time period (ii) Time between failures. Models are also classified based on failure history. The different categories here are: Time between failures model, Fault count models, Failure rate models, Baysian models etc. Musa and Okumoto classified models in terms of 5 different attributes. They are:
1. Time domain: Wall clock versus execution time
2. Category: The total number of failures experienced by time t.
3. Class (Finite failures category only): Functional form of failure intensity expressed in terms of time.
4. Family (Infinite failures category models): Functional form of failure intensity function expressed in terms of the expected number of failures experienced.
Based on the number of failures, the models are classified into two major categories:
(i) Finite failures category models, where the number of failures is assumed to be finite.
(ii) Infinite failures category models, where number of failures is assumed to be infinite.

The scope of this research is limited to finite failures category models only.
There are four general ways of characterizing failure occurences in time:
1. Time of failure
2. Time interval between failures
3. Cumulative failures experienced up to a given time
4. Failures experienced in a time interval.
Depending on the distribution of the time of failure, finite failures category models are further classified into the following four classes of models (Musa et al. (1991)):

Exponential class models: In this class of models, the failure time T is assumed to follow exponential distribution with parameter Φ, the failure rate. Thus, the probability density function of T is given by
$f(t) = \Phi e^{-\Phi t}, \quad t > 0.$

Weibull class models: This class consists of models with failure time distribution as Weibull with shape parameter β and scale parameter Φ, the failure rate. Thus, if T denotes the failure time, then, its probability density function is given by
$f(t) = \Phi\beta t^{\beta-1} e^{-\Phi t^{\beta}}, \quad t > 0.$

Gamma class models: The model of Yamada, Ohba and Osaki (Yamada et al. (1983)) assumes that the time to failure T follows gamma distribution with shape parameter 2 and scale parameter Φ, the failure rate. Thus, the probability density function of T is given by
$f(t) = \Phi^2 t e^{-\Phi t}, \quad t > 0.$

Pareto class models: This class of models have Pareto failure time distribution with shape parameter α and scale parameter β, which denotes the failure rate. Thus, if T denotes the failure time, then its probability density function is given by
$f(t) = \frac{\alpha}{\beta}\left(1 + \frac{t}{\beta}\right)^{-\alpha-1}, \quad t > 0.$
This model is due to Littlewood (Littlewood (1981)).

Each of these classes of models are again classified into different types of models, depending upon the distribution of the number of failures. Model classification scheme for finite failures category models, proposed by Musa and Okumoto (Musa et al. (1991)) is shown in Table 1.0.1.

Table 1.0.1 Musa model classification scheme - finite failures category models

Class	Type		
	Poisson	Binomial	Others
Exponential	Musa (1975) Moranda (1975) Schneidewind (1975) Goel-Okumoto (1979)	Jelinski-Moranda (1972) Shooman (1972)	Goel-Okumoto (1978) Musa (1979) Keiller-Littlewood (1983)
Weibull		Wagoner (1973) Schick-Wolverton (1973)	
Gamma	Yamada- Ohba- Osaki (1983)		
Pareto		Littlewood (1981)	

Eventhough the models in the finite failures category are old, many of them are still in use and research is still going on related to these models (some of them being Turk et al. (2016), Ledoux (2002), Ledoux (2002), Xu and Yao (2016), Nagar and Thankachan (2012), Lavanya et al. (2017), Turk and Alsolami (2016), Kantham and Rao (2009)). Herein, since the focus is on estimating the software reliability, which needs distribution of time of failures, models based on other time measures are not considered. Further, the statistical techniques used herein, require only the distribution of failure times and hence the focus is on finite failures category, wherein, the distributions of the times of failures are specified.

1.1 ESTIMATION

The procedure of finding the value of the unknown parameter of the given distribution by using statistical technique, is called estimation. The value of the parameter so obtained is called the estimator. To obtain this estimator, a sample $(X_1, X_2, \ldots, X_n)$ of size n is taken from the given distribution. Any function of this sample is called a statistic. A particular value of the estimator, obtained using the sample observations is called the estimate.

Terminologies and Definitions:

Conditional pdf: The conditional pdf of the random variable X given the random variable Y, denoted by $f(x|y)$ is defined as $f(x|y) = \dfrac{g(x,y)}{h(y)}$, where $g(x,y)$ is the joint pdf of the random variable (X,Y) and $h(y)$ is the marginal pdf of the random variable Y. If the random variables X and Y are independent, then, the joint pdf of (X,Y) is the product of the marginal pdfs of X and Y. i.e. $g(x,y) = f(x) \cdot h(y)$.

Likelihood Function: Let (X_1, X_2, X_3,$\ldots X_n$) be a random sample of size n from a population with probability density function $f(x;\theta)$ (or $f(x)$), where θ is the parameter. Then, the likelihood function of this sample, denoted by L or $L(\theta)$, is given by

$$L = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} f(x_i).$$

Expectation: Let X be a continuous random variable with pdf $f(x)$. Then, the expectation or the expected value of X or the mean of X, denoted by $E(X)$ (or μ) is defined as $E(X) = \int\limits_{-\infty}^{\infty} xf(x)dx$.

If $q(x)$ is any function of this random variable X, then, its expected value is defined as

$$E(q(X)) = \int\limits_{-\infty}^{\infty} q(x)f(x)dx.$$

Variance: The variance of the random variable X, denoted by $V(X)$ (or σ^2) is a measure of the dispersion among its values from the mean and is defined as $V(X) = E(X - E(X))^2 = E(X)^2 - (E(X))^2$. The positive square root of this variance is called the standard deviation, denoted by σ.

Conditional Expectation: Conditional expectation of the random variable X given the random variable Y, denoted by $E(X|Y)$ is defined as $E(X|Y) = \int\limits_{-\infty}^{\infty} xf(x|y)dx$, where $f(x|y)$ is the conditional pdf of X given Y.

Reliability: Let the random variable T denote the time to failure of individual faults, with pdf $f(t)$. Then, the reliability function of the software, denoted by $R(t)$, is the probability of its failure free operation upto time t and is given by $R(t) = P(T > t)$.

Properties of Estimators:

The estimators are expected to possess the following statistical properties (Gupta and Kapoor (1996)):

(i) Unbiasedness: An estimator $\dot{\theta}$, based on a sample of size n from a distribution with parameter θ, is said to be unbiased for θ, if $E(\dot{\theta}) = \theta$. Otherwise, the estimator is said to be biased and its bias is given by Bias = $E(\dot{\theta}) - \theta$.

(ii) Consistency: An estimator $\dot{\theta}$, based on a sample of size n from a distribution with parameter θ, is said to be consistent for θ, if $\dot{\theta}$ converges in probability to θ. i.e. if

$P(\mid \dot{\theta} - \theta \mid < \epsilon) \longrightarrow 1$ as $n \longrightarrow, \infty$ for every $\epsilon > 0$.

(iii) Efficiency: If $\dot{\theta}_1$ and $\dot{\theta}_2$ are any two estimators of a parameter θ, with variances V_1 and V_2 respectively, then, the efficiency of $\dot{\theta}_2$ relative to $\dot{\theta}_1$ is defined as Eff$(\dot{\theta}_2) = \dfrac{V_1}{V_2}$. Also, $\dot{\theta}_2$ is said to be more efficient than $\dot{\theta}_1$, if Eff$(\dot{\theta}_2) > 1$.

(iv) Sufficiency: An estimator $\dot{\theta}$ is said to be a sufficient estimator for the parameter θ, if it contains all the information in the sample regarding that parameter. i.e. an estimator $\dot{\theta}$ is sufficient for θ, if the conditional distribution of the sample, given $\dot{\theta}$, is independent of θ.

Factorization theorem: An estimator $T(x)$ based on a sample of size n is sufficient for the parameter θ, iff the likelihood function L can be expressed in the form
$L = g_\theta(T(x)) \cdot h(x)$, where $g_\theta(T(x))$ depends on θ and x only through the value of $T(x)$ and $h(x)$ is independent of θ.

Complete estimator: Let $\dot{\theta}$ be an estimator and let T be any function of $\dot{\theta}$. Then, $\dot{\theta}$ is said to be the complete estimator, if $E(T) = 0$ implies $T = 0$. Or in other words, an estimator $\dot{\theta}$ is complete, if and only if the only unbiased estimator of zero that is a function of $\dot{\theta}$ is the statistic that is identically zero with probability one. More precisely, an estimator $\dot{\theta}$ that is not complete will have at least one value of the parameter, such that the function T of $\dot{\theta}$ is not almost surely zero for that value and yet its expected value is zero for all values of the parameter. Or equivalently, $\dot{\theta}$ is a complete estimator, if any unbiased estimator of zero based on $\dot{\theta}$ is identically zero.

Complete Sufficient Estimator: A sufficient estimator which is also complete is called the complete sufficient estimator.

Methods of Estimation:
The performance of the software can be assessed by several measures such as - the failure rate, the failure intensity function, the Mean Time To Failure (MTTF) etc. Software reliability is one such measure. Estimation of software reliability provides a better knowledge of the reliability of the software. Several methods of estimation exist. Two such important methods of estimation are - the method of maximum likelihood estimation and the method of minimum variance unbiased estimation. The estimators obtained using the above two methods are respectively called Maximum Likelihood Estimator (MLE) and Minimum Variance Unbiased Estimator (MVUE).

Method of Maximum Likelihood Estimation (Gupta and Kapoor (1996)): Let L denote the likelihood function of the given sample of size n from a distribution with pdf $f(x, \theta)$. Then, the MLE of the parameter θ is that value of the parameter, which maximizes this likelihood function L. Thus, the Maximum Lilkelihood Estimator (MLE) of θ, denoted by $\hat{\theta}$ is the solution of $\dfrac{\partial(lnL)}{\partial\theta} = 0$ with $\dfrac{\partial^2(lnL)}{\partial\theta^2} < 0$.

Method of Minimum Variance Unbiased Estimation (Gupta and Kapoor (1996)): If a statistic M based on a sample of size n is such that M
(i) is unbiased for the parameter θ and
(ii) has the smallest variance among the class of all unbiased estimators of θ,
then, M is called the Minimum Variance Unbiased Estimator (MVUE) of θ.
More precisely, M is MVUE of θ, if
(i) $E(M) = \theta$ and (ii) $V(M) \leq V(M')$ where V denotes the variance amd M' is any other unbiased estimator of θ. Such an MVUE is always unique. This MVUE can be found using a result stated below:
If T is a complete sufficient estimator of θ and if $g(T)$ is a function of T, which is unbiased for θ, then $g(T)$ is MVUE of θ.
As an example, let (X_1, X_2, X_3,$\ldots X_n$) be a random sample of size n from the Poisson distribution $P(\theta)$. Then, the likelihood function is

$$L = \prod_{i=1}^{n} f(x_i, \theta) = \frac{e^{-n\theta}\theta^{\sum\limits_{i=1}^{n} x_i}}{\prod\limits_{i=1}^{n} x_i!}$$

Thus, $\sum\limits_{i=1}^{n} x_i$ is complete sufficient for θ. Now, consider a function of this sufficient statistic as $\bar{X} = \dfrac{\sum\limits_{i=1}^{n} x_i}{n}$. Then, obviously, $E(\bar{X}) = \dfrac{n\theta}{n} = \theta$. Thus, $\bar{X}$ is unbiased for θ. Hence from the stated result, $\bar{X}$ is the MVUE of θ.
The MVUE can also be found by finding an unbiased estimator of the parameter and then improving upon it by defining a function of the sufficient statistic. This procedure is called Blackwellization and is explained below:

Blackwellization (Gupta and Kapoor (1996)): Let U be an unbiased estimator of the parameter θ and let S be the complete sufficient estimator of θ. Then, The Minimum Variance Unbiased Estimator (MVUE) of θ, denoted by $\tilde{\theta}$ is defined as the conditional expectation of U given S. i.e. $\tilde{\theta} = E(U|S)$.

Theorem 1 (Gupta and Kapoor (1996)): If T is a complete sufficient statistic for the parameter θ and if $h(x_1, x_2, \ldots, x_n)$ is any unbiased estimator of $r(\theta)$, then, $E(h(x_1, x_2, \ldots, x_n)|T)$ is the MVUE of $r(\theta)$.

Invariance property of MLEs (Gupta and Kapoor (1996)): If $\hat{\theta}$ is the MLE of θ and if $W(\theta)$ is a one to one function of θ, then, the MLE of $W(\theta)$, denoted by $\hat{W}(\theta)$ is obtained as $W(\hat{\theta})$.

Comparison of estimators: When there are several estimators, the most efficient among them is chosen by comparing the properties satisfied by these estimators (unbiasedness, consistency, sufficiency, efficiency). An estimator that satisfies maximum number of these properties is declared as the best estimator. It has been established that MLEs are always consistent and sufficient, while they need not be unbiased. MVUEs are always unbiased and sufficient, while they need not be consistent ((Gupta and Kapoor (1996))). However, the efficiency property can be checked by using some measures of dispersion, like variance, coefficient of variation, quartile coefficient of dispersion etc. The estimate with least value of this measure of dispersion is considered as the efficient estimator. Further, since the current problem under study deals with finite failures only, the consistency property, which is an asymptotic property, is not applicable for comparing the estimators.

Coefficient of variation: When comparison between biased and unbiased estimators are to be done, the variances are not considered for comparison; instead, the coefficient of variation, which considers both the mean and the standard deviation, is considered as a measure of dispersion among the estimators to be compared. The coefficient of variation is defined as the ratio of the standard deviation to the mean. Thus, if σ and μ denote the standard deviation and mean of an estimator say $\hat{\theta}$ respectively, then, the coefficient of variation of $\hat{\theta}$, denoted by $CV(\hat{\theta})$ is given by $CV(\hat{\theta}) = \dfrac{\sigma}{\mu}$.

Quartile coefficient of dispersion: Another measure of comparison between the estimators is the quartile coefficient of dispersion. It is computed using the first (Q_1) and third (Q_3) quartiles for each data set. The first quartile marks where 25 percent of the data is below or to the left of it, while the third quartile is where 75 percent of the data lies below this point. The quartile coefficient of dispersion, denoted by QD is computed as $QD = \dfrac{Q_3 - Q_1}{Q_3 + Q_1}$.

The desirable properties satisfied by MLE and MVUE for the current research work are provided in Table 1.1.1.

Table 1.1.1 Properties satisfied by MLE and MVUE

	Unbiased	Sufficient	Efficient
MLE	Not always	Yes	*
MVUE	Yes	Yes	*

(*) - To be checked by comparing the measure of dispersion.

1.2 RELIABILITY ESTIMATION

Software is an essential component of all computer systems. These systems depend on the reliable operation of the software components. Reliability is a measure of how well the software provides the services expected by the customer.

Software Reliability Engineering process consists of the following stages:
1. Defining reliability objective: Usually, the user wants the software to sustain till some specified time, which is set as reliability objective. Using the reliability curve obtained by estimation, it can be checked whether the objective as set by the user is achieved or not.
2. Modeling expected system usage: It is necessary to model how users employ the software. i.e., the environment, type of installation etc.
3. Prepare test cases: According to usage model, test cases are selected randomly.
4. Execute test, collect failure data: Once the test cases are selected, they are executed. The failures are detected and the data pertaining to these failures, such as the time of failure or time between failures are obtained.
5. Perform reliability certification and monitor reliability growth: Based on the failure data, the reliability can be estimated.

Several questions that may arise during the testing of the software can be answered by estimating the reliability of the software. Estimation of software reliability helps the software developers and users to have an idea of its durability in the long run. Reliability estimates help software developers to ensure that the user requirements are met and also to decide when to release the software. On the other hand, the reliability estimates help the users of the software in deciding whether or not to accept the software. Thus, there is a strong need for estimating the reliability of the software. Herein, it is intended to obtain the estimators of software reliability using the methods of maximum likelihood estimation and minimum variance unbiased estimation. Since MVUE are always unbiased, while MLEs need not be, it is intended to obtain the bias present in MLE, if any, by using the MVUE and to improve the MLE by removing this bias. Let this estimator be called an Improved Estimator of the reliability. It is intended to compare the three estimators, namely, the MLE, the MVUE and the Improved Estimator of the reliability and choose the best among them. A few software failure time data are used for this purpose. The properties as stated in Table 1.1.1 are considered for the comparison, for all finite failures category models, using the wide range of software failure time data sets. It has been found that the Improved Estimator satisfies maximum number of these desirable properties, as compared to that of MLE and MVUE for all finite failures category models. Thus, the Improved Estimator gives more accurate value of reliability and hence is preferred over MLE and MVUE of reliability.

Chapter two deals with literature review, outcome of literature review, scope for research, research objectives, problem statement and methodology. Chapters three through six deal with obtaining the MLE, the MVUE and the Improved Estimator of software reliability for Exponential, Weibull, Gamma and Pareto class models respectively and comparing these estimators using failure data sets and choosing the best estimator. The report concludes with conclusion in chapter seven and scope for improvement thereby.

Chapter 2

LITERATURE REVIEW

A lot of research has been carried out in the area of estimation of the parameters and other measures of reliability for different software reliability models, in literature. Some of such research work related to estimation of the parameters and other estimates of reliability measures for different classes of finite failures models, considered in this research work are provide below:

P. A. Keiller et. al (Keiller et al. (1983)) have proposed different ways of analyzing the quality of software reliability predictions using software failure data sets. The work was carried out on different models, including the Jelinski-Moranda model of exponential class. But, the results of applying maximum likelihood estimation on this model showed extremely poor prediction.
H. Joe and N. Reid (Joe and Reid (1985)) have formulated Jelinski-Moranda model and Littlewood model in terms of failure times and obtained maximum likelihood estimators of parameters of Jelinski-Moranda model. Also, an improved estimator for the initial number of faults in the software was obtained.
Jun Hishitani et. al (Hishitani et al. (1990)) have adopted mean time between software failures as a reliability assessment measure and proposed a method of software reliability assessment by considering the mean time between failures for exponential and delayed S-shaped software reliability growth models based on Non Homogeneous Poisson Process (NHPP).
Bev Littlewood (Littlewood (1991)) mentions that there are quite stringent requirements for the testing regime in which data is collected. Only if these requirements are satisfied, then it is possible to obtain accurate reliability estimates.
Chris Dale (Dale (1991)) has explained actual and potential uses of statistical methods in the assessment of software reliability. The work identifies reasons for software reliability assessment and discusses the role of statistics.

Raymond Jacoby and Yoshihiro Tohma (Jacoby and Tohma (1991)) have presented the Hyper-Geometric Distribution Software Reliability Growth Model (HGDM) for estimating of the number of software faults at the beginning of the test-and-debugging phase. The method of least squares was used to estimate the parameters.
Hossain and Dahiya (Hossain and Dahiya (1993)) have obtained a necessary and sufficient condition for the likelihood estimates of the parameters of Goel-Okumoto (G-O) model to be finite, positive and unique. A modification of the G-O model was suggested and the performance measures of the new model were discussed.
Mark and Anne (Yang and Chao (1995)) have obtained reliability estimation and stopping rules for software testing, based on repeated appearances of bugs. The rules worked well on Musa execution time models (which belong to the exponential class) and Logarithmic Poisson models, which works well for small bugs, i.e, bugs with very low occurence rates.
Singpurwalla and Soyer (Singpurwalla and Soyer (1996)) have described briefly, several well known probability models for assessing the reliability of the software. Statistical methods were used to estimate the parameters of the model.
Qureshi and Daniel (Quereshi and Jeske (1997)) have introduced the concept of proxy failure times and showed how to simulate proxy failure times. Jelinski-Moranda model was considered and a graphical diagnosis for testing goodness-of-fit was done to show the improvement in it.
Balwant Singh et. al (Singh et al. (1997)) have estimated the software reliability using inverse sampling and also showed how these estimates can be used to determine the testing time and whether or not the software is acceptable.
In the thesis entitled "Accurate Software Reliability Estimation", Jason Allen Denton (Denton (1999)) has examined the impact of the parameter estimation technique on model accuracy and showed that the maximum likelihood method provides better estimates than the least squares method. The method is applicable to all types of software reliability models.
Srinivasan et. al (Ramani et al. (2000)) have developed architecture-based approaches to assess the reliability and performance of the systems. The work presented high-level design of a Software Reliability Estimation and Prediction Tool (SREPT), to assist in the evaluation of software reliability during all phases of software life-cycle.
Frank Padberg (Padberg (2001)) has presented a fast and exact algorithm to compute the maximum likelihood estimates for the number of faults initially contained in software, using hyper geometric software reliability model. The key idea leading to the algorithm was to study the growth quotient of the likelihood function instead of the likelihood function itself.

Katerina and Trivedi (Goseva-Popstojanova and Trivedi (2001)) have developed architecture based approach to reliability assessment of software systems. The work also describes how it can be used to examine software behavior from beginning to implementation stage. Different models, including the Shooman model were considered.
Frenkel et. al (Frenkel et al. (2003)) have considered a Non-Homogeneous Poisson Process (NHPP) with intensity function $\lambda(t)$ to estimate the parameters by the maximum likelihood procedure and showed that there is one and the only one solution to the maximum likelihood equation for the log-linear form of $\lambda(t)$.
Hiroyuki et. al (Okamura et al. (2003)) have proposed an estimation method for the model parameters of a software reliability model based on the EM (Expectation-Maximization) principle. The method was applied to models like exponential, Rayleigh, Pareto etc. Real time data were used to compare the method with other methods.
Guen et. al (Guen et al. (2004)) have developed a testing technique called statistical usage testing (SUT) to estimate the reliability of the software. They have also presented new approaches to estimate the reliability from Markov chains. The reliability estimation has been implemented in a tool for SUT, called MaTeLo.
Kuo and Yang (Kuo and Yang (1995)) has used Baysian approach to obtain the estimate the reliability of Jelinski-Moranda and Littlewood model. In addition, prediction of future failure times and future reliabilities were also examined.
Hiroyuki Okamura et. al (Okamura et al. (2007)) have proposed a parameter estimation method which combines the EM (Expectation-Maximization) algorithm and a heuristic solution method, as an effective parameter estimation method for the generalized gamma software reliability model.
In the research work carried out by Raj Kiran and Ravi (Kiran and Ravi (2008)), an ensemble-based approach, viz, Back Propagation trained Neural Network (BPNN) is proposed in predicting software reliability. Statistical techniques like multiple linear regression were also used. Experiments were performed on software reliability data obtained from literature.
Khalaf and Mustafa (Khalaf and Mustafa (2009)) in their work, explored a model that can be used for software reliability prediction, using the fuzzy logic technique. Data sets were used in predicting reliability.
RajPal and Kapil Sharma (Garg and Sharma (2010)) have presented a computational methodology based on matrix operations for a computer based solution to the problem of performance analysis of Software Reliability Models (SRMs). A set of seven comparison criteria have been formulated to rank various non-homogeneous Poisson process software reliability models proposed during the past 30 years to estimate software reliability measures such as the number of remaining faults, software failure

rate and software reliability.
Aasia Quyoum et. al (Quyoum et al. (2010)) have focused on using software engineering principles in the software development and maintenance so that reliability of software could be improved. They suggested that sufficient testing and proper maintenance could improve software reliability to a great extent.
Krishna Mohan et. al (Mohan et al. (2010)) have used an effective sampling method of testing to measure software reliability. They have used Black box testing to obtain failure data and white box testing to quantify the reliability.
Sinda Rebello and Neeraj Kumar Goyal (Rebello and Goyal (2010)) have presented a methodology for assessing the reliability and safety of a software based on extended Failure Modes and Effects Analysis (FMEA) approach.
Sagelietti et. al (Sohnlein et al. (2010)) have assessed software reliability based on evaluation of operational experience. The work also elaborates the possibility of assessing software reliability at system level by the combination of component-specific software reliability estimates.
Latha and Lilly (Shanmugam and Florence (2012)) have compared the parameter estimation procedures of maximum likelihood estimation and least square estimation method with other methods, viz, Expectation-Maximization principle, Genetic Algorithm and Particle Swarm Optimization Algorithms. The parameter estimation method based on Ant Colony Optimization algorithm was proposed to overcome the drawbacks of all these methods.
Mandeep Kaur et. al (Kaur et al. (2013)) have discussed the role of reliability metrics in improving the software reliability. Mean Time To Failure (MTTF) was considered as the most commonly used reliability metric.
Chris Bambey et. al (Guure et al. (2013)) have proposed Baysian parameter and reliability estimate of two parameter Weibull failure time distribution.
The work done by Gurpreet Kaur and Kailash Bahl (Kaur and Bahl (2014)) discusses about software reliability, metrics and reliability improvement using Agile process.
In the thesis entitled "Reliability Estimation of Open Source Software based Computational Systems", Shelbi Joseph (Joseph (2014)) presents a model for evaluating the reliability of systems based on Open Source Software. The method involves in identifying the failure data for hardware components, software components and building a model based on it, to predict the reliability.
In their research paper, Gee Yong Park and Seung Cheol Jang (Park and Jang (2014)) have used two types of modeling schemes of software reliability growth models to estimate and predict the number of software defects based on software failure data. The Bayesian statistical inference is employed to estimate the model parameters.

Taehyoun et. al (Kim et al. (2015)) have proposed an effective approach to estimate the parameters of Software Reliability Growth Model (SRGM) using a Real-valued Genetic Algorithm (RGA). The advantage of using this method over other methods like the method of maximum likelihood estimation and the method of least squares is that, it is free from constraints on the parameter estimation of SRGM. Two real-valued genetic operators, heuristic crossover and non-uniform mutation, were applied to improve the accuracy and performance of the parameter estimation of SRGM. Eight real world datasets were considered for comparing the method with other methods.
Sanjay Kumar Chauhan et. al (Chauhan et al. (2016)) have used software which is a collection of bits, to estimate software reliability in terms of time analysis. Performance of the system was justified by analyzing bits and reliability was estimated in a better way. Some statistical concepts were also used to examine this.
Parveen Sehgal and Meenal (Sehgal and Meenal (2016)) have used artificial intelligence based techniques for estimating the reliability. The framework was based on the use of artificial neural networks for estimating the software reliability based upon historical data sets.
Subburaj Ramasamy and Indhurani Lakshmanan (Ramasamy and Lakshmanan (2017)) have used Machine Learning Approach for Software Reliability Growth Modeling to describe software failure data. They have also compared the performance of the proposed model with existing model using practical software failure data sets.

2.1 OUTCOME OF LITERATURE REVIEW

As the need for a reliable software grows, there is a need for having methods that can give more accurate values of reliability. Keeping this in mind, many researchers have come up with various ideas of measuring reliability. The assessment of reliability measurements considers various techniques involving estimation of parameters as well as estimating the reliability directly. Estimation of reliability is often done to ensure that the requirements of the user are met. Reliability estimation tools have been obtained by many researchers. Statistical techniques like Baysian approach have also been used. It is also found that most of the other techniques involve either the method of MLE or the method of least squares. Methods like statistical usage technique, fuzzy logic technique, artificial intelligence techniques, neural network techniques etc have also been used for different types of models. Confidence intervals have also been obtained for the reliability of the given software. Many researchers have also developed different models for estimating the software reliability. Many researchers have used different measures of reliability such as mean time to failure, time of next failure, mean time between failures, failure rate, mean value function etc. Other methods like genetic algorithm, ant colony method etc are restricted only for particular types of models and not for all finite failures category models. Most of these deal with exponential, gamma and Weibull models. Also, these methods do not yield the assessment of reliability and are practically not applicable for all types of models. Thus, there is a strong need for estimating more accurate value of reliability for any given software product.

2.2 SCOPE FOR RESEARCH

There is a need for reliable software because of many reasons. The software failure may cause economic damage and even loss of life. As the need for reliable software grows, there is a need to have methods that can give more accurate values of reliability. Software developers are often concerned with the problem of when to release the software and to decide whether or not the user requirements are met. On the other hand, the users of the software are concerned with whether or not to accept the software. As solutions to these problems, researchers have come up with the idea of assessing the reliability of the software, which will help both the developers and the users of the software to answer the above questions. There are various ways of assessing the performance of the software. Some of such assessments are prediction and estimation of reliability, estimation of Mean Time To Failure, failure rate, failure intensity function, mean value function etc. Lot of work pertaining to assessment of these have been done in literature as detailed above. Estimation of measures such as Mean Time To Failure, failure rate, failure intensity function, mean value function etc have widely been used. Research is still going on, in obtaining the best solution to the problems faced by developers and users of the software in answering such questions. Another best alternative to solve this problem would be to estimate the reliability of the software, using the minimum possible information available. A better option in this case would be to go for some statistical methods of estimation, which require minimum information of some probability distributions. One such piece of information needed for estimating the reliability is the information regarding the failure time distribution. Hence there is a scope for using the probability distribution of failure times in estimating the reliability using statistical methods. The most commonly used methods are the method of maximum likelihood estimation and the method of least squares. It has been proved in literature, as mentioned above, that MLEs are preferred over least square estimators. Reliability can also be estimated using the estimated values of the parameters obtained using the method of MLE. However, such estimators need some desirable properties of estimators, to be satisfied by them. It has been known that

MLEs satisfy most of the properties of good estimators. But, most of the time, they are biased and are not efficient. Method of Minimum Variance Unbiased Estimation (MVUE), on the other hand always provides unbiased and sufficient estimators. Thus, there is a scope for finding the reliability estimates using the methods of MLE and MVUE and to improve the estimators. Since MVUE are always unbiased, the MVUEs can be used in finding the bias, if any, present in MLE. The MLE is then improved by removing its bias to obtain the Improved Estimator of reliability.

2.3 PROBLEM STATEMENT

The estimator of reliability is found using the methods of MLE and MVUE for all finite failures category software reliability models. An Improved Estimator is obtained using these two estimators and the best estimator is chosen from these three estimators, through the properties satisfied by these estimators.

2.4 RESEARCH OBJECTIVES

The objectives of the present work are:

(i) To obtain the MLE and MVUE of $R(t)$ for the following finite failures category software reliability models:

(a) Exponential class models.

(b) Weibull class models.

(c) Gamma class models.

(d) Pareto class models.

(ii) To find the bias, if any, in MLEs, by using MVUE, based on given sample failure data.

(iii) To remove this bias from MLE and obtain the Improved Estimator.

(iv) To compare the properties satisfied by the three estimators, viz, MLE, MVUE and Improved Estimators of $R(t)$ by using sample failure data and choose the best estimator of reliability, based on the number of properties satisfied by them, for all finite failures category software reliability models.

2.5 METHODOLOGY

(i) Find the expression for the reliability function $R(t)$ for each of the finite failures category software reliability models, using the probability density function of the failure time of that model.

(ii) Estimate the parameters and hence obtain MLE of reliability function $R(t)$.

(iii) Obtain MVUE of reliability for each of the models, using the method of Blackwellization.

(iv) Obtain the bias, if any, in MLE by using MVUE.

(v) Improve the MLE by reducing the bias and hence obtain the Improved Estimator of reliability.

(vi) Use sample failure data and obtain the estimated values of MLE, MVUE and Improved Estimator of $R(t)$.

(vii) Decide on the best estimator based on the properties satisfied by them, for all finite failures category software reliability models.

Chapter 3

EXPONENTIAL CLASS MODELS

The software reliability models, which belong to the exponential class, have the failure times (T), which are assumed to have Exponential distribution with probability density function (pdf), given by

$$f(t) = \Phi e^{-\Phi t}, \quad t > 0 \tag{3.0.1}$$

where the parameter Φ denotes the failure rate.

Many models in literature belong to this class. Work on one such model viz, the Jelinski- Moranda model, was done by P. A. Keiller et. al (Keiller et al. (1983)) about predicting software reliability, which showed extremely poor results, by using the method of maximum likelihood estimation. However, an effort in this area can be done by applying this method of MLE in estimating the reliability. Further, the same method was also used by H. Joe and N. Reid (Joe and Reid (1985)) to estimate the parameters of Jelinski-Moranda model, which belings to the exponential class. But the estimate was obtained by considering the time between failures distribution rather than failure time distributions. As an alternative, failure time distributions can be used in estimating the parameters and hence the reliability using the method of MLE.

Another exponential class model, viz., the Goel-Okumoto model was considered by Hossain and Dahiya (Hossain and Dahiya (1993)) and the method of MLE was considered in estimation of parameters, by using the distribution of time between failures. Jason Allen Denton (Denton (1999)) obtained and proved that MLEs provide better parameter estimates than least square estimates and are more accurate. The method was applicable to all types of software reliability models. Thus, it is applicable for all the four classes of models, considered in this research work.

However, these works focus on estimation of parameters, based on either time between failures or number of failures and do not focus on the estimation of reliability. Estimation of parameters may help in estimation of reliability, but limited to only MLEs, since invariance property is applicable for MLEs only. Further, it has also been proved that

MLEs provide better estimates than least squares.Thus, there is a strong need for looking into other statistical techniques of estimation in reliability. A better option in this regard would be to opt for minimum variance unbiased technique in estimation of reliability. The following sections focus on this area:

If T has an exponential distribution with parameter Φ, it is then denoted as $T \sim \mathscr{E}(\Phi)$. For this model, the reliability function at time t, denoted by$R(t)$ (as explained in Chapter 1), is obtained as

$$R(t) = P(T > t) = \int_{t}^{\infty} f(t)dt = \int_{t}^{\infty} \Phi e^{-\Phi t}dt = e^{-\Phi t} \tag{3.0.2}$$

Consider obtaining the estimates of this reliability using the methods of MLE and MVUE.

3.1 MLE OF $R(t)$

Since MLEs satisfy the invariance property, the MLE of $R(t)$, denoted by $\hat{R}(t)$ is obtained as

$$\hat{R}(t) = e^{-\hat{\Phi}t} \tag{3.1.1}$$

where $\hat{\Phi}$ denotes the MLE of Φ.

To find the MLE of Φ: Let $(T_1, T_2, \ldots T_n)$ be a sample of size n from exponential distribution with pdf as given in (3.0.1). Then, the likelihood function is given by

$$L = \prod_{i=1}^{n} f(t_i) = \Phi^n e^{-\Phi \sum\limits_{i=1}^{n} t_i} \tag{3.1.2}$$

Maximizing this likelihood function using the principle of differential calculus, the MLE of Φ, denoted by $\hat{\Phi}$, is obtained as the solution of $\frac{\partial \ln L}{\partial \Phi} = 0$, with $\frac{\partial^2 \ln L}{\partial \Phi^2} < 0$. This gives $\frac{n}{\Phi} - \sum\limits_{i=1}^{n} t_i = 0$, from which, the MLE of Φ is obtained as

$$\hat{\Phi} = \frac{n}{\sum\limits_{i=1}^{n} t_i} \tag{3.1.3}$$

Using (3.1.3) in (3.1.1), the MLE of $R(t)$ is obtained as

$$\hat{R}(t) = e^{-\left(\frac{nt}{\sum\limits_{i=1}^{n} t_i}\right)} \tag{3.1.4}$$

3.2 MVUE OF $R(t)$

To find the MVUE of $R(t)$, it is intended to obtain the unbiased estimator of $R(t)$ and the complete sufficient estimator of the parameter Φ.

Unbiased Estimator of $R(t)$: Define a function of the random variable T_1 as (Sinha and Kale (1980))

$$U(t_1) = \begin{cases} 1 & \text{if } t_1 > t \\ 0 & \text{otherwise} \end{cases}$$

Then, $E[U(t_1)] = 1.P(T_1 > t) + 0.P(T_1 \leq t) = P(T_1 > t) = R(t)$.
Hence, $U(t_1)$ is unbiased for $R(t)$.

Complete Sufficient Estimator of Φ: Applying the factorization theorem (as explained in Chapter 1) to the likelihood function given in (3.1.2), it can be seen that the likelihood function depends on Φ and t_i, only through the value of $\sum\limits_{i=1}^{n} t_i$. Thus, $\sum\limits_{i=1}^{n} t_i$ is the sufficient estimator of the parameter Φ.
Since each $T_i \sim \mathscr{E}(\Phi)$, $\sum\limits_{i=1}^{n} T_i \sim G(n, \Phi)$ (Section 1 of Appendix A). Further, by Result 3 of Appendix A, the estimator $\sum\limits_{i=1}^{n} t_i$ is also the complete statistic of Φ and hence, $\sum\limits_{i=1}^{n} t_i$ is the complete sufficient estimator of Φ.
Further, $U(t_1)$ is an unbiased estimator of $R(t)$ and $R(t)$ is a function of Φ, as given in (3.0.2). Hence, by Theorem 1 of Chapter 1, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = E(U(t_1)|\sum_{i=1}^{n} t_i) = \int_{t}^{\infty} f(t_1|\sum_{i=1}^{n} t_i)dt_1 \tag{3.2.1}$$

where $f(t_1|\sum\limits_{i=1}^{n} t_i)$ denotes the conditional pdf of T_1 given $\sum\limits_{i=1}^{n} T_i$ and is given by $f(t_1|\sum\limits_{i=1}^{n} t_i) = \dfrac{g(t_1, \sum\limits_{i=1}^{n} t_i)}{h(\sum\limits_{i=1}^{n} t_i)}$, where $g(t_1, \sum\limits_{i=1}^{n} t_i)$ denotes the joint pdf of $(T_1\ , \sum\limits_{i=1}^{n} T_i)$

and $h(\sum_{i=1}^{n} t_i)$ denotes the marginal pdf of $\sum_{i=1}^{n} T_i$.
Thus, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_{t}^{\infty} \frac{g(t_1, \sum_{i=1}^{n} t_i)}{h(\sum_{i=1}^{n} t_i)} dt_1 \tag{3.2.2}$$

Since each $T_i \sim \mathscr{E}(\Phi)$, $\sum_{i=1}^{n} T_i \sim G(n, \Phi)$ (Section 1 of Appendix A). Hence, the pdf of $\sum_{i=1}^{n} T_i$ is given by

$$h(\sum_{i=1}^{n} t_i) = \frac{\Phi^n}{\Gamma(n)} e^{-\Phi \sum_{i=1}^{n} t_i} (\sum_{i=1}^{n} t_i)^{n-1} \tag{3.2.3}$$

To find the pdf $g(t_1, \sum_{i=1}^{n} t_i)$, split the sample $(T_1, T_2, T_3, \ldots T_n)$ into two samples as - T_1 of size one and $(T_2, T_3, T_4, \ldots T_n)$ of size $(n-1)$.
Since T_1 and $\sum_{i=2}^{n} T_i$ are independent, the joint pdf of T_1 and $\sum_{i=2}^{n} T_i$ is obtained as

$$g(t_1, \sum_{i=2}^{n} t_i) = f(t_1) \cdot h(\sum_{i=2}^{n} t_i) \tag{3.2.4}$$

where $f(t_1)$ and $h(\sum_{i=2}^{n} t_i)$ denote the pdfs of T_1 and $\sum_{i=2}^{n} T_i$ respectively.
Since $T_1 \sim \mathscr{E}(\Phi)$, the pdf of T_1 is given by $f(t_1) = \Phi e^{-\Phi t_1}$.
Also, $\sum_{i=2}^{n} T_i \sim G(n-1, \Phi)$ and hence its pdf is given by
$h(\sum_{i=2}^{n} t_i) = \frac{\Phi^{n-1}}{\Gamma(n-1)} e^{-\Phi \sum_{i=2}^{n} t_i} (\sum_{i=2}^{n} t_i)^{n-2}$.
Hence, from (3.2.4), the joint pdf of T_1 and $\sum_{i=2}^{n} T_i$ is obtained as

$$g(t_1, \sum_{i=2}^{n} t_i) = \Phi e^{-\Phi t_1} \frac{\Phi^{n-1}}{\Gamma(n-1)} e^{-\Phi \sum_{i=2}^{n} t_i} (\sum_{i=2}^{n} t_i)^{n-2} = \frac{e^{-\Phi[t_1 + \sum_{i=2}^{n} t_i]} \Phi^n}{\Gamma(n-1)} (\sum_{i=2}^{n} t_i)^{n-2}.$$

Hence, $g(t_1, \sum_{i=2}^{n} t_i) = \frac{e^{-\Phi \sum_{i=1}^{n} t_i} \Phi^n}{\Gamma(n-1)} (\sum_{i=2}^{n} t_i)^{n-2}$.

Considering the transformation $\sum_{i=1}^{n} T_i = T_1 + \sum_{i=2}^{n} T_i$ and noting that the absolute value of the Jacobian of the inverse transformation is one (Appendix B), the joint pdf of (T_1 , $\sum_{i=1}^{n} T_i$) is obtained as

$$g(t_1, \sum_{i=1}^{n} t_i) = \frac{e^{-\Phi \sum_{i=1}^{n} t_i}}{\Gamma(n-1)} \Phi^n (\sum_{i=1}^{n} t_i - t_1)^{n-2} \tag{3.2.5}$$

Substituting (3.2.3) and (3.2.5) in (3.2.2), the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_{t}^{\infty} \frac{\frac{\Phi^n}{\Gamma(n-1)} e^{-\Phi \sum_{i=1}^{n} t_i} (\sum_{i=1}^{n} t_i - t_1)^{n-2}}{\frac{\Phi^n}{\Gamma(n)} e^{-\Phi \sum_{i=1}^{n} t_i} (\sum_{i=1}^{n} t_i)^{n-1}} dt_1$$

This reduces to

$$\tilde{R}(t) = \int_{t}^{\infty} \frac{\Gamma(n)}{\Gamma(n-1)} \left(\frac{\sum_{i=1}^{n} t_i - t_1}{\sum_{i=1}^{n} t_i} \right)^{n-2} \frac{1}{\sum_{i=1}^{n} t_i} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_{t}^{\infty} \frac{(n-1)\Gamma(n-1)}{\Gamma(n-1)} \left(\frac{\sum_{i=1}^{n} t_i - t_1}{\sum_{i=1}^{n} t_i} \right)^{n-2} \frac{1}{\sum_{i=1}^{n} t_i} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_{t}^{\infty} \frac{(n-1)}{\sum_{i=1}^{n} t_i} \left(1 - \frac{t_1}{\sum_{i=1}^{n} t_i} \right)^{n-2} dt_1$$

Noting that this integral converges if $t_1 < \sum_{i=1}^{n} t_i$, the MVUE of reliability is obtained as

$$\tilde{R}(t) = \int_{t}^{\sum_{i=1}^{n} t_i} \frac{(n-1)}{\sum_{i=1}^{n} t_i} \left(1 - \frac{t_1}{\sum_{i=1}^{n} t_i} \right)^{n-2} dt_1,$$

The term inside the integral is the conditional pdf of t_1 given that $\sum_{i=1}^{n} t_i$ has occurred, as given in (3.2.1). Thus, for a given $\sum_{i=1}^{n} t_i$, the integral depends only on t_1. Hence, taking $\sum_{i=1}^{n} t_i = s$, the integral reduces to

$$\tilde{R}(t) = \int_{t}^{s} \frac{(n-1)}{s}\left(1-\frac{t_1}{s}\right)^{n-2} dt_1 = \left(-\left(1-\frac{t_1}{s}\right)^{n-1}\right)\Bigg|_{t}^{s} = \left(0+(1-\frac{t}{s})^{n-1}\right)$$

Replacing s by $\sum_{i=1}^{n} t_i$, the MVUE of reliability is obtained as

$$\tilde{R}(t) = \begin{cases} \left(1-\dfrac{t}{\sum\limits_{i=1}^{n} t_i}\right)^{n-1} & \text{if } t < \sum\limits_{i=1}^{n} t_i \\ 0 & \text{otherwise} \end{cases} \tag{3.2.6}$$

Here, t is any time instance and $\sum_{i=1}^{n} t_i$ is the sum of such time instances. Hence, if we have a sample failure data of n time instances, say $(t_1, t_2, \ldots, t_n)$, then t (or t_i) is a member of the set $(t_1, t_2, \ldots, t_n)$.

3.3 IMPROVED ESTIMATOR OF $R(t)$

The true reliability function of exponential models is given by

$$R(t) = e^{-\Phi t} = 1 - \Phi t + \frac{(\Phi t)^2}{2!} - \frac{(\Phi t)^3}{3!} + \ldots \tag{3.3.1}$$

$\hat{R}(t)$ and $\tilde{R}(t)$ are unbiased for $R(t)$, if (i) $E(\hat{R}(t)) = R(t)$ and (ii) $E(\tilde{R}(t)) = R(t)$ respectively.

To check whether $\hat{R}(t)$ is unbiased for $R(t)$ or not, consider $E(\hat{R}(t)) = E\Big(e^{-\{\frac{nt}{\sum_{i=1}^{n} t_i}\}}\Big)$.

Taking $Y = \sum_{i=1}^{n} t_i$, we have, $E(\hat{R}(t)) = E\Big(e^{-\{\frac{nt}{Y}\}}\Big)$.

Since $Y \sim G(n, \Phi)$ (Section 1 of Appendix A), it can be observed that,

$$E\Big(\frac{1}{Y}\Big) = \int_0^\infty \frac{1}{y}\frac{\Phi^n}{\Gamma(n)}e^{-y\Phi}y^{n-1}dy$$

$$= \frac{\Phi^n}{\Gamma(n)}\int_0^\infty e^{-\Phi y}y^{n-2}dy$$

$$= \frac{\Phi^n}{\Gamma(n)}\frac{\Gamma(n-1)}{\Phi^{n-1}} \qquad \text{(Result 4 of Appendix A)}$$

$$= \frac{\Phi\Gamma(n-1)}{(n-1)\Gamma(n-1)} = \frac{\Phi}{n-1}.$$

Similarly, $$E\Big(\frac{1}{Y^2}\Big) = \int_0^\infty \frac{1}{y^2}\frac{\Phi^n}{\Gamma(n)}e^{-y\Phi}y^{n-1}dy$$

$$= \frac{\Phi^n}{\Gamma(n)}\int_0^\infty e^{-\Phi y}y^{n-3}dy$$

$$= \frac{\Phi^n}{\Gamma(n)}\frac{\Gamma(n-2)}{\Phi^{n-2}} \qquad \text{(Result 4 of Appendix A)}$$

$$= \frac{\Phi^2\Gamma(n-2)}{(n-1)(n-2)\Gamma(n-2)} = \frac{\Phi^2}{(n-1)(n-2)}.$$

$$E\Big(\frac{1}{Y^3}\Big) = \int_0^\infty \frac{1}{y^3}\frac{\Phi^n}{\Gamma(n)}e^{-y\Phi}y^{n-1}dy$$

$$= \frac{\Phi^n}{\Gamma(n)}\int_0^\infty e^{-\Phi y}y^{n-4}dy$$

$$= \frac{\Phi^n}{\Gamma(n)}\frac{\Gamma(n-3)}{\Phi^{n-3}} \qquad \text{(Result 4 of Appendix A)}$$

$$= \frac{\Phi^3\Gamma(n-3)}{(n-1)(n-2)(n-3)\Gamma(n-3)}$$

$$= \frac{\Phi^3}{(n-1)(n-2)(n-3)}$$

and so on.

$$
\begin{aligned}
\text{Thus, } E(\hat{R}(t)) &= E\Big(1 - \frac{nt}{1!Y} + \frac{(nt)^2}{2!Y^2} - \frac{(nt)^3}{3!Y^3} + \dots\Big) \\
&= 1 - \frac{nt}{1!}E(\frac{1}{Y}) + \frac{(nt)^2}{2!}E(\frac{1}{Y^2}) - \frac{(nt)^3}{3!}E(\frac{1}{Y^3}) + \dots \\
&= 1 - \frac{\Phi t}{1!}\frac{n}{(n-1)} + \frac{(\Phi t)^2}{2!}\frac{n^2}{(n-1)(n-2)} \\
&\quad - \frac{(\Phi t)^3}{3!}\frac{n^3}{(n-1)(n-2)(n-3)} + \dots \\
&\neq R(t) \quad \text{(as given in 3.3.1)}
\end{aligned}
$$

Hence, $\hat{R}(t)$ is not unbiased for $R(t)$.

To verify that $\tilde{R}(t)$ is unbiased for $R(t)$, consider $E(\tilde{R}(t)) = E\left(1 - \frac{t}{\sum_{i=1}^{n} t_i}\right)^{n-1}$.

Since $Y = \sum_{i=1}^{n} t_i$, we have,

$$
\begin{aligned}
E(\tilde{R}(t)) &= E\left(1 - \frac{t}{Y}\right)^{n-1} \\
&= E\Big(1 - \frac{(n-1)t}{Y} + \frac{(n-1)(n-2)}{2!}\frac{t^2}{Y^2} \\
&\qquad - \frac{(n-1)(n-2)(n-3)}{3!}\frac{t^3}{Y^3} + \dots\Big) \\
&= 1 - (n-1)tE(\frac{1}{Y}) + \frac{(n-1)(n-2)t^2}{2!}E(\frac{1}{Y^2}) \\
&\quad - \frac{(n-1)(n-2)(n-3)t^3}{3!}E(\frac{1}{Y^3}) + \dots \\
&= 1 - (n-1)t\frac{\Phi}{n-1} + \frac{(n-1)(n-2)t^2}{2!}\frac{\Phi^2}{(n-1)(n-2)} \\
&\quad - \frac{(n-1)(n-2)(n-3)t^3}{3!}\frac{\Phi^3}{(n-1)(n-2)(n-3)} + \dots \\
&= 1 - \Phi t + \frac{(\Phi t)^2}{2!} - \frac{(\Phi t)^3}{3!} + \dots \\
&= R(t)
\end{aligned}
$$

Thus, $\tilde{R}(t)$ is unbiased for $R(t)$.

Since $\hat{R}(t)$ is biased for $R(t)$ and $\tilde{R}(t)$ is unbiased for $R(t)$, the bias in $\hat{R}(t)$ is obtained as

$$\text{Bias}(\hat{R}(t)) = E(\hat{R}(t)) - e^{-\Phi t} = E(\hat{R}(t)) - E(\tilde{R}(t)) \tag{3.3.2}$$

$$\text{Hence, Bias } (\hat{R}(t)) = -\frac{\Phi t}{(n-1)} + \frac{(\Phi t)^2}{2!}\frac{(3n-2)}{(n-1)(n-2)} - \frac{(\Phi t)^3}{3!}\frac{(6n^2-11n+6)}{(n-1)(n-2)(n-3)} + \dots \tag{3.3.3}$$

Thus, the above bias in $\hat{R}(t)$ can be found for the given sample failure data, by using the estimated values of $\hat{R}(t)$ and $\tilde{R}(t)$.
Hence, if $\mathscr{T} = \{t_1, t_2, \dots t_n\}$ is the given sample failure data set of size n, then the bias is obtained by taking the difference in the means of $\hat{R}(t)$ and $\tilde{R}(t)$ and is obtained as

$$\text{Bias}(\hat{R}(t)) = \frac{\sum\limits_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n} \tag{3.3.4}$$

Removing this bias from $\hat{R}(t)$, the Improved Estimator of $R(t)$ denoted by $\breve{R}(t)$, is obtained as $\breve{R}(t) = \hat{R}(t) - \text{Bias}(\hat{R}(t)) = \hat{R}(t) - \left(\frac{\sum\limits_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n}\right)$.

$$\text{i.e. } \breve{R}(t) = e^{-\left(\frac{nt}{\sum\limits_{i=1}^{n} t_i}\right)} - \left(\frac{\sum\limits_{t\epsilon\mathscr{T}} e^{-\left(\frac{nt}{\sum\limits_{i=1}^{n} t_i}\right)}}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}} \left(1 - \frac{t}{\sum\limits_{i=1}^{n} t_i}\right)^{n-1}}{n}\right) \tag{3.3.5}$$

In all the above calculations, t is any time instance. For a sample failure time data set $\mathscr{T}$, as given above, t is a member of $\mathscr{T}$.

3.4 COMPARISON OF ESTIMATES

The three estimators of reliability are to be compared by comparing the properties satisfied by them. The Improved Estimator of $R(t)$ is unbiased and sufficient, as it is obtained from MLE of $R(t)$, by removing the bias present in it. The only property to be checked thus, is the efficiency property. Since MLE of $R(t)$ is biased as shown above, while MVUE of $R(t)$ and Improved Estimator of $R(t)$ are unbiased, coefficient of variation is used as a measure of dispersion instead of the variance, as mentioned in Section 1.1 of Chapter 1. The estimate with the least value of the coefficient of variation is considered as the efficient estimator. The comparison is also done by considering the quartile coefficient of dispersion, as mentioned in Section 1.1 of Chapter 1. Even with this measure, the estimate with the least value of the quartile coefficient of dispersion is considered as the efficient estimator. For this purpose, the following case studies have been considered and the three estimates have been found. The coefficients of variation and the quartile coefficient of dispersion for these three estimates have also been obtained. For all the case studies, CV($\hat{R}(t)$), CV($\tilde{R}(t)$) and CV($\breve{R}(t)$) are respectively obtained using

$$\text{CV}(\hat{R}(t)) = \frac{S_{\hat{R}(t)}}{\overline{\hat{R}}(t)} \tag{3.4.1}$$

$$\text{CV}(\tilde{R}(t)) = \frac{S_{\tilde{R}(t)}}{\overline{\tilde{R}}(t)} \tag{3.4.2}$$

$$\text{CV}(\breve{R}(t)) = \frac{S_{\breve{R}(t)}}{\overline{\breve{R}}(t)} \tag{3.4.3}$$

Here, the sample variances $S^2_{\hat{R}(t)}$, $S^2_{\tilde{R}(t)}$ and $S^2_{\breve{R}(t)}$ are respectively obtained using

$$S^2_{\hat{R}(t)} = \sum_{t\epsilon\mathscr{T}} \frac{\left(\hat{R}(t) - \overline{\hat{R}}(t)\right)^2}{(n-1)} \tag{3.4.4}$$

$$S^2_{\tilde{R}(t)} = \sum_{t\epsilon\mathscr{T}} \frac{\left(\tilde{R}(t) - \overline{\tilde{R}}(t)\right)^2}{(n-1)} \tag{3.4.5}$$

$$S^2_{\breve{R}(t)} = \sum_{t\epsilon\mathscr{T}} \frac{\left(\breve{R}(t) - \overline{\breve{R}}(t)\right)^2}{(n-1)} \tag{3.4.6}$$

Further, the sample means $\overline{\hat{R}}(t)$, $\overline{\tilde{R}}(t)$, and $\overline{\breve{R}}(t)$ are respectively obtained using

$$\overline{\hat{R}}(t) = \frac{\sum_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} \tag{3.4.7}$$

$$\overline{\tilde{R}}(t) = \frac{\sum_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n} \tag{3.4.8}$$

$$\overline{\breve{R}}(t) = \frac{\sum_{t\epsilon\mathscr{T}} \breve{R}(t)}{n} \tag{3.4.9}$$

The above equations are used in calculations pertaining to all the four classes of models.

The failure data are obtained using different methods in literature. Qureshi and Daniel (Quereshi and Jeske (1997)) used proxy failure times, by simulating them. Jelinski-Morand model of exponential class was considered and goodness of fit was done. Herein, instead of such proxy failure data, some past bench mark data obtained during testing have been used for comparison.

Case study 1: On-Line Data Entry IBM Software Package

The data reported by Ohba (Ohba (1984)) are recorded from testing an on-line data entry software package developed at IBM. There are 15 failures, with failure times as indicated in Table 3.4.1.

Table 3.4.1 On-Line Data Entry IBM Software Package

Failure Number	1	2	3	4	5	6	7	8	9	10
Failure Time	10	19	32	43	58	70	88	103	125	150
Failure Number	11	12	13	14	15					
Failure Time	169	199	231	256	296					

Table 3.4.2 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in Table 3.4.2 in equations (3.4.4) to (3.4.9), we get,
$\overline{\hat{R}}(t) = 0.456803$; $S^2_{\hat{R}(t)} = 0.0756$; $\overline{\tilde{R}}(t) = 0.465465$; $S^2_{\tilde{R}(t)} = 0.0775$;
$\overline{\breve{R}}(t) = 0.465463$; $S^2_{\breve{R}(t)} = 0.0755$.
Hence, using (3.3.2), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = E(\hat{R}(t)) - E(\tilde{R}(t)) = \overline{\hat{R}}(t) - \overline{\tilde{R}}(t) = 0.456803 - 0.465465 = -0.00866$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.00866) = \hat{R}(t) + 0.00866$.
Thus, using equations (3.4.1) to (3.4.3), the coefficient of variation of the three estimators are respectively obtained as,
CV($\hat{R}(t)$)=0.6017, CV($\tilde{R}(t)$)=0.5981 and **CV($\boldsymbol{\breve{R}(t)}$)=0.5905**.
It is observed that the Improved Estimator $\breve{R}(t)$ has the least value of the coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.1990 and Q_3=0.7055. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as

Table 3.4.2 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Exponential Model (Case Study 1)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	SD$_{\hat{R}(t)}$	SD$_{\tilde{R}(t)}$	SD$_{\breve{R}(t)}$
1	10	0.92207	0.92688	0.93073	0.21648	0.21291	0.21648
2	19	0.85715	0.86536	0.86581	0.16028	0.15991	0.16028
3	32	0.77136	0.78316	0.78002	0.09894	0.10093	0.09894
4	43	0.70550	0.71933	0.71416	0.06185	0.06444	0.06185
5	58	0.62467	0.64006	0.63333	0.02818	0.03048	0.02818
6	70	0.56672	0.58256	0.57538	0.01208	0.01371	0.01208
7	88	0.48973	0.50525	0.49839	0.00108	0.00158	0.00108
8	103	0.43361	0.44823	0.44227	0.00053	0.00029	0.00053
9	125	0.36274	0.37532	0.37140	0.00884	0.00812	0.00884
10	150	0.29615	0.30590	0.30481	0.02580	0.02545	0.02580
11	169	0.25385	0.26134	0.26251	0.04119	0.04166	0.04119
12	199	0.19901	0.20307	0.20767	0.06645	0.06884	0.06645
13	231	0.15351	0.15437	0.16217	0.09198	0.09677	0.09198
14	256	0.12533	0.12413	0.13399	0.10987	0.11650	0.10987
15	296	0.09060	0.08695	0.09926	0.13410	0.14327	0.13410

QD($\hat{R}(t)$)=0.5600. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2030 and Q_3=0.7193. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.5598. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2077 and Q_3=0.7142. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.5494**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 3.4.1.

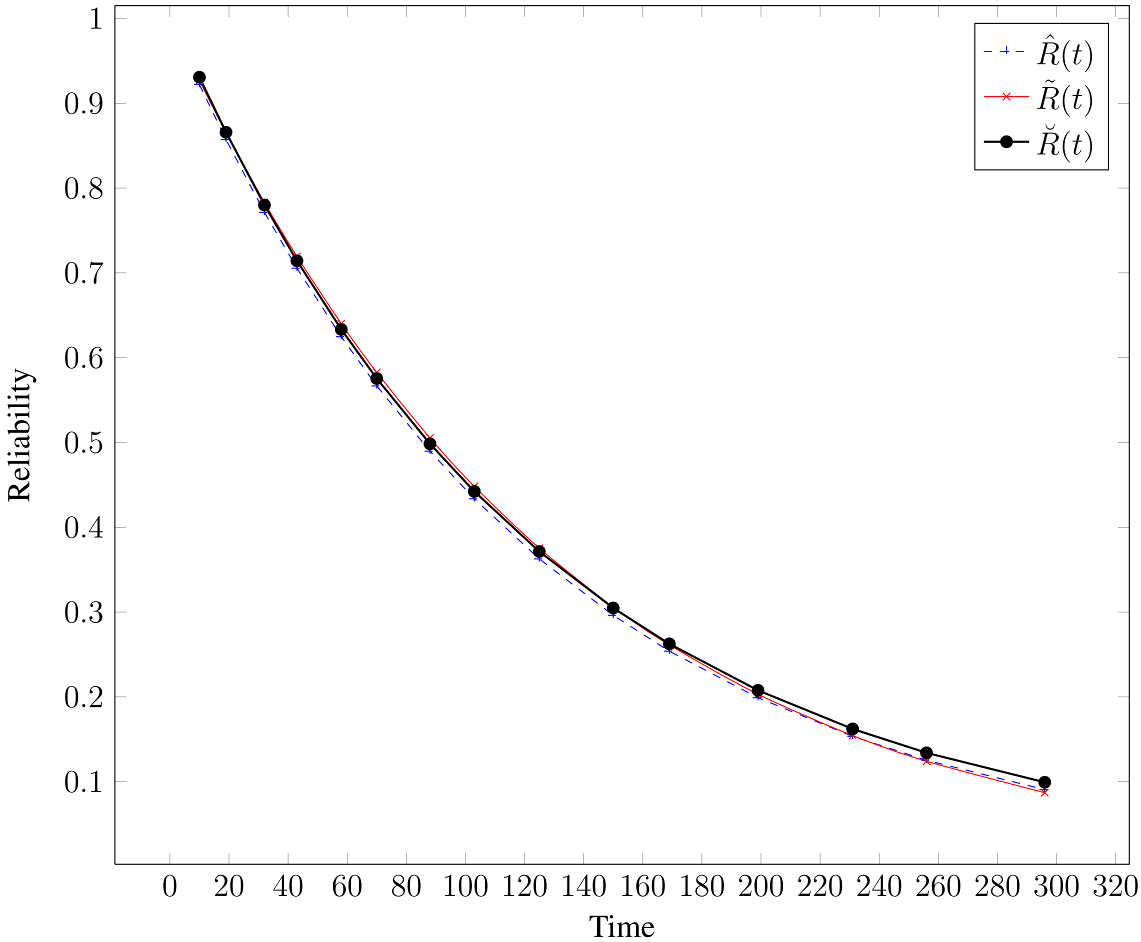

Figure 3.4.1 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Exponential Model (Case Study 1)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 2: Nuclear Power Agency

A nuclear power agency uses a computer-based monitoring system for its reactors. The operating system for the computer is employed for this and other applications in an estimated 5000 installations throughout the world. A total of 17 failures that occurred with failure times are listed in Table 3.4.3 (Musa et al. (1991)).

Table 3.4.3 Nuclear Power Agency

Failure number	1	2	3	4	5	6
Failure time	932	4035	4696	4893	6369	6524
Failure number	7	8	9	10	11	12
Failure time	7882	8170	9339	10400	10542	11036
Failure number	13	14	15	16	17	
Failure time	11696	11905	12266	12954	14000	

Table 3.4.4 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ respectively denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means.

Table 3.4.4 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Exponential Model (Case study 2)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$\text{SD}_{\hat{R}(t)}$	$\text{SD}_{\tilde{R}(t)}$	$\text{SD}_{\breve{R}(t)}$
1	932	0.89824	0.90364	0.90774	0.24666	0.24260	0.24666
2	4035	0.62837	0.64186	0.63787	0.05143	0.05325	0.05143
3	4696	0.58232	0.59619	0.59182	0.03266	0.03426	0.03266
4	4893	0.56926	0.58318	0.57876	0.02811	0.02961	0.02811
5	6369	0.48029	0.49383	0.48979	0.00619	0.00684	0.00619
6	6524	0.47179	0.48523	0.48129	0.00492	0.00549	0.00492
7	7882	0.40350	0.41568	0.41300	0.00000	0.00002	0.00000
8	8170	0.39033	0.40218	0.39983	0.00012	0.00007	0.00012
9	9339	0.34117	0.35150	0.35067	0.00364	0.00354	0.00364
10	10400	0.30194	0.31075	0.31144	0.00992	0.01006	0.00992
11	10542	0.29704	0.30565	0.30654	0.01092	0.01111	0.01092
12	11036	0.28062	0.28850	0.29012	0.01463	0.01502	0.01463
13	11696	0.26008	0.26698	0.26958	0.02002	0.02076	0.02002
14	11905	0.25390	0.26049	0.26340	0.02181	0.02267	0.02181
15	12266	0.24356	0.24962	0.25306	0.02497	0.02606	0.02497
16	12954	0.22501	0.23008	0.23451	0.03117	0.03276	0.03117
17	14000	0.19948	0.20310	0.20898	0.04084	0.04325	0.04084

Using the values of Table 3.4.4 in equations (3.4.4) to (3.4.9), we get,
$\overline{\hat{R}}(t) = 0.401586$; $S^2_{\hat{R}(t)} = 0.0343$; $\overline{\tilde{R}}(t) = 0.411091$; $S^2_{\tilde{R}(t)} = 0.0348$;
$\overline{\breve{R}}(t) = 0.411087$; $S^2_{\breve{R}(t)} = 0.0343$.
Also, using (3.3.2), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = 0.401586 - 0.411091 = -0.0095$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.0095) = \hat{R}(t) + 0.0095$.

Thus, using equations (3.4.1) to (3.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.4609, CV($\tilde{R}(t)$)=0.4538 and **CV($\breve{\boldsymbol{R}}(\boldsymbol{t})$)=0.4503**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.2570 and Q_3=0.5248. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.3425. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2637 and Q_3=0.5385. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.3426. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2665 and Q_3=0.5343. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{\boldsymbol{R}}(\boldsymbol{t})$)=0.3344**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 3.4.2.

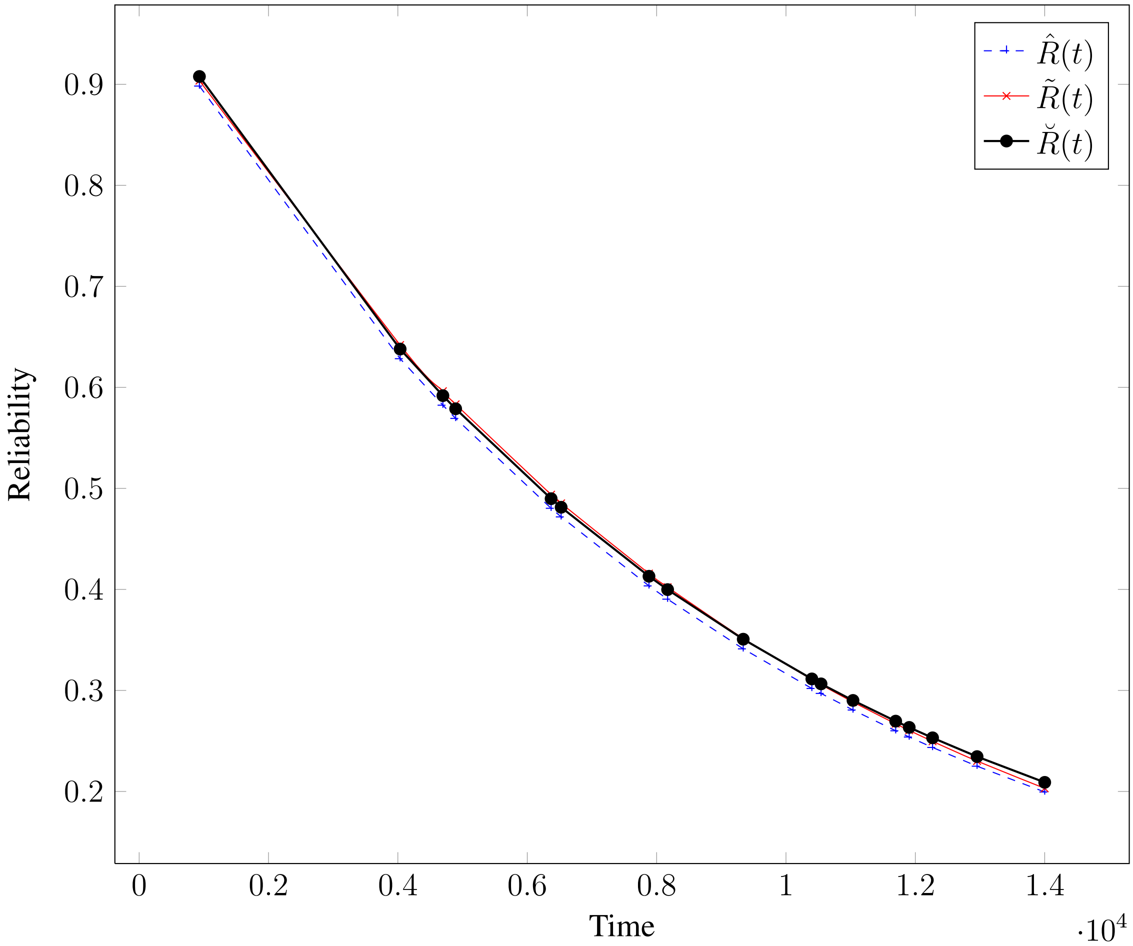

Figure 3.4.2 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Exponential Model (Case study 2)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 3: Failure data set of Lyu

The failure time data for 10 failures obtained by Lyu (Lyu (2004)) are given in Table 3.4.5.

Table 3.4.5 Failure data set of Lyu

Failure number	1	2	3	4	5	6	7	8	9	10
Failure time	7	18	26	36	51	73	93	118	146	181

Table 3.4.6 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively.

Table 3.4.6 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Exponential Model (Case study 3)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$\text{SD}_{\hat{R}(t)}$	$\text{SD}_{\tilde{R}(t)}$	$\text{SD}_{\breve{R}(t)}$
1	7	0.91189	0.91999	0.92359	0.06800	0.06327	0.20022
2	18	0.78886	0.80572	0.79919	0.01897	0.01884	0.10436
3	26	0.70995	0.73073	0.71953	0.0034	0.00387	0.05924
4	36	0.62231	0.64575	0.63120	0.0008	0.00051	0.02404
5	51	0.51071	0.53471	0.51897	0.01971	0.01788	0.00183
6	73	0.38220	0.40247	0.39015	0.07231	0.07074	0.00739
7	93	0.29367	0.30838	0.30172	0.12776	0.12965	0.03041
8	118	0.21125	0.21854	0.21973	0.1934	0.20241	0.06574
9	146	0.14608	0.14620	0.15519	0.25505	0.27274	0.10300
10	181	0.09211	0.08613	0.10204	0.31247	0.33910	0.13994

Using the data obtained in Table 3.4.6 in equations (3.4.4) to (3.4.9), we get,
$\overline{\hat{R}}(t) = 0.463317$; $S^2_{\hat{R}(t)} = 0.0818$; $\overline{\tilde{R}}(t) = 0.476136$; $S^2_{\tilde{R}(t)} = 0.0853$;

$\overline{\breve{R}}(t) = 0.476137$; $S^2_{\breve{R}(t)} = 0.0818$.
Using (3.3.2), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = 0.463317 - 0.476136 = -0.01282$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.01282) = \hat{R}(t) + 0.01282$.

Thus, using equations (3.4.1) to (3.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.6173, CV($\tilde{R}(t)$)=0.6134 and **CV($\breve{R}(t)$)=0.6007**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has less value of coefficient of variation than those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.21125 and Q_3=0.70995. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.5414. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.21854 and Q_3=0.73073. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.5400. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.21973 and Q_3=0.71953. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.5321**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 3.4.3.

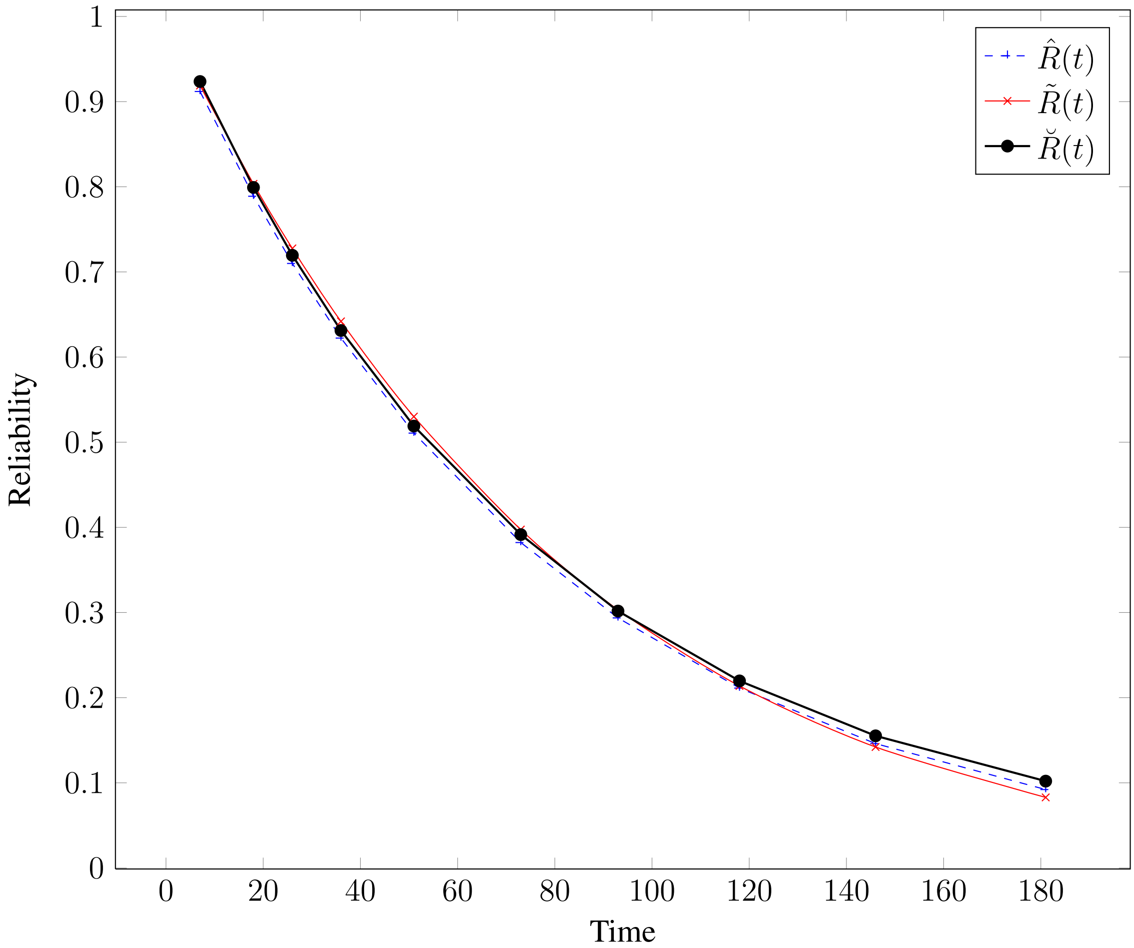

Figure 3.4.3 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Exponential Model (Case study 3)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Table 3.4.7 shows the consolidated values of the coefficients of variation (CV) of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for all the three case studies.

Table 3.4.7 Exponential Models (Consolidated1)

Case study	CV($\hat{R}(t)$)	CV($\tilde{R}(t)$)	CV($\breve{R}(t)$)
1	0.6017	0.5981	**0.5905**
2	0.4609	0.4538	**0.4503**
3	0.6173	0.6134	**0.6007**

Table 3.4.8 shows the consolidated values of the quartile coefficient of dispersions (QD) of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for all the three case studies.

Table 3.4.8 Exponential Models (Consolidated2)

Case study	QD($\hat{R}(t)$)	QD($\tilde{R}(t)$)	QD($\breve{R}(t)$)
1	0.5600	0.5598	**0.5494**
2	0.3425	0.3426	**0.3344**
3	0.5414	0.5400	**0.5321**

From Tables 3.4.7 and 3.4.8, it is observed that, the Improved Estimator ($\breve{R}(t)$) has the least values of the coefficient of variation and quartile coefficient of dispersion than those of MLE ($\hat{R}(t)$) and MVUE ($\tilde{R}(t)$) in all the three case studies, which means that $\breve{R}(t)$ is more efficient than $\hat{R}(t)$ and $\tilde{R}(t)$.

However, to choose the best estimator among the three estimates, the desirable properties of good estimators as mentioned in Section 1.1 of Chapter 1 are to be considered.

Unbiasedness of $\breve{R}(t)$: It has been shown above that $\hat{R}(t)$ is biased for $R(t)$, while $\tilde{R}(t)$ is unbiased for $R(t)$. Since the Improved Estimators are obtained from MLEs just by removing the bias present in the MLEs, they satisfy the unbiasedness property. Thus, $\breve{R}(t)$ is unbiased for $R(t)$.

Sufficiency of $\breve{R}(t)$: Improved Estimator is a function of MLE, which is sufficient. Since any function of sufficient estimator is also sufficient, Improved Estimator is sufficient.

Now, to compare the biased estimator $\hat{R}(t)$ with unbiased estimators $\tilde{R}(t)$ and $\breve{R}(t)$, the coefficient of variation and the quartile coefficient of dispersion are considered as measures of dispersion to check the efficiency property. The sample results of comparison of coefficients of variation and the quartile coefficients of dispersion for the three estimators indicate that Improved Estimator has least values of coefficient of variation and the quartile coefficient of dispersion as compared to those of MLE and MVUE of $R(t)$, which indicates that the Improved estimators are efficient compared to MLE and MVUE.

Thus, by referring to Table 1.1.1 of Chapter 1, Table 3.4.9 provides the statistical properties satisfied by MLE, MVUE and the Improved Estimator of reliability for exponential class models.

Table 3.4.9 Exponential class models- Properties satisfied by estimators of reliability

	Unbiased	**Sufficient**	**Efficient**
MLE	No	Yes	No
MVUE	Yes	Yes	No
Improved Estimator	**Yes**	**Yes**	**Yes**

It can be seen from Table 3.4.9 that the Improved Estimator satisfies maximum number of properties of estimators as compared to MLE and MVUE of $R(t)$. Hence, it can be inferred that the estimate of reliability obtained using the Improved Estimator, is more efficient than those estimated using the methods of MLE and MVUE.

Hence, it is concluded that $\breve{R}(t)$ gives more accurate value of reliability than $\hat{R}(t)$ and $\tilde{R}(t)$, for exponential class software reliability models.

Chapter 4

WEIBULL CLASS MODELS

In this class of models, the failure times (T) are assumed to have Weibull distribution with probability density function, given by

$$f(t) = \Phi\beta t^{\beta-1}e^{-\Phi t^{\beta}}, t > 0 \tag{4.0.1}$$

where β denotes the shape parameter and the scale parameter Φ denotes the failure rate. Among works on estimation of parameters and reliability, the one by Chris Bambey et. al (Guure et al. (2013)) focussed on Baysian parameter estimation technique. The estimation technique was applied to two parameter Weibull failure time distribution. However, its superiority over other methods has not been carried out. With the use of computer algorithms and developments of computer based techniques, the estimation of software reliability has attracted the attention of computer scientists too. In this direction, Parveen Sehgal and Meenal (Sehgal and Meenal (2016)) used artificial intelligence techniques in estimating the reliability. The software reliability was estimated based on historical data sets. Eventhough no specific model was considered, it was intended to apply the method for all types and classes of software reliability models and so also the Weibull model. Algorithms like genetic algorithms were also used by Taehyoun et. al (Kim et al. (2015)) to estimate the reliability of any model and so also the Weibull model. However, it is not very feasible to compare these algorithms with other methods because of the complexity involved in obtaining their time complexity. Hence, the role of statistics still plays a major part of estimation procedures, as they also enable to compare various techniques using statistical testing procedures. Thus, it is intended to apply the methods of MLE and MVUE in the reliability estimation procedure of this class of models.
If T has Weibull distribution with parameters Φ and β, it is denoted as $T \sim W(\Phi, \beta)$.

The reliability function at time t, denoted by$R(t)$, is obtained as

$$R(t) = P(T > t) = \int_t^\infty f(t)dt = \int_t^\infty \Phi\beta t^{\beta-1} e^{-\Phi t^\beta} dt = e^{-\Phi t^\beta} \tag{4.0.2}$$

Schick and Wolverton obtained a Weibull model, where the value of the shape parameter β is found to be 2 (Schick and Wolverton (1973)). Thus, β is assumed to be 2 here.

4.1 MLE OF $R(t)$

Since MLEs satisfy the invariance property (as explained in Chapter 1), the MLE of $R(t)$, denoted by $\hat{R}(t)$ is obtained as

$$\hat{R}(t) = e^{-\hat{\Phi} t^\beta} \tag{4.1.1}$$

where $\hat{\Phi}$ is the MLE of Φ.

To find the MLE of Φ: Let $(T_1, T_2, \ldots T_n)$ be a sample of size n from Weibull distribution as given in (4.0.1). Then, the likelihood function of this sample is given by

$$L = \prod_{i=1}^{n} f(t_i) = \Phi^n \beta^n e^{-\Phi \sum\limits_{i=1}^{n} t_i^\beta} \prod_{i=1}^{n} t_i^{\beta-1} \tag{4.1.2}$$

Maximizing L using the concept of differential calculus, the MLE of Φ, denoted by $\hat{\Phi}$, is obtained as the solution of $\dfrac{\partial \ln L}{\partial \Phi} = 0$, with $\dfrac{\partial^2 \ln L}{\partial \Phi^2} < 0$. This gives $\dfrac{n}{\Phi} - \sum\limits_{i=1}^{n} t_i^\beta = 0$, from which, the MLE of Φ is obtained as

$$\hat{\Phi} = \frac{n}{\sum\limits_{i=1}^{n} t_i^\beta} \tag{4.1.3}$$

Using (4.1.3) in (4.1.1), the MLE of $R(t)$ is obtained as

$$\hat{R}(t) = e^{-\left(\frac{nt^\beta}{\sum_{i=1}^{n} t_i^\beta}\right)} \tag{4.1.4}$$

4.2 MVUE OF $R(t)$

To find the MVUE of $R(t)$, it is intended to obtain the unbiased estimator of $R(t)$ and the complete sufficient estimator of the parameter Φ.

Unbiased Estimator of $R(t)$: Define a function of the random variable T_1 as

$$U(t_1) = \begin{cases} 1 & \text{if } t_1 > t \\ 0 & \text{otherwise} \end{cases}$$

Then, $E[U(t_1)] = 1.P(T_1 > t) + 0.P(T_1 \leq t) = P(T_1 > t) = R(t)$.
Therefore, $U(t_1)$ is unbiased for $R(t)$.

Complete Sufficient Estimator of Φ: Applying the factorization theorem (as explained in Chapter 1) to the likelihood function given in (4.1.2), it can be seen that the likelihood function depends on Φ and t_i, only through the value of $\sum_{i=1}^{n} t_i^\beta$. Thus, $\sum_{i=1}^{n} t_i^\beta$ is the sufficient estimator of the parameter Φ.
Since each $T_i \sim W(\Phi, \beta,)$, each $T_i^\beta \sim \mathscr{E}(\Phi)$ (Section 2 of Appendix A). Hence, $\sum_{i=1}^{n} T_i^\beta \sim G(n, \Phi)$ (Section 1 of Appendix A). Further, by Result 3 of Appendix A, this estimator $\sum_{i=1}^{n} t_i^\beta$ is also the complete statistic of Φ and hence, $\sum_{i=1}^{n} t_i^\beta$ is the complete sufficient estimator of Φ.
Further, $U(t_1)$ is an unbiased estimator of $R(t)$ and $R(t)$ is a function of Φ, as given in (4.0.2). Hence, by Theorem 1 of Chapter 1, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = E(U(t_1)|\sum_{i=1}^{n} t_i^\beta) = \int_{t}^{\infty} f(t_1|\sum_{i=1}^{n} t_i^\beta)dt_1 \tag{4.2.1}$$

where $f(t_1|\sum_{i=1}^{n} t_i^\beta)$ denotes the conditional pdf of T_1 given $\sum_{i=1}^{n} T_i^\beta$ and is given by $f(t_1|\sum_{i=1}^{n} t_i^\beta) = \frac{g(t_1, \sum_{i=1}^{n} t_i^\beta)}{h(\sum_{i=1}^{n} t_i^\beta)}$, where $g(t_1, \sum_{i=1}^{n} t_i^\beta)$ denotes the joint pdf of T_1 and $\sum_{i=1}^{n} T_i^\beta$ and $h(\sum_{i=1}^{n} t_i^\beta)$ denotes the marginal pdf of $\sum_{i=1}^{n} T_i^\beta$.

Hence, the MVUE of $R(t)$ is given by

$$\tilde{R}(t) = \int_t^\infty \frac{g(t_1, \sum_{i=1}^n t_i^\beta)}{h(\sum_{i=1}^n t_i^\beta)} dt_1 \tag{4.2.2}$$

Since each $T_i \sim W(\Phi, \beta,)$, each $T_i^\beta \sim \mathscr{E}(\Phi)$ (Section 2 of Appendix A). Hence, $\sum_{i=1}^n T_i^\beta \sim G(n, \Phi)$ (Section 1 of Appendix A).
Therefore, the pdf of this random variable is given by

$$h(\sum_{i=1}^n t_i^\beta) = \frac{\Phi^n}{\Gamma(n)} e^{-\Phi \sum_{i=1}^n t_i^\beta} (\sum_{i=1}^n t_i^\beta)^{n-1} \tag{4.2.3}$$

To find the probability function $g(t_1, \sum_{i=1}^n t_i^\beta)$, split the sample $(T_1, T_2, T_3, \dots T_n)$ into two samples as T_1 of size one and $(T_2, T_3, T_4, \dots T_n)$ of size $(n-1)$.
Since T_1 and $\sum_{i=2}^n T_i^\beta$ are independent, the joint pdf of T_1 and $\sum_{i=2}^n T_i^\beta$ is obtained as

$$g(t_1, \sum_{i=2}^n t_i^\beta) = f(t_1) \cdot h(\sum_{i=2}^n t_i^\beta),$$

where $f(t_1)$ and $h(\sum_{i=2}^n t_i^\beta)$ denote the pdfs of T_1 and $\sum_{i=2}^n T_i^\beta$ respectively.
Since $T_1 \sim W(\Phi, \beta)$, the pdf of T_1 is given by $f(t_1) = \Phi\beta t_1^{\beta-1} e^{-\Phi t_1^\beta}$.
Also, $\sum_{i=2}^n T_i^\beta \sim G(n-1, \Phi)$ and hence its pdf is given by
$h(\sum_{i=2}^n t_i^\beta) = \frac{\Phi^{n-1}}{\Gamma(n-1)} e^{-\Phi \sum_{i=2}^n t_i^\beta} (\sum_{i=2}^n t_i^\beta)^{n-2}$.
Hence, the joint pdf of T_1 and $\sum_{i=2}^n T_i$ is obtained as
$g(t_1, \sum_{i=2}^n t_i^\beta) = f(t_1) \cdot h(\sum_{i=2}^n t_i^\beta) = \Phi\beta t_1^{\beta-1} e^{-\Phi t_1^\beta} \frac{\Phi^{n-1}}{\Gamma(n-1)} e^{-\Phi \sum_{i=2}^n t_i^\beta} (\sum_{i=2}^n t_i^\beta)^{n-2}$,
which simplifies to $g(t_1, \sum_{i=2}^n t_i^\beta) = \frac{\Phi^n e^{-\Phi \sum_{i=1}^n t_i^\beta}}{\Gamma(n-1)} \left(\sum_{i=2}^n t_i^\beta \right)^{n-2} \beta t_1^{\beta-1}$.
Considering the transformation $\sum_{i=1}^n T_i^\beta = T_1^\beta + \sum_{i=2}^n T_i^\beta$ and noting that the modulus

of the Jacobian of the inverse transformation is one(Appendix B), the joint probability density function of T_1 and $\sum_{i=1}^{n} T_i^{\beta}$ is obtained as

$$g(t_1, \sum_{i=1}^{n} t_i^{\beta}) = \frac{\Phi^n e^{-\Phi \sum_{i=1}^{n} t_i^{\beta}}}{\Gamma(n-1)} \left(\sum_{i=1}^{n} t_i^{\beta} - t_1^{\beta} \right)^{n-2} \beta t_1^{\beta-1} \tag{4.2.4}$$

Substituting (4.2.3) and (4.2.4) in (4.2.2), the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_t^{\infty} \frac{\frac{\Phi^n}{\Gamma(n-1)} e^{-\Phi \sum_{i=1}^{n} t_i^{\beta}} (\sum_{i=1}^{n} t_i^{\beta} - t_1^{\beta})^{n-2}}{\frac{\Phi^n}{\Gamma(n)} e^{-\Phi \sum_{i=1}^{n} t_i^{\beta}} (\sum_{i=1}^{n} t_i^{\beta})^{n-1}} \beta t_1^{\beta-1} dt_1$$

This reduces to

$$\tilde{R}(t) = \int_t^{\infty} \frac{\Gamma(n)}{\Gamma(n-1)} \frac{1}{(\sum_{i=1}^{n} t_i^{\beta})^{n-1}} \left(\sum_{i=1}^{n} t_i^{\beta} - t_1^{\beta} \right)^{n-2} \beta t_1^{\beta-1} dt_1,$$

$$\text{i.e., } \tilde{R}(t) = \int_t^{\infty} \frac{(n-1)}{(\sum_{i=1}^{n} t_i^{\beta})^{n-1}} \left(\sum_{i=1}^{n} t_i^{\beta} - t_1^{\beta} \right)^{n-2} \beta t_1^{\beta-1} dt_1,$$

$$\text{i.e., } \tilde{R}(t) = \int_t^{\infty} \frac{\beta(n-1) t_1^{\beta-1}}{\sum_{i=1}^{n} t_i^{\beta}} \left(\frac{\sum_{i=1}^{n} t_i^{\beta} - t_1^{\beta}}{\sum_{i=1}^{n} t_i^{\beta}} \right)^{n-2} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_t^{\infty} \frac{\beta(n-1)}{\sum_{i=1}^{n} t_i^{\beta}} \left(1 - \frac{t_1^{\beta}}{\sum_{i=1}^{n} t_i^{\beta}} \right)^{n-2} t_1^{\beta-1} dt_1$$

This integral converges if $t_1^{\beta} < \sum_{i=1}^{n} t_i^{\beta}$. i.e., if $t_1 < \left(\sum_{i=1}^{n} t_i^{\beta} \right)^{\frac{1}{\beta}}$.
Thus, the MVUE of reliability is obtained as

$$\tilde{R}(t) = \int_t^{(\sum_{i=1}^{n} t_i^{\beta})^{\frac{1}{\beta}}} \frac{\beta(n-1)}{\sum_{i=1}^{n} t_i^{\beta}} \left(1 - \frac{t_1^{\beta}}{\sum_{i=1}^{n} t_i^{\beta}} \right)^{n-2} t_1^{\beta-1} dt_1$$

The term inside the integral is the conditional pdf of t_1 given that $\sum_{i=1}^{n} t_i^{\beta}$ has occurred, as given in (4.2.1). Thus, for a given $\sum_{i=1}^{n} t_i^{\beta}$, the integral depends only on t_1^{β}. Hence, taking $\sum_{i=1}^{n} t_i^{\beta} = p$, the integral becomes

$$\tilde{R}(t) = \int_{t}^{p^{\frac{1}{\beta}}} \frac{\beta(n-1)}{p}\left(1-\frac{t_1^{\beta}}{p}\right)^{n-2} t_1^{\beta-1} dt_1 = \left(-\left(1-\frac{t_1^{\beta}}{p}\right)^{n-1}\right)\Bigg|_{t}^{p^{\frac{1}{\beta}}}$$

$$\text{i.e., } \tilde{R}(t) = \left(0+(1-\frac{t^{\beta}}{p})^{n-1}\right) = \left(1-\frac{t^{\beta}}{p}\right)^{n-1}$$

Replacing p by $\sum_{i=1}^{n} t_i^{\beta}$, MVUE of reliability is obtained as

$$\tilde{R}(t) = \begin{cases} \left(1-\dfrac{t^{\beta}}{\sum_{i=1}^{n} t_i^{\beta}}\right)^{n-1} & \text{if } t < \sum_{i=1}^{n} t_i^{\beta} \\ 0 & \text{otherwise} \end{cases} \tag{4.2.5}$$

4.3 IMPROVED ESTIMATOR OF $R(t)$

The true reliability function is given by

$$R(t) = e^{-\Phi t^{\beta}} = 1 - \Phi t^{\beta} + \frac{(\Phi t^{\beta})^2}{2!} - \frac{(\Phi t^{\beta})^3}{3!} + \ldots \tag{4.3.1}$$

$\hat{R}(t)$ and $\tilde{R}(t)$ are unbiased for $R(t)$, if (i) $E(\hat{R}(t)) = R(t)$ and (ii) $E(\tilde{R}(t)) = R(t)$ respectively.

To check whether $\hat{R}(t)$ is unbiased or not, consider $E(\hat{R}(t)) = E\Big(e^{-\{\frac{nt^{\beta}}{\sum_{i=1}^{n} t_i^{\beta}}\}}\Big)$.

Taking $Y = \sum_{i=1}^{n} t_i^{\beta}$, we have, $E(\hat{R}(t)) = E\Big(e^{-\{\frac{nt^{\beta}}{Y}\}}\Big)$.

Since the above random variable Y is the sum of n independent exponential random variates, we have $Y \sim G(n, \Phi)$. (Section 1 of Appendix A). Further, we have,

$$E\Big(\frac{1}{Y}\Big) = \int_0^\infty \frac{1}{y}\frac{\Phi^n}{\Gamma(n)} e^{-y\Phi} y^{n-1} dy$$
$$= \frac{\Phi^n}{\Gamma(n)} \int_0^\infty e^{-\Phi y} y^{n-2} dy$$
$$= \frac{\Phi^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\Phi^{n-1}} \quad \text{(Result 4 of Appendix A)}$$
$$= \frac{\Phi\Gamma(n-1)}{(n-1)\Gamma(n-1)} = \frac{\Phi}{n-1}.$$

Similarly, $$E\Big(\frac{1}{Y^2}\Big) = \int_0^\infty \frac{1}{y^2}\frac{\Phi^n}{\Gamma(n)} e^{-y\Phi} y^{n-1} dy$$
$$= \frac{\Phi^n}{\Gamma(n)} \int_0^\infty e^{-\Phi y} y^{n-3} dy$$
$$= \frac{\Phi^n}{\Gamma(n)} \frac{\Gamma(n-2)}{\Phi^{n-2}} \quad \text{(Result 4 of Appendix A)}$$
$$= \frac{\Phi^2\Gamma(n-2)}{(n-1)(n-2)\Gamma(n-2)} = \frac{\Phi^2}{(n-1)(n-2)}.$$

$$E\Big(\frac{1}{Y^3}\Big) = \int_0^\infty \frac{1}{y^3}\frac{\Phi^n}{\Gamma(n)} e^{-y\Phi} y^{n-1} dy$$
$$= \frac{\Phi^n}{\Gamma(n)} \int_0^\infty e^{-\Phi y} y^{n-4} dy$$
$$= \frac{\Phi^n}{\Gamma(n)} \frac{\Gamma(n-3)}{\Phi^{n-3}} \quad \text{(Result 4 of Appendix A)}$$
$$= \frac{\Phi^3\Gamma(n-3)}{(n-1)(n-2)(n-3)\Gamma(n-3)}$$
$$= \frac{\Phi^3}{(n-1)(n-2)(n-3)}$$

and so on.

$$\text{So, } E(\hat{R}(t)) = E\Big(1 - \frac{nt^\beta}{Y} + \frac{(nt^\beta)^2}{2!Y^2} - \frac{(nt^\beta)^3}{3!Y^3} + \dots\Big)$$
$$= 1 - \frac{\Phi t^\beta}{1!}\frac{n}{(n-1)} + \frac{(\Phi t^\beta)^2}{2!}\frac{n^2}{(n-1)(n-2)} - \frac{(\Phi t^\beta)^3}{3!}\frac{n^3}{(n-1)(n-2)(n-3)} + \dots$$
$$\neq R(t)$$

Hence, $\hat{R}(t)$ is not unbiased for $R(t)$.

To verify that $\tilde{R}(t)$ is unbiased, consider $E(\tilde{R}(t)) = E\left(1 - \frac{t^\beta}{\sum_{i=1}^{n} t_i^\beta}\right)^{n-1}$.

But, $Y = \sum_{i=1}^{n} t_i^\beta$ and hence,

$$E(\tilde{R}(t)) = E\Big(1 - \frac{t^\beta}{Y}\Big)^{n-1}$$
$$= E\Big(1 - \frac{(n-1)t^\beta}{Y} + \frac{(n-1)(n-2)}{2!}\frac{(t^\beta)^2}{Y^2} - \frac{(n-1)(n-2)(n-3)}{3!}\frac{(t^\beta)^3}{Y^3} + \dots\Big)$$
$$= 1 - \Phi t^\beta + \frac{(\Phi t^\beta)^2}{2!} - \frac{(\Phi t^\beta)^3}{3!} + \dots$$
$$= R(t)$$

Thus, $\tilde{R}(t)$ is unbiased for $R(t)$.

Since $\hat{R}(t)$ is biased for $R(t)$ while $\tilde{R}(t)$ is unbiased for $R(t)$, the bias of $\hat{R}(t)$ is given by

Bias $(\hat{R}(t)) = E(\hat{R}(t)) - e^{-\Phi t^\beta} = E(\hat{R}(t)) - E(\tilde{R}(t))$ and hence is obtained as

$$\text{Bias}(\hat{R}(t)) = -\frac{\Phi t^\beta}{(n-1)} + \frac{(\Phi t^\beta)^2}{2!}\frac{(3n-2)}{(n-1)(n-2)} - \frac{(\Phi t^\beta)^3}{3!}\frac{(6n^2-11n+6)}{(n-1)(n-2)(n-3)} + \dots \tag{4.3.2}$$

Thus, the above bias can be found for the given sample failure data, by using the estimated values of $\hat{R}(t)$ and $\tilde{R}(t)$.

Hence, if $\mathscr{T} = \{t_1, t_2, \dots t_n\}$ is the given sample failure data set of size n, then the bias is obtained by taking the difference in the means of $\hat{R}(t)$ and $\tilde{R}(t)$ and is obtained as

$$\text{Bias}(\hat{R}(t)) = \frac{\sum\limits_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n} \tag{4.3.3}$$

Removing this bias from $\hat{R}(t)$, the Improved Estimator of $R(t)$ denoted by $\breve{R}(t)$, is obtained as $\breve{R}(t) = \hat{R}(t) - \text{Bias}(\hat{R}(t)) = \hat{R}(t) - \left(\frac{\sum\limits_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n} \right)$.

$$i.e.\breve{R}(t) = e^{-\left(\frac{nt^{\beta}}{\sum\limits_{i=1}^{n} t_i^{\beta}}\right)} - \left(\frac{\sum\limits_{t\epsilon\mathscr{T}} e^{-\left(\frac{nt^{\beta}}{\sum\limits_{i=1}^{n} t_i^{\beta}}\right)}}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}} \left(1 - \frac{t^{\beta}}{\sum\limits_{i=1}^{n} t_i^{\beta}}\right)^{n-1}}{n} \right) \tag{4.3.4}$$

In all the above calculations, t is any time instance. For a sample failure time data set $\mathscr{T}$, as given above, t is a member of $\mathscr{T}$.

4.4 COMPARISON OF ESTIMATES

The three estimators of reliability are to be compared by comparing the properties satisfied by them. The Improved Estimator of $R(t)$ is unbiased and sufficient, as it is obtained from MLE of $R(t)$, by removing the bias present in it. The only property to be checked thus, is the efficiency property. Since MLE of $R(t)$ is biased as shown above, while MVUE of $R(t)$ and Improved Estimator of $R(t)$ are unbiased, coefficient of variation is used as a measure of dispersion instead of the variance, as mentioned in Section 1.1 of Chapter 1. The estimate with the least value of the coefficient of variation is considered as the efficient estimator. The comparison is also done by considering the quartile coefficient of dispersion, as mentioned in Section 1.1 of Chapter 1. Even with this measure, the estimate with the least value of the quartile coefficient of dispersion is considered as the efficient estimator. For this purpose, the following case studies have been considered and the three estimates have been found. The coefficients of variation and the quartile coefficient of dispersion for these three estimates have also been obtained. For all the case studies, CV($\hat{R}(t)$), CV($\tilde{R}(t)$) and CV($\breve{R}(t)$) are respectively obtained using

$$\text{CV}(\hat{R}(t)) = \frac{S_{\hat{R}(t)}}{\bar{\hat{R}}(t)} \tag{4.4.1}$$

$$\text{CV}(\tilde{R}(t)) = \frac{S_{\tilde{R}(t)}}{\bar{\tilde{R}}(t)} \tag{4.4.2}$$

$$\mathrm{CV}(\breve{R}(t)) = \frac{S_{\breve{R}(t)}}{\overline{\breve{R}}(t)} \tag{4.4.3}$$

Here, the sample variances $S^2_{\hat{R}(t)}$, $S^2_{\tilde{R}(t)}$ and $S^2_{\breve{R}(t)}$ are respectively obtained using

$$S^2_{\hat{R}(t)} = \sum_{t \epsilon \mathscr{T}} \frac{\left(\hat{R}(t) - \overline{\hat{R}}(t)\right)^2}{(n-1)} \tag{4.4.4}$$

$$S^2_{\tilde{R}(t)} = \sum_{t \epsilon \mathscr{T}} \frac{\left(\tilde{R}(t) - \overline{\tilde{R}}(t)\right)^2}{(n-1)} \tag{4.4.5}$$

$$S^2_{\breve{R}(t)} = \sum_{t \epsilon \mathscr{T}} \frac{\left(\breve{R}(t) - \overline{\breve{R}}(t)\right)^2}{(n-1)} \tag{4.4.6}$$

Further, the sample means $\overline{\hat{R}}(t)$, $\overline{\tilde{R}}(t)$, and $\overline{\breve{R}}(t)$ are respectively obtained using

$$\overline{\hat{R}}(t) = \frac{\sum_{t \epsilon \mathscr{T}} \hat{R}(t)}{n} \tag{4.4.7}$$

$$\overline{\tilde{R}}(t) = \frac{\sum_{t \epsilon \mathscr{T}} \tilde{R}(t)}{n} \tag{4.4.8}$$

$$\overline{\breve{R}}(t) = \frac{\sum_{t \epsilon \mathscr{T}} \breve{R}(t)}{n} \tag{4.4.9}$$

As mentioned in literature above, some computer methodologies have also been used in generating failure data, such as the one by Subburaj Ramasamy and Indhurani Lakshmanan (Ramasamy and Lakshmanan (2017)) used machine learning approach, which also compared the method with practical failure data set. Herein, few practical data sets have been used.

Case study 1: On-Line Data Entry IBM Software Package

The data reported by Ohba (Ohba (1984)) are recorded from testing an on-line data entry software package developed at IBM. There are 15 failures, with failure times as indicated in Table 4.4.1.

Table 4.4.1 On-Line Data Entry IBM Software Package

Failure Number	1	2	3	4	5	6	7	8	9	10
Failure Time	10	19	32	43	58	70	88	103	125	150
Failure Number	11	12	13	14	15					
Failure Time	169	199	231	256	296					

Table 4.4.2 denotes the MLE, MVUE and the Improved Estimator of reliability functions for Weibull class models. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using

Table 4.4.2 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Weibull Model (Case study 1)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$\text{SD}_{\hat{R}(t)}$	$\text{SD}_{\tilde{R}(t)}$	$\text{SD}_{\breve{R}(t)}$
1	10	0.99563	0.99592	1.00112	0.1906	0.1861	0.19057
2	19	0.98432	0.98535	0.98981	0.1808	0.1771	0.18083
3	32	0.95617	0.95897	0.96166	0.1577	0.1555	0.15768
4	43	0.92226	0.92706	0.92775	0.1319	0.1314	0.1319
5	58	0.86310	0.87102	0.86859	0.0924	0.0939	0.09242
6	70	0.80699	0.81743	0.81248	0.0615	0.0639	0.06146
7	88	0.7125	0.72619	0.71803	0.0235	0.0261	0.02355
8	103	0.62858	0.64391	0.63407	0.0048	0.0063	0.00483
9	125	0.50469	0.52036	0.51018	0.003	0.002	0.00296
10	150	0.37355	0.38651	0.37904	0.0344	0.0317	0.03442
11	169	0.28652	0.29579	0.29201	0.0743	0.0722	0.07429
12	199	0.17673	0.17926	0.18222	0.1462	0.1485	0.14619
13	231	0.09678	0.09355	0.10227	0.2137	0.2219	0.21373
14	256	0.05680	0.05125	0.06229	0.2523	0.2635	0.25229
15	296	0.02161	0.01603	0.02710	0.2889	0.3009	0.28888

the values obtained in Table 4.4.2 in equations (4.4.4) to (4.4.9), we get,
$\overline{\hat{R}}(t) = 0.559090$; $S^2_{\hat{R}(t)} = 0.13257$; $\overline{\tilde{R}}(t) = 0.564578$; $S^2_{\tilde{R}(t)} = 0.13435$;
$\overline{\breve{R}}(t) = 0.564578$; $S^2_{\breve{R}(t)} = 0.1326$.
Hence, using (4.3.3), the bias in $\hat{R}(t)$ is obtained as
Bias$(\hat{R}(t)) = 0.559090 - 0.564578 = -0.005488$.

Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.005488) = \hat{R}(t) + 0.005488$.

Thus, using equations (4.4.1) to (4.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.6512, CV($\tilde{R}(t)$)=0.6492 and **CV($\breve{R}(t)$)=0.6449**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.1767 and Q_3=0.9223. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.6784. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.1793 and Q_3=0.9271. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.6759. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.1822 and Q_3=0.9277. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.6717**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 4.4.1.

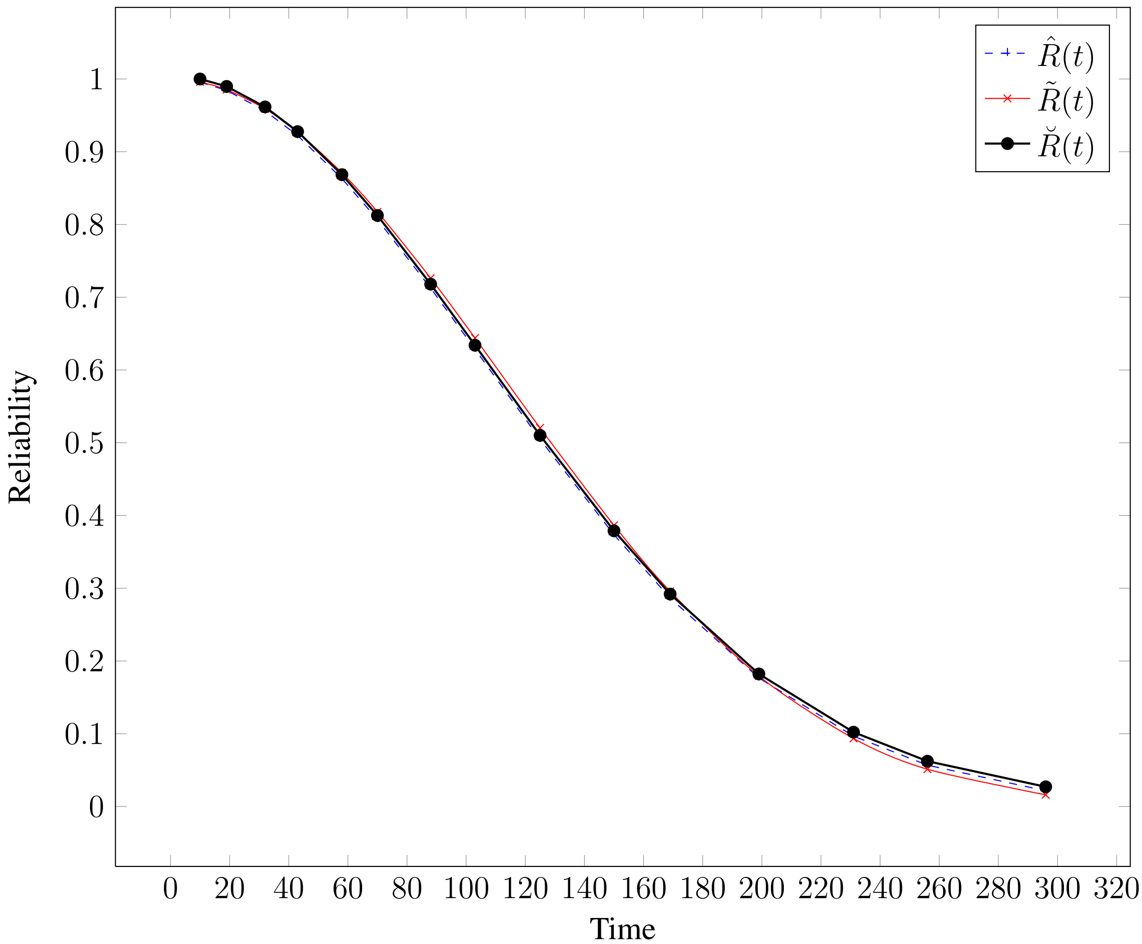

Figure 4.4.1 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Weibull Model (Case study 1)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 2: Nuclear Power Agency

A nuclear power agency uses a computer-based monitoring system for its reactors. The operating system for the computer is employed for this and other applications in an estimated 5000 installations throughout the world. A total of 17 failures have occurred with failure times as listed in Table 4.4.3 (Musa et al. (1991)).

Table 4.4.3 Nuclear Power Agency

Failure number	1	2	3	4	5	6
Failure time	932	4035	4696	4893	6369	6524
Failure number	7	8	9	10	11	12
Failure time	7882	8170	9339	10400	10542	11036
Failure number	13	14	15	16	17	
Failure time	11696	11905	12266	12954	14000	

Table 4.4.4 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in Table 4.4.4 in equations (4.4.4) to (4.4.9), we get,

$\overline{\hat{R}}(t) = 0.44805$; $S^2_{\hat{R}(t)} = 0.0752$; $\overline{\tilde{R}}(t) = 0.45551$; $S^2_{\tilde{R}(t)} = 0.0764$;

$\overline{\breve{R}}(t) = 0.45551$; $S^2_{\breve{R}(t)} = 0.0752$.

Hence, using (4.3.3), the bias in $\hat{R}(t)$ is obtained as

$\text{Bias}(\hat{R}(t)) = 0.44805 - 0.45551 = -0.0075.$

Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as

$\breve{R}(t) = \hat{R}(t) - (-0.0075) = \hat{R}(t) + 0.0075.$

Thus, using equations (4.4.1) to (4.4.3), the coefficient of variation of the three estimators, are respectively obtained as

CV($\hat{R}(t)$)=0.6120, CV($\tilde{R}(t)$)=0.6066 and **CV($\boldsymbol{\breve{R}(t)}$)=0.6019**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

Table 4.4.4 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Weibull Model (Case study 2)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$SD_{\hat{R}(t)}$	$SD_{\tilde{R}(t)}$	$SD_{\breve{R}(t)}$
1	932	0.9902	0.9908	0.9976	0.2939	0.2865	0.2939
2	4035	0.8311	0.8394	0.8386	0.1467	0.1474	0.1467
3	4696	0.7784	0.7885	0.7858	0.1091	0.1109	0.1091
4	4893	0.7618	0.7725	0.7693	0.0985	0.1005	0.0985
5	6369	0.6307	0.6442	0.6382	0.0334	0.0356	0.0334
6	6524	0.6166	0.6302	0.624	0.0284	0.0305	0.0284
7	7882	0.4937	0.5074	0.5011	0.0021	0.0027	0.0021
8	8170	0.4684	0.4818	0.4759	0.0004	0.0007	0.0004
9	9339	0.3712	0.3825	0.3787	0.0059	0.0053	0.0059
10	10400	0.2926	0.301	0.3001	0.0242	0.0239	0.0242
11	10542	0.2829	0.2909	0.2904	0.0273	0.0271	0.0273
12	11036	0.2506	0.2571	0.2581	0.039	0.0394	0.039
13	11696	0.2113	0.2157	0.2188	0.056	0.0575	0.056
14	11905	0.1998	0.2035	0.2073	0.0616	0.0635	0.0616
15	12266	0.181	0.1835	0.1884	0.0713	0.074	0.0713
16	12954	0.1486	0.1491	0.1561	0.0897	0.0939	0.0897
17	14000	0.1079	0.1058	0.1153	0.1157	0.1223	0.1157

Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.2056 and Q_3=0.6963. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.5440. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2096 and Q_3=0.7084. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.5434. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2131 and Q_3=0.7038. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.5352**.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 4.4.2.

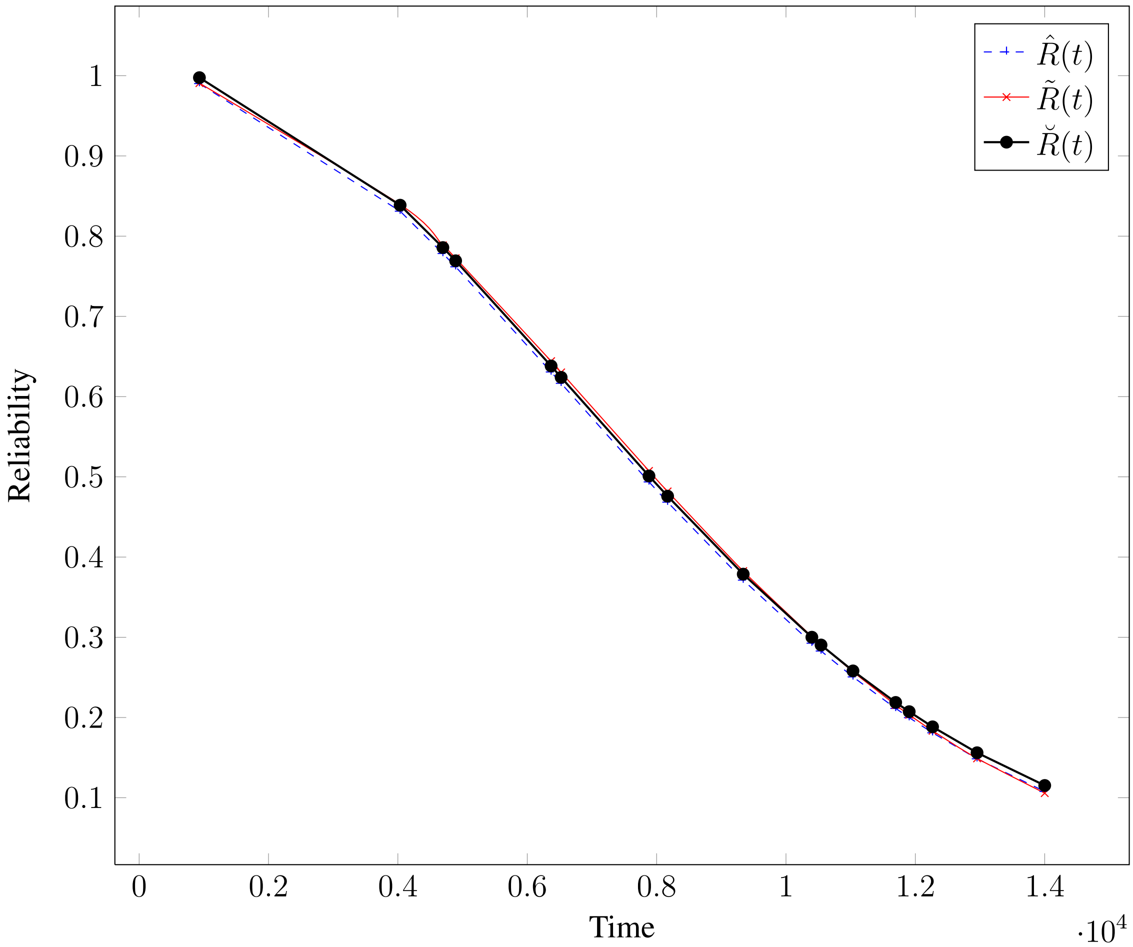

Figure 4.4.2 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Weibull Model (Case study 2)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 3: Failure data set of Lyu

The failure time data for 10 failures obtained by Lyu (Lyu (2004)) are given in Table 4.4.5.

Table 4.4.5 Failure data set of Lyu

Failure number	1	2	3	4	5	6	7	8	9	10
Failure time	7	18	26	36	51	73	93	118	146	181

Table 4.4.6 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in

Table 4.4.6 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Weibull Model (Case study 3)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$\text{SD}_{\hat{R}(t)}$	$\text{SD}_{\tilde{R}(t)}$	$\text{SD}_{\breve{R}(t)}$
1	7	0.9944	0.9949	1.00	0.17942	0.17342	0.17942
2	18	0.9634	0.9669	0.9711	0.15415	0.15089	0.15415
3	26	0.9252	0.9321	0.9329	0.12559	0.12506	0.12559
4	36	0.8615	0.8735	0.8692	0.0845	0.08705	0.0845
5	51	0.7414	0.7608	0.7491	0.0291	0.03322	0.0291
6	73	0.5417	0.5659	0.5494	0.00085	0.00016	0.00085
7	93	0.3697	0.3894	0.3774	0.04043	0.03577	0.04043
8	118	0.2015	0.2078	0.2092	0.13636	0.13742	0.13636
9	146	0.0861	0.0795	0.0938	0.23493	0.24901	0.23493
10	181	0.0231	0.0142	0.0308	0.3	0.31849	0.3

Table 4.4.6 in equations (4.4.4) to (4.4.9), we get,
$\overline{\hat{R}}(t) = 0.570797$; $S^2_{\hat{R}(t)} = 0.1428$; $\overline{\tilde{R}}(t) = 0.578506$; $S^2_{\tilde{R}(t)} = 0.1456$;
$\overline{\breve{R}}(t) = 0.578497$; $S^2_{\breve{R}(t)} = 0.1428$.
Hence, using (4.3.3), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = 0.570797 - 0.578506 = -0.0077$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.0077) = \hat{R}(t) + 0.0077$.

Thus, using equations (4.4.1) to (4.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.6620, CV($\tilde{R}(t)$)=0.6596 and **CV($\boldsymbol{\breve{R}(t)}$)=0.6532**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.2015 and Q_3=0.9252. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.6423. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2078 and Q_3=0.9321. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.6354. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2092 and Q_3=0.9329. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\boldsymbol{\breve{R}(t)}$)=0.6337**.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 4.4.3.

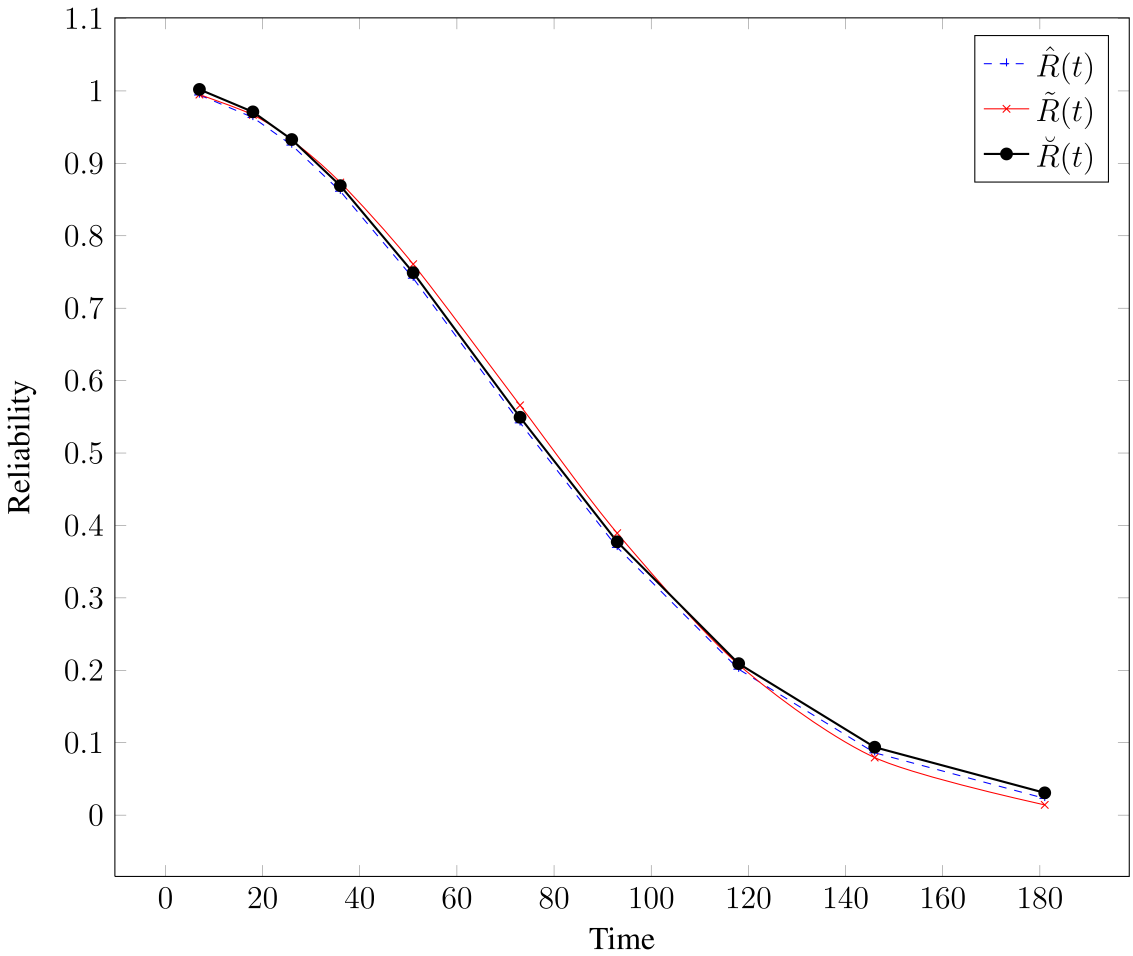

Figure 4.4.3 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Weibull Model (Case study 3)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Table 4.4.7 shows the consolidated values of the coefficient of variation (CV) of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for all the three case studies.
Table 4.4.8 shows the consolidated values of the quartile coefficient of dispersion (QD) of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for all the three case studies. From Tables 4.4.7 and 4.4.8, it is observed that, the Improved Estimator ($\breve{R}(t)$) has the least values of the coefficient of

Table 4.4.7 Weibull Models (Consolidated1)

Case study	CV($\hat{R}(t)$)	CV($\tilde{R}(t)$)	CV($\breve{R}(t)$)
1	0.6512	0.6492	**0.6449**
2	0.6120	0.6066	**0.6019**
3	0.6620	0.6596	**0.6532**

Table 4.4.8 Weibull Models (Consolidated2)

Case study	QD($\hat{R}(t)$)	QD($\tilde{R}(t)$)	QD($\breve{R}(t)$)
1	0.6784	0.6759	**0.6717**
2	0.5440	0.5434	**0.5352**
3	0.6423	0.6354	**0.6337**

variation and quartile coefficient of dispersion than those of MLE ($\hat{R}(t)$) and MVUE ($\tilde{R}(t)$) in all the three case studies, which means that $\breve{R}(t)$ is more efficient than $\hat{R}(t)$ and $\tilde{R}(t)$.

However, to choose the best estimator among the three estimates, the desirable properties of good estimators as mentioned in Section 1.1 of Chapter 1 are to be considered.

Unbiasedness of $\breve{R}(t)$: It has been shown above that $\hat{R}(t)$ is biased for $R(t)$, while $\tilde{R}(t)$ is unbiased for $R(t)$. Since the Improved Estimators are obtained from MLEs just by removing the bias present in the MLEs, they satisfy the unbiasedness property. Thus, $\breve{R}(t)$ is unbiased for $R(t)$.

Sufficiency of $\breve{R}(t)$: Improved Estimator is a function of MLE, which is sufficient. Since any function of sufficient estimator is also sufficient, Improved Estimator is sufficient.

Now, to compare the biased estimator $\hat{R}(t)$ with unbiased estimators $\tilde{R}(t)$ and $\breve{R}(t)$, the coefficient of variation and the quartile coefficient of dispersion are considered as measures of dispersion to check the efficiency property. The sample results of comparison of coefficients of variation and the quartile coefficients of dispersion for the three estimators indicate that Improved Estimator has least values of coefficient of variation and the quartile coefficient of dispersion as compared to those of MLE and MVUE of $R(t)$, which indicates that the Improved estimators are efficient compared to MLE and MVUE.

Thus, by referring to Table 1.1.1 of Chapter 1, Table 4.4.9 provides the statistical properties satisfied by MLE, MVUE and the Improved Estimator of reliability for Weibull class models.

It can be seen from Table 4.4.9 that the Improved Estimator satisfies maximum number of properties of estimators as compared to MLE and MVUE of $R(t)$. Hence, it can be inferred that the estimate of reliability obtained using the Improved Estimator, is more

Table 4.4.9 Weibull class models- Properties satisfied by estimators of reliability

	Unbiased	**Sufficient**	**Efficient**
MLE	No	Yes	No
MVUE	Yes	Yes	No
Improved Estimator	**Yes**	**Yes**	**Yes**

efficient than those estimated using the methods of MLE and MVUE.
Thus, it can be concluded that $\breve{R}(t)$ gives more accurate value of reliability than $\hat{R}(t)$ and $\tilde{R}(t)$, for Weibull class software reliability models.

Chapter 5

GAMMA CLASS MODELS

Gamma class models are the software reliability models, wherein the failure times (T) are assumed to have Gamma distribution with probability density function, given by

$$f(t) = \Phi^2 t e^{-\Phi t}, \quad t > 0 \tag{5.0.1}$$

where Φ is the failure rate. This model is due to Yamada, Ohba and Osaki (Yamada et al. (1983)).
Some of the work related to estimation of reliability from literature are provided below: The use of inverse sampling in estimating the reliability for any model and so also Gamma model was proposed by Balwant Singh et. al (Singh et al. (1997)). The intention was to determine the testing time and also to decide whether or not to accept the software. A combined parameter estimation procedure using Expectation maximization algorithm and heuristic solution method was developed for Gamma class models by Hiroyuki Okamura et. al (Okamura et al. (2007)). However, the method has not been used in estimation of reliability. Moreover, by considering the failure time distribution, many statistical procedures of estimation can be used in estimating the reliability and for comparing them statistically.

If T has an Gamma distribution with parameters 2 and Φ, it is then denoted as $T \sim G(2, \Phi)$. The reliability at time t, denoted by$R(t)$, is obtained as

$$R(t) = P(T > t) = \int_t^\infty f(t)dt = \int_t^\infty \Phi^2 t e^{-\Phi t} dt = e^{-\Phi t}(t\Phi + 1) \tag{5.0.2}$$

In the following sections, the estimates of this reliability are obtained using the methods of MLE and MVUE.

5.1 MLE OF $R(t)$

Using the invariance property satisfied by the MLEs, the MLE of $R(t)$, denoted by $\hat{R}(t)$ is obtained as

$$\hat{R}(t) = e^{-\hat{\Phi}t}(t\hat{\Phi} + 1) \tag{5.1.1}$$

where $\hat{\Phi}$ is the MLE of Φ.

To find the MLE of Φ: Let $(T_1, T_2, \ldots T_n)$ be a sample of size n from Gamma distribution as given in (5.0.1). Then, the likelihood function of this sample is given by

$$L = \prod_{i=1}^{n} f(t_i) = \Phi^{2n} e^{-\Phi \sum_{i=1}^{n} t_i} \prod_{i=1}^{n} t_i \tag{5.1.2}$$

Using the principle of calculus, maximizing this likelihood function, the MLE of Φ, denoted by $\hat{\Phi}$, is obtained as the solution of $\dfrac{\partial \ln L}{\partial \Phi} = 0$, with $\dfrac{\partial^2 \ln L}{\partial \Phi^2} < 0$.
This gives $\dfrac{2n}{\Phi} - \sum_{i=1}^{n} t_i = 0$, from which, the MLE of Φ is obtained as

$$\hat{\Phi} = \frac{2n}{\sum_{i=1}^{n} t_i} \tag{5.1.3}$$

Using (5.1.3) in (5.1.1), the MLE of $R(t)$ is obtained as

$$\hat{R}(t) = e^{-\frac{2nt}{\sum_{i=1}^{n} t_i}} \left(\frac{2nt}{\sum_{i=1}^{n} t_i} + 1 \right) \tag{5.1.4}$$

5.2 MVUE OF $R(t)$

To find the MVUE of $R(t)$, consider a function of the random variable T_1, given by

$$U(t_1) = \begin{cases} 1 & \text{if } t_1 > t \\ 0 & \text{otherwise} \end{cases}$$

Then, $E[U(t_1)] = 1.P(T_1 > t) + 0.P(T_1 \leq t) = P(T_1 > t) = R(t)$.
Thus, $U(t_1)$ is unbiased for $R(t)$.

Complete Sufficient Estimator of Φ: Applying the factorization theorem (as explained in Chapter 1) to the likelihood function given in (5.1.2), it can be seen that the likelihood function depends on Φ and t_i, only through the value of $\sum_{i=1}^{n} t_i$. Thus, $\sum_{i=1}^{n} t_i$ is the sufficient estimator of the parameter Φ.

Also, since each $T_i \sim G(2, \Phi)$, $\sum_{i=1}^{n} T_i \sim G(2n, \Phi)$ (Section 3 of Appendix A). Further, by Result 3 of Appendix A, the estimator $\sum_{i=1}^{n} t_i$ is also the complete statistic and hence, $\sum_{i=1}^{n} t_i$ is the complete sufficient estimator of Φ.

Further, $U(t_1)$ is an unbiased estimator of $R(t)$ and $R(t)$ is a function of Φ, as given in (5.0.2). Hence, by Theorem 1 of Chapter 1, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = E(U(t_1)|\sum_{i=1}^{n} t_i) = \int_{t}^{\infty} f(t_1|\sum_{i=1}^{n} t_i) dt_1$$

where $f(t_1|\sum_{i=1}^{n} t_i)$ denotes the conditional pdf of T_1 given $\sum_{i=1}^{n} T_i$ and is given by $f(t_1|\sum_{i=1}^{n} t_i) = \frac{g(t_1, \sum_{i=1}^{n} t_i)}{h(\sum_{i=1}^{n} t_i)}$, where $g(t_1, \sum_{i=1}^{n} t_i)$ denotes the joint pdf of T_1 and $\sum_{i=1}^{n} T_i$. $h(\sum_{i=1}^{n} t_i)$ denotes the marginal pdf of $\sum_{i=1}^{n} T_i$.

Hence, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_{t}^{\infty} \frac{g(t_1, \sum_{i=1}^{n} t_i)}{h(\sum_{i=1}^{n} t_i)} dt_1 \tag{5.2.1}$$

Since each $T_i \sim G(2, \Phi)$, $\sum_{i=1}^{n} T_i \sim G(2n, \Phi)$ (Section 3 of Appendix A). Hence, the pdf of $\sum_{i=1}^{n} T_i$ is obtained as

$$h(\sum_{i=1}^{n} t_i) = \frac{\Phi^{2n}}{\Gamma(2n)} e^{-\Phi \sum_{i=1}^{n} t_i} (\sum_{i=1}^{n} t_i)^{2n-1} \tag{5.2.2}$$

To find the pdf $g(t_1, \sum_{i=1}^{n} t_i)$, split the sample $(T_1, T_2, T_3, \ldots T_n)$ into two samples as T_1 of size one and $(T_2, T_3, T_4, \ldots T_n)$ of size $(n-1)$.

Since T_1 and $\sum_{i=2}^{n} T_i$ are independent, the joint pdf of T_1 and $\sum_{i=2}^{n} T_i$ is obtained as

$$g(t_1, \sum_{i=2}^{n} t_i) = f(t_1) \cdot h(\sum_{i=2}^{n} t_i),$$

where $f(t_1)$ and $h(\sum_{i=2}^{n} t_i)$ denote the pdfs of T_1 and $\sum_{i=2}^{n} T_i$ respectively.

Since $T_1 \sim G(2, \Phi)$, the pdf of T_1 is given by $f(t_1) = \Phi^2 t_1 e^{-\Phi t_1}$.

Also, $\sum_{i=2}^{n} T_i \sim G\big(2(n-1), \Phi\big) = G(2n-2, \Phi)$ and hence its pdf is obtained as $h(\sum_{i=2}^{n} t_i) = \frac{\Phi^{2n-2}}{\Gamma(2n-2)} (\sum_{i=2}^{n} t_i)^{2n-3} e^{-\Phi \sum_{i=2}^{n} t_i}$.

Hence, the joint pdf of T_1 and $\sum_{i=2}^{n} T_i$ is obtained as

$g(t_1, \sum_{i=2}^{n} t_i) = f(t_1) \cdot h(\sum_{i=2}^{n} t_i) = \Phi^2 t_1 e^{-\Phi t_1} \frac{\Phi^{2n-2}}{\Gamma(2n-2)} (\sum_{i=2}^{n} t_i)^{2n-3} e^{-\Phi \sum_{i=2}^{n} t_i}$,

which simplifies to $g(t_1, \sum_{i=2}^{n} t_i) = \frac{e^{-\Phi \sum_{i=1}^{n} t_i} \Phi^{2n}}{\Gamma(2n-2)} t_1 (\sum_{i=2}^{n} t_i)^{2n-3}$.

Considering the transformation $\sum_{i=1}^{n} T_i = T_1 + \sum_{i=2}^{n} T_i$ and noting that the modulus of the Jacobian of the inverse transformation is one (Appendix B), the joint pdf of T_1 and $\sum_{i=1}^{n} T_i$ is obtained as

$$g(t_1, \sum_{i=1}^{n} t_i) = \frac{e^{-\Phi \sum_{i=1}^{n} t_i} \Phi^{2n}}{\Gamma(2n-2)} t_1 (\sum_{i=1}^{n} t_i - t_1)^{2n-3} \tag{5.2.3}$$

Substituting (5.2.2) and (5.2.3) in (5.2.1), the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_{t}^{\infty} \frac{\frac{e^{-\Phi\sum_{i=1}^{n} t_i}\Phi^{2n}}{\Gamma(2n-2)} t_1 (\sum_{i=1}^{n} t_i - t_1)^{2n-3}}{\frac{e^{-\Phi\sum_{i=1}^{n} t_i}\Phi^{2n}}{\Gamma(2n)} (\sum_{i=1}^{n} t_i)^{2n-1}} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_{t}^{\infty} \frac{\Gamma(2n)t_1}{\Gamma(2n-2)} \frac{(\sum_{i=1}^{n} t_i - t_1)^{2n-3}}{(\sum_{i=1}^{n} t_i)^{2n-1}} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_{t}^{\infty} \frac{(2n-1)(2n-2)t_1}{(\sum_{i=1}^{n} t_i)^2} \frac{(\sum_{i=1}^{n} t_i - t_1)^{2n-3}}{(\sum_{i=1}^{n} t_i)^{2n-3}} dt_1$$

This simplifies to

$$\tilde{R}(t) = \int_{t}^{\infty} (2n-1)(2n-2) \left(1 - \frac{t_1}{\sum_{i=1}^{n} t_i}\right)^{2n-3} \frac{t_1}{(\sum_{i=1}^{n} t_i)^2} dt_1$$

Noting that this integral converges if $t_1 < \sum_{i=1}^{n} t_i$, the MVUE of reliability is obtained as

$$\tilde{R}(t) = \int_{t}^{\sum_{i=1}^{n} t_i} (2n-1)(2n-2) \left(1 - \frac{t_1}{\sum_{i=1}^{n} t_i}\right)^{2n-3} \frac{t_1}{(\sum_{i=1}^{n} t_i)^2} dt_1$$

The term inside the definite integral is the conditional pdf $f(t_1 | \sum_{i=1}^{n} t_i)$, which is the conditional pdf or t_1 given that $\sum_{i=1}^{n} t_i$ has occurred. i.e., this conditional pdf is evaluated under $\sum_{i=1}^{n} t_i$. Thus, for a given $\sum_{i=1}^{n} t_i$, the integral depends only on t_1.

Hence, taking $\sum_{i=1}^{n} t_i = s$, the integral becomes

$$\tilde{R}(t) = \int_{t}^{s} (2n-1)(2n-2)\left(1 - \frac{t_1}{s}\right)^{2n-3} \frac{t_1}{s^2} dt_1$$

$$\text{i.e. }, \tilde{R}(t) = \frac{(2n-1)(2n-2)}{s^2} \int_{t}^{s} t_1 \left(1 - \frac{t_1}{s}\right)^{2n-3} dt_1$$

Integrating by parts, by taking t_1 as the first function and $\left(1 - \frac{t_1}{s}\right)^{2n-3}$ as the second function,

$$\tilde{R}(t) = \frac{(2n-1)(2n-2)}{s^2}\left(\frac{t_1(-s)}{(2n-2)}\left(1 - \frac{t_1}{s}\right)^{2n-2}\Bigg|_{t}^{s} + \int_{t}^{s} \frac{s}{(2n-2)}(1 - \frac{t_1}{s})^{2n-2} dt_1 \right).$$

The subsequent steps of simplification are provided below:

$$\begin{aligned}
\tilde{R}(t) &= \frac{(2n-1)(2n-2)}{s^2}\left(\frac{ts}{(2n-2)}\left(1 - \frac{t}{s}\right)^{2n-2} - \frac{s}{(2n-2)}\left(\frac{s(1 - \frac{t_1}{s})^{2n-1}}{(2n-1)}\right)\Bigg|_{t}^{s} \right) \\
&= \frac{(2n-1)(2n-2)}{s^2}\left(\frac{ts}{(2n-2)}\left(1 - \frac{t}{s}\right)^{2n-2} + \frac{s^2}{(2n-1)(2n-2)}\left(1 - \frac{t}{s}\right)^{2n-1} \right) \\
&= \frac{(2n-1)t}{s}\left(1 - \frac{t}{s}\right)^{2n-2} + \left(1 - \frac{t}{s}\right)^{2n-1} \\
&= \left(1 - \frac{t}{s}\right)^{2n-2}\left(\frac{(2n-1)t}{s} + \left(1 - \frac{t}{s}\right)\right) = \left(1 - \frac{t}{s}\right)^{2n-2}\left(\frac{2nt}{s} - \frac{t}{s} + 1 - \frac{t}{s}\right) \\
&= \left(1 - \frac{t}{s}\right)^{2n-2}\left(\frac{2nt}{s} - \frac{2t}{s} + 1\right) = \left(1 - \frac{t}{s}\right)^{2n-2}\left(\frac{(2n-2)t}{s} + 1\right)
\end{aligned}$$

Replacing s by $\sum_{i=1}^{n} t_i$ again, the MVUE of reliability is obtained as

$$\tilde{R}(t) = \begin{cases} \left(1 - \frac{t}{\sum_{i=1}^{n} t_i}\right)^{2n-2} \left(\frac{(2n-2)t}{\sum_{i=1}^{n} t_i} + 1\right) & \text{if } t < \sum_{i=1}^{n} t_i \\ 0 & \text{otherwise} \end{cases} \tag{5.2.4}$$

5.3 IMPROVED ESTIMATOR OF $R(t)$

The reliability function at time t for Gamma class models is given by

$$R(t) = e^{-\Phi t}(\Phi t + 1) = 1 - \frac{(\Phi t)^2}{2} + \frac{(\Phi t)^3}{3} - \dots \tag{5.3.1}$$

$\hat{R}(t)$ and $\tilde{R}(t)$ are unbiased for $R(t)$, if (i) $E(\hat{R}(t)) = R(t)$ and (ii) $E(\tilde{R}(t)) = R(t)$ respectively.

To check whether $\hat{R}(t)$ is unbiased or not, consider $E(\hat{R}(t)) = E\left(e^{-\frac{2nt}{\sum_{i=1}^{n} t_i}}\left(\frac{2nt}{\sum\limits_{i=1}^{n} t_i}+1\right)\right)$.

Taking $Y = \sum\limits_{i=1}^{n} t_i$, we have, $E(\hat{R}(t)) = E\left(e^{-\{\frac{2nt}{Y}\}}\left(\frac{2nt}{Y} + 1\right)\right)$.

Also, since $Y \sim G(2n, \Phi)$, (Section 3 of Appendix A), it can be observed that,

$$\begin{aligned}
E\left(\frac{1}{Y}\right) &= \int_0^\infty \frac{1}{y} \frac{\Phi^{2n}}{\Gamma(2n)} e^{-y\Phi} y^{2n-1} dy \\
&= \frac{\Phi^{2n}}{\Gamma(2n)} \int_0^\infty e^{-\Phi y} y^{2n-2} dy \\
&= \frac{\Phi^{2n}}{\Gamma(2n)} \frac{\Gamma(2n-1)}{\Phi^{2n-1}} \qquad \text{(Result 4 of Appendix A)} \\
&= \frac{\Phi\Gamma(2n-1)}{(2n-1)\Gamma(2n-1)} = \frac{\Phi}{2n-1}.
\end{aligned}$$

$$\begin{aligned}
\text{Similarly, } E\left(\frac{1}{Y^2}\right) &= \int_0^\infty \frac{1}{y^2} \frac{\Phi^{2n}}{\Gamma(2n)} e^{-y\Phi} y^{2n-1} dy \\
&= \frac{\Phi^{2n}}{\Gamma(2n)} \int_0^\infty e^{-\Phi y} y^{2n-3} dy \\
&= \frac{\Phi^{2n}}{\Gamma(2n)} \frac{\Gamma(2n-2)}{\Phi^{2n-2}} \qquad \text{(Result 4 of Appendix A)} \\
&= \frac{\Phi^2\Gamma(2n-2)}{(2n-1)(2n-2)\Gamma(2n-2)} = \frac{\Phi^2}{(2n-1)(2n-2)}.
\end{aligned}$$

$$
\begin{aligned}
E\Big(\frac{1}{Y^3}\Big) &= \int_0^\infty \frac{1}{y^3}\frac{\Phi^{2n}}{\Gamma(2n)}e^{-y\Phi}y^{2n-1}dy \\
&= \frac{\Phi^{2n}}{\Gamma(2n)}\int_0^\infty e^{-\Phi y}y^{2n-4}dy \\
&= \frac{\Phi^{2n}}{\Gamma(2n)}\frac{\Gamma(2n-3)}{\Phi^{2n-3}} \qquad \text{(Result 4 of Appendix A)} \\
&= \frac{\Phi^3\Gamma(2n-3)}{(2n-1)(2n-2)(2n-3)\Gamma(2n-3)} \\
&= \frac{\Phi^3}{(2n-1)(2n-2)(2n-3)}
\end{aligned}
$$

and so on.

$$
\begin{aligned}
\text{Hence, } E(\hat{R}(t)) &= E\Bigg(\Big(1-\frac{2nt}{Y}+\frac{(2nt)^2}{2!Y^2}-\frac{(2nt)^3}{3!Y^3}+\dots\Big)\Big(\frac{2nt}{Y}+1\Big)\Bigg) \\
&= 1-\frac{(\Phi t)^2}{2}\frac{(2n)^2}{(2n-1)(2n-2)} \\
&\qquad\qquad + \frac{(\Phi t)^3}{3}\frac{(2n)^3}{(2n-1)(2n-2)(2n-3)} - \dots \\
&\neq R(t)
\end{aligned}
$$

Hence, $\hat{R}(t)$ is not unbiased for $R(t)$.

Now, to verify that $\tilde{R}(t)$ is unbiased, consider

$$
\begin{aligned}
E(\tilde{R}(t)) &= E\Bigg(\Bigg(1-\frac{t}{\sum_{i=1}^{n} t_i}\Bigg)^{2n-2}\Bigg(\frac{(2n-2)t}{\sum_{i=1}^{n} t_i}+1\Bigg)\Bigg) \\
&= E\Bigg(\Big(1-\frac{t}{Y}\Big)^{2n-2}\Big(\frac{(2n-2)t}{Y}+1\Big)\Bigg)
\end{aligned}
$$

where $Y = \sum_{i=1}^{n} t_i$.

$$\text{So, } E(\tilde{R}(t)) = E\Bigg(1 - \frac{(2n-1)(2n-2)}{2}(\frac{t}{Y})^2 + \frac{(2n-1)(2n-2)(2n-3)}{3}(\frac{t}{Y})^3 - \dots\Bigg)$$

$$= 1 - \frac{(\Phi t)^2}{2} + \frac{(\Phi t)^3}{3} - \dots$$

$$= R(t)$$

Hence, $\tilde{R}(t)$ is unbiased for $R(t)$.

Since $\hat{R}(t)$ is biased for $R(t)$ and $\tilde{R}(t)$ is unbiased for $R(t)$, the bias of $\hat{R}(t)$ is given by $\text{Bias}(\hat{R}(t)) = E(\hat{R}(t)) - R(t) = E(\hat{R}(t)) - E(\tilde{R}(t))$ and hence is obtained as

$$\text{Bias}(\hat{R}(t)) = -\frac{(\Phi t)^2}{2}\Big(\frac{(2n)^2}{(2n-1)(2n-2)} - 1\Big) + \frac{(\Phi t)^3}{3}\Big(\frac{(2n)^3}{(2n-1)(2n-2)(2n-3)} - 1\Big) - \dots \tag{5.3.2}$$

This bias can be found for the given sample failure data, by using the estimated values of $\hat{R}(t)$ and $\tilde{R}(t)$.

Hence, if $\mathscr{T} = \{t_1, t_2, \dots t_n\}$ is the given sample failure data of size n, then the estimated bias is obtained by taking the difference in the means of $\hat{R}(t)$ and $\tilde{R}(t)$ and is given by

$$\text{Bias}(\hat{R}(t)) = \frac{\sum_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} - \frac{\sum_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n} \tag{5.3.3}$$

Removing this bias from $\hat{R}(t)$, the Improved Estimator of $R(t)$ denoted by $\breve{R}(t)$, is obtained as $\breve{R}(t) = \hat{R}(t) - \text{Bias}(\hat{R}(t)) = \hat{R}(t) - \left(\frac{\sum_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} - \frac{\sum_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n}\right)$.

$$i.e., \breve{R}(t) = e^{-\left(\frac{2nt}{\sum\limits_{i=1}^{n} t_i}\right)}\left(\frac{2nt}{\sum\limits_{i=1}^{n} t_i} + 1\right) - \left(\frac{\sum\limits_{t\epsilon\mathscr{T}} e^{-\left(\frac{2nt}{\sum\limits_{i=1}^{n} t_i}\right)}\left(\frac{2nt}{\sum\limits_{i=1}^{n} t_i} + 1\right)}{n}\right.$$

$$\left. - \frac{\sum\limits_{t\epsilon\mathscr{T}}\left(1 - \frac{t}{\sum\limits_{i=1}^{n} t_i}\right)^{2n-2}\left(\frac{(2n-2)t}{\sum\limits_{i=1}^{n} t_i} + 1\right)}{n}\right) \quad (5.3.4)$$

In all the above calculations, t is any time instance. For a sample failure time data set $\mathscr{T}$, as given above, t is a member of $\mathscr{T}$.

5.4 COMPARISON OF ESTIMATES

The three estimators of reliability are to be compared by comparing the properties satisfied by them. The Improved Estimator of $R(t)$ is unbiased and sufficient, as it is obtained from MLE of $R(t)$, by removing the bias present in it. The only property to be checked thus, is the efficiency property. Since MLE of $R(t)$ is biased as shown above, while MVUE of $R(t)$ and Improved Estimator of $R(t)$ are unbiased, coefficient of variation is used as a measure of dispersion instead of the variance, as mentioned in Section 1.1 of Chapter 1. The estimate with the least value of the coefficient of variation is considered as the efficient estimator. The comparison is also done by considering the quartile coefficient of dispersion, as mentioned in Section 1.1 of Chapter 1. Even with this measure, the estimate with the least value of the quartile coefficient of dispersion is considered as the efficient estimator. For this purpose, the following case studies have been considered and the three estimates have been found. The coefficients of variation and the quartile coefficient of dispersion for these three estimates have also been obtained. For all the case studies, CV($\hat{R}(t)$), CV($\tilde{R}(t)$) and CV($\breve{R}(t)$) are respectively obtained using

$$\text{CV}(\hat{R}(t)) = \frac{S_{\hat{R}(t)}}{\bar{\hat{R}}(t)} \quad (5.4.1)$$

$$\text{CV}(\tilde{R}(t)) = \frac{S_{\tilde{R}(t)}}{\bar{\tilde{R}}(t)} \quad (5.4.2)$$

$$\text{CV}(\breve{R}(t)) = \frac{S_{\breve{R}(t)}}{\overline{\breve{R}}(t)} \tag{5.4.3}$$

Here, the sample variances $S^2_{\hat{R}(t)}$, $S^2_{\tilde{R}(t)}$ and $S^2_{\breve{R}(t)}$ are respectively obtained using

$$S^2_{\hat{R}(t)} = \sum_{t\epsilon\mathscr{T}} \frac{\left(\hat{R}(t) - \overline{\hat{R}}(t)\right)^2}{(n-1)} \tag{5.4.4}$$

$$S^2_{\tilde{R}(t)} = \sum_{t\epsilon\mathscr{T}} \frac{\left(\tilde{R}(t) - \overline{\tilde{R}}(t)\right)^2}{(n-1)} \tag{5.4.5}$$

$$S^2_{\breve{R}(t)} = \sum_{t\epsilon\mathscr{T}} \frac{\left(\breve{R}(t) - \overline{\breve{R}}(t)\right)^2}{(n-1)} \tag{5.4.6}$$

Further, the sample means $\overline{\hat{R}}(t)$, $\overline{\tilde{R}}(t)$, and $\overline{\breve{R}}(t)$ are respectively obtained using

$$\overline{\hat{R}}(t) = \frac{\sum\limits_{t\epsilon\mathscr{T}} \hat{R}(t)}{n} \tag{5.4.7}$$

$$\overline{\tilde{R}}(t) = \frac{\sum\limits_{t\epsilon\mathscr{T}} \tilde{R}(t)}{n} \tag{5.4.8}$$

$$\overline{\breve{R}}(t) = \frac{\sum\limits_{t\epsilon\mathscr{T}} \breve{R}(t)}{n} \tag{5.4.9}$$

Case study 1: On-Line Data Entry IBM Software Package

The data reported by Ohba (Ohba (1984)) are recorded from testing an on-line data entry software package developed at IBM. There are 15 failures, with failure times as indicated in Table 5.4.1.

Table 5.4.1 On-Line Data Entry IBM Software Package

Failure Number	1	2	3	4	5	6	7	8	9	10
Failure Time	10	19	32	43	58	70	88	103	125	150
Failure Number	11	12	13	14	15					
Failure Time	169	199	231	256	296					

Table 5.4.2 denotes the MLE, MVUE and the Improved Estimator of reliability functions for Gamma class models. In this table, SD$_{\hat{R}(t)}$, SD$_{\tilde{R}(t)}$ and SD$_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively.

Table 5.4.2 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Gamma Model (Case study 1)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	SD$_{\hat{R}(t)}$	SD$_{\tilde{R}(t)}$	SD$_{\breve{R}(t)}$
1	10	0.98818	0.98922	0.99272	0.23347	0.2301	0.23347
2	19	0.96121	0.96432	0.96574	0.20813	0.20683	0.20813
3	32	0.90392	0.91056	0.90846	0.15914	0.16083	0.15914
4	43	0.845	0.85438	0.84954	0.11561	0.11892	0.11561
5	58	0.75743	0.7695	0.76197	0.06372	0.06759	0.06372
6	70	0.68596	0.69914	0.6905	0.03275	0.03595	0.03275
7	88	0.58227	0.59549	0.58681	0.00597	0.00739	0.00597
8	103	0.50225	0.51429	0.50679	7.5E-06	2.3E-05	7.5E-06
9	125	0.39845	0.40752	0.40298	0.01135	0.01041	0.01135
10	150	0.30117	0.30615	0.3057	0.04155	0.04137	0.04155
11	169	0.24114	0.2431	0.24567	0.06962	0.07099	0.06962
12	199	0.16749	0.16552	0.17202	0.11391	0.11835	0.11391
13	231	0.11189	0.1072	0.11643	0.15453	0.16187	0.15453
14	256	0.08095	0.07515	0.08549	0.17981	0.18869	0.17981
15	296	0.04763	0.04145	0.05217	0.20918	0.2191	0.20918

Using the values obtained in Table 5.4.2 in equations (5.4.4) to (5.4.9), we get,
$\overline{\hat{R}}(t) = 0.505$; $S^2_{\hat{R}(t)} = 0.1142$; $\overline{\tilde{R}}(t) = 0.50953$; $S^2_{\tilde{R}(t)} = 0.117$;
$\overline{\breve{R}}(t) = 0.50953$; $S^2_{\breve{R}(t)} = 0.1142$.

Hence, using (5.3.3), the bias in $\hat{R}(t)$ is obtained as
Bias$(\hat{R}(t)) = 0.505 - 0.50953 = -0.00454$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.00454) = \hat{R}(t) + 0.00454$.

Thus, using equations (5.4.1) to (5.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.6692, CV($\tilde{R}(t)$)=0.6714 and **CV($\breve{R}(t)$)=0.6632**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.1675 and Q_3=0.8450. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as

QD($\hat{R}(t)$)=0.6691. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.16552 and Q_3=0.85438. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.6754. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.17202 and Q_3=0.84954. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.6632**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 5.4.1.

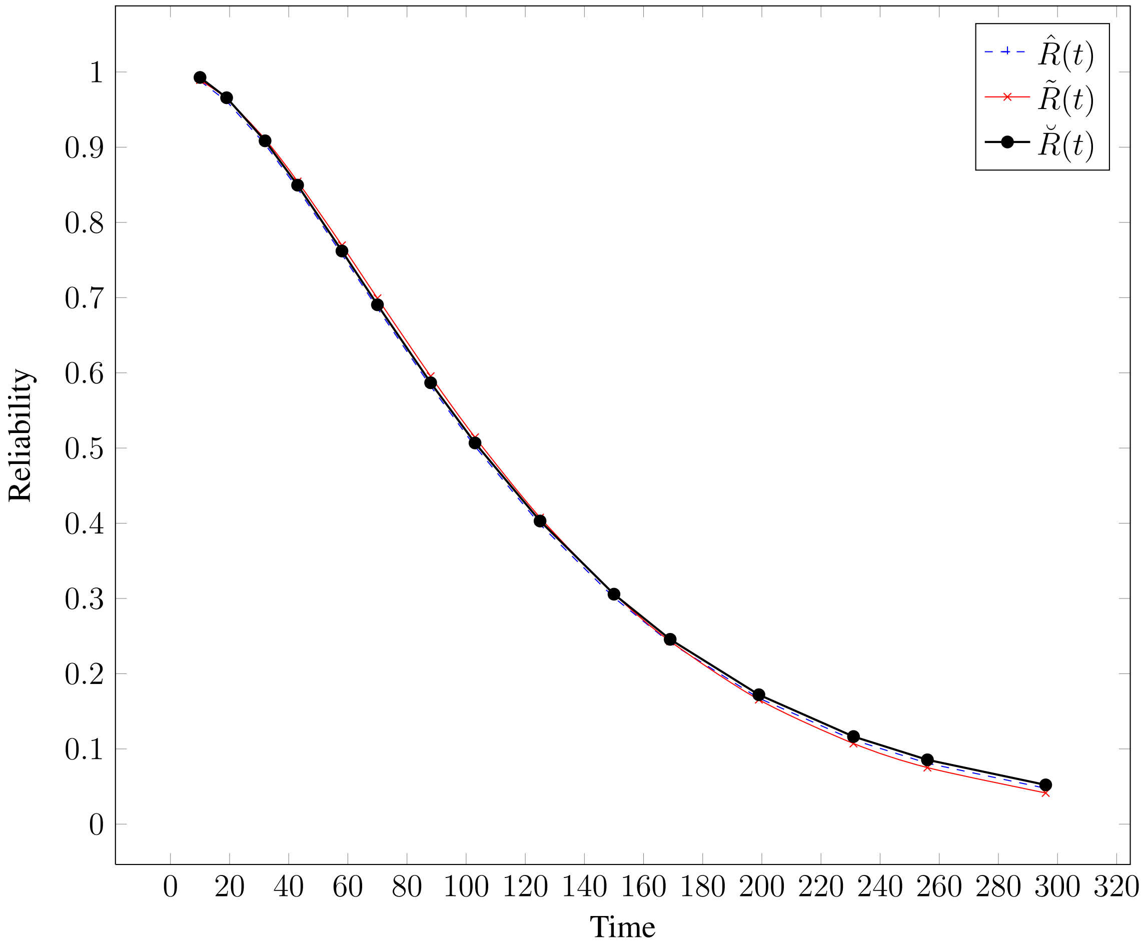

Figure 5.4.1 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Gamma Model (Case study 1)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 2: Nuclear Power Agency

A nuclear power agency uses a computer-based monitoring system for its reactors. The operating system for the computer is employed for this and other applications in an estimated 5000 installations throughout the world. A total of 17 failures have occurred with failure times as listed in Table 5.4.3 (Musa et al. (1991)).

Table 5.4.3 Nuclear Power Agency

Failure number	1	2	3	4	5	6
Failure time	932	4035	4696	4893	6369	6524
Failure number	7	8	9	10	11	12
Failure time	7882	8170	9339	10400	10542	11036
Failure number	13	14	15	16	17	
Failure time	11696	11905	12266	12954	14000	

Table 5.4.4 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in Table 5.4.4 in equations (5.4.4) to (5.4.9), we get,

$\overline{\hat{R}}(t) = 0.441504$; $S^2_{\hat{R}(t)} = 0.05485$; $\overline{\tilde{R}}(t) = 0.447354$; $S^2_{\tilde{R}(t)} = 0.05615$;

$\overline{\breve{R}}(t) = 0.0.447354$; $S^2_{\breve{R}(t)} = 0.05485$.

Hence, using (5.3.3), the bias in $\hat{R}(t)$ is obtained as

$\text{Bias}(\hat{R}(t)) = 0.441504 - 0.447354 = -0.00585$.

Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as

$\breve{R}(t) = \hat{R}(t) - (-0.00585) = \hat{R}(t) + 0.00585$.

Using equations (5.4.1) to (5.4.3), the coefficient of variation of the three estimators, are respectively obtained as

CV($\hat{R}(t)$)=0.5305, CV($\tilde{R}(t)$)=0.5296 and **CV($\breve{R}(t)$)=0.5235**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.2455 and Q_3=0.6291. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.4386. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2475 and Q_3=0.6407. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is

Table 5.4.4 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Gamma Model (Case study 2)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$SD_{\hat{R}(t)}$	$SD_{\tilde{R}(t)}$	$SD_{\breve{R}(t)}$
1	932	0.98001	0.98152	0.98586	0.28999	0.28533	0.28999
2	4035	0.76177	0.77232	0.76762	0.10257	0.1056	0.10257
3	4696	0.70583	0.71724	0.71168	0.06987	0.07284	0.06987
4	4893	0.68922	0.70079	0.69507	0.06136	0.06423	0.06136
5	6369	0.56902	0.58053	0.57487	0.01626	0.01773	0.01626
6	6524	0.55702	0.5684	0.56287	0.01334	0.01465	0.01334
7	7882	0.45834	0.46798	0.46419	0.00028	0.00043	0.00028
8	8170	0.43903	0.44818	0.44488	0.00000	0.0000	0.0000
9	9339	0.36675	0.37367	0.3726	0.00559	0.00543	0.00559
10	10400	0.30952	0.31426	0.31537	0.01742	0.01772	0.01742
11	10542	0.30245	0.30689	0.3083	0.01934	0.01973	0.01934
12	11036	0.27889	0.28232	0.28474	0.02644	0.02723	0.02644
13	11696	0.24984	0.25199	0.25569	0.03673	0.03817	0.03673
14	11905	0.24121	0.24296	0.24706	0.04012	0.04178	0.04012
15	12266	0.2269	0.22798	0.23275	0.04606	0.04812	0.04606
16	12954	0.20167	0.20158	0.20752	0.05752	0.06041	0.05752
17	14000	0.16809	0.16642	0.17394	0.07476	0.07892	0.07476

obtained as QD($\tilde{R}(t)$)=0.4427. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2514 and Q_3=0.6350. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.4327**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 5.4.2.

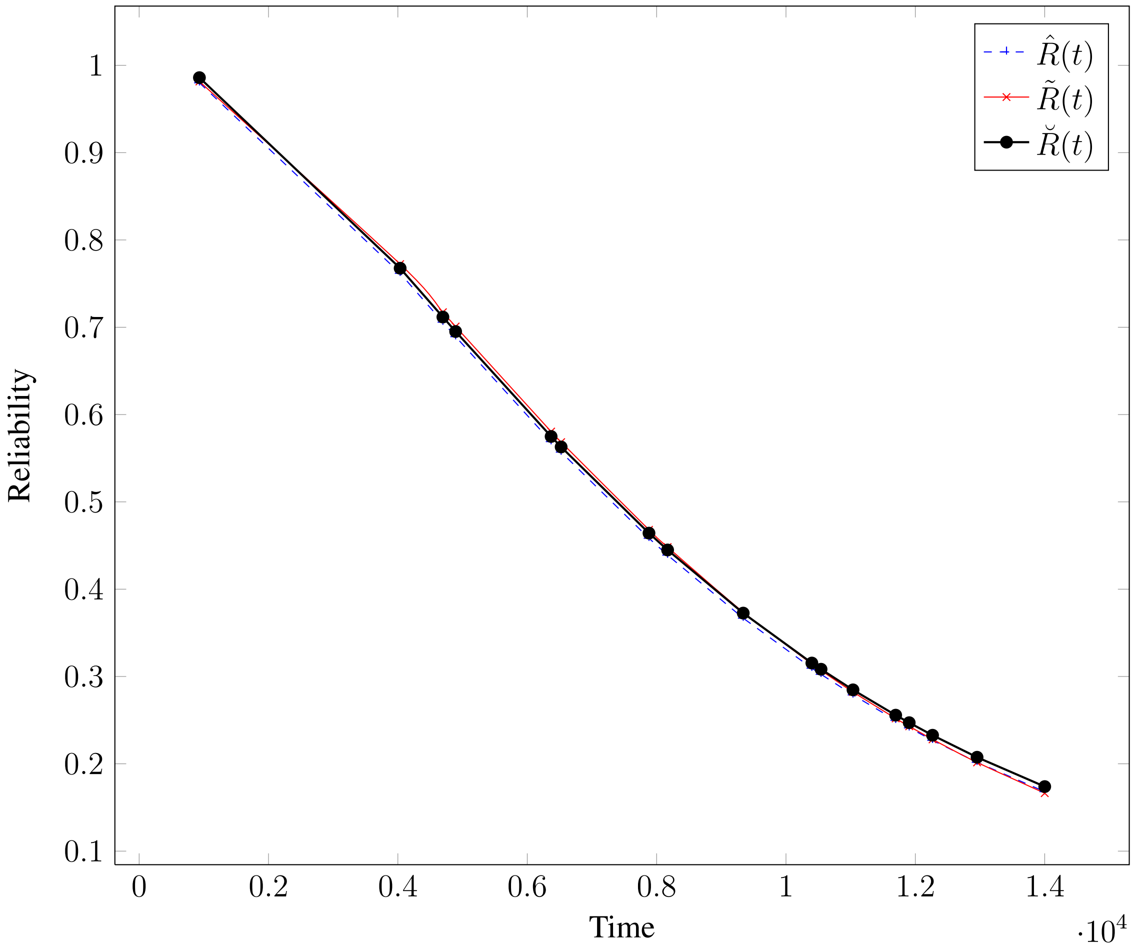

Figure 5.4.2 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Gamma Model (Case study 2)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 3: Failure data set of Lyu

The failure time data for 10 failures obtained by Lyu (Lyu (2004)) are given in Table 5.4.5.

Table 5.4.5 Failure data set of Lyu

Failure number	1	2	3	4	5	6	7	8	9	10
Failure time	7	18	26	36	51	73	93	118	146	181

Table 5.4.6 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in

Table 5.4.6 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Gamma Model (Case study 3)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$\text{SD}_{\hat{R}(t)}$	$\text{SD}_{\tilde{R}(t)}$	$\text{SD}_{\breve{R}(t)}$
1	7	0.9846	0.9866	0.9911	0.22602	0.22367	0.22185
2	18	0.9156	0.9246	0.9221	0.16521	0.16891	0.16165
3	26	0.8462	0.8603	0.8527	0.1136	0.12015	0.11064
4	36	0.75	0.7685	0.7565	0.05801	0.06497	0.05591
5	51	0.6051	0.6254	0.6116	0.0092	0.0125	0.00838
6	73	0.4199	0.4349	0.4264	0.00796	0.0062	0.00877
7	93	0.2907	0.2975	0.2973	0.0477	0.04672	0.04965
8	118	0.1777	0.1753	0.1842	0.10984	0.11449	0.11279
9	146	0.0993	0.091	0.1058	0.16797	0.17861	0.17161
10	181	0.0464	0.0368	0.053	0.21409	0.22736	0.2182

Table 4.4.6 in equations (5.4.4) to (5.4.9), we get,
$\overline{\hat{R}}(t) = 0.51356$; $S^2_{\hat{R}(t)} = 0.12444$; $\overline{\tilde{R}}(t) = 0.52008$; $S^2_{\tilde{R}(t)} = 0.12928$;
$\overline{\breve{R}}(t) = 0.52008$; $S^2_{\breve{R}(t)} = 0.12438$.
Hence, using (5.3.3), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = 0.51356 - 0.52008 = -0.00653$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.00653) = \hat{R}(t) + 0.00653$.

Hence, using equations (5.4.1) to (5.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.6868, CV($\tilde{R}(t)$)=0.6914 and **CV($\breve{R}(t)$)=0.6781**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.1777 and Q_3=0.8462. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.6529. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.1753 and Q_3=0.8603. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.6615. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.1842 and Q_3=0.8527. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.6447**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 5.4.3.

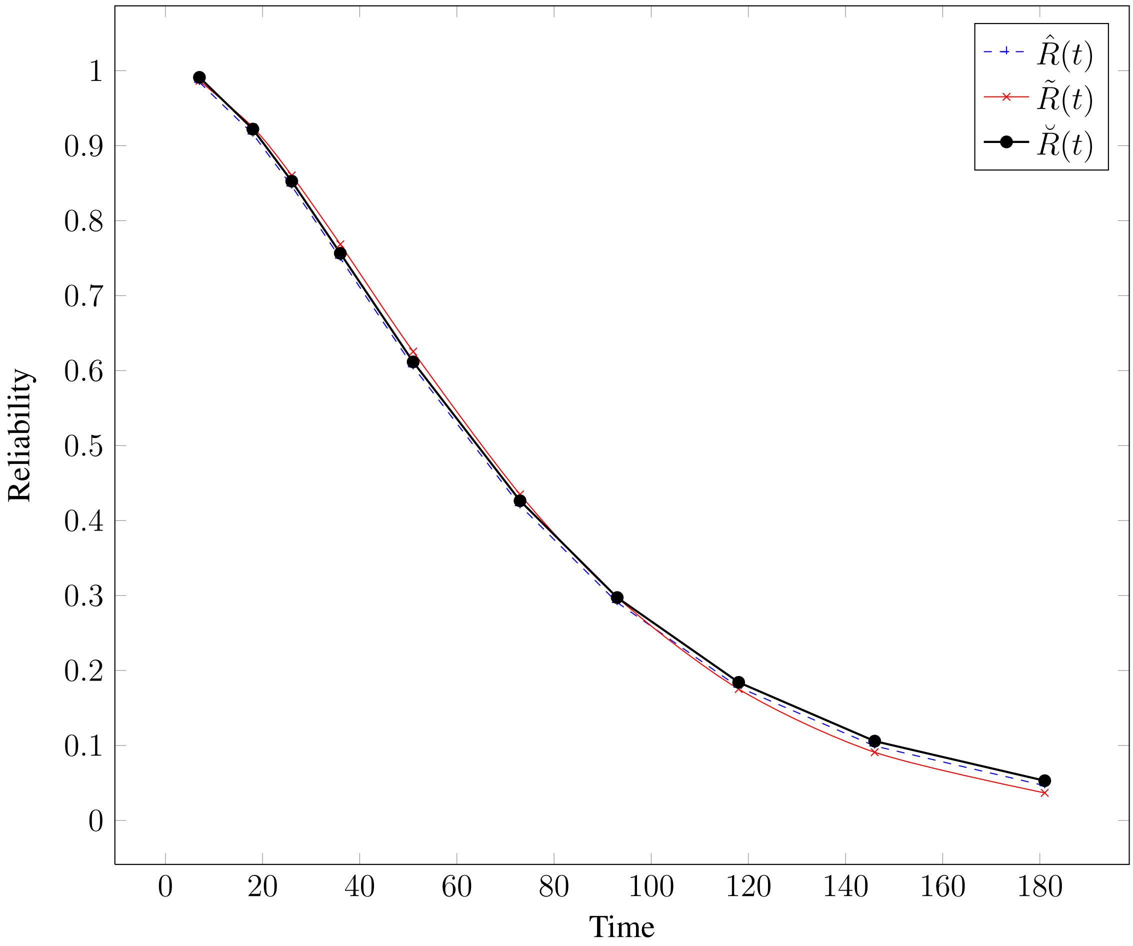

Figure 5.4.3 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Gamma Model (Case study 3)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Table 5.4.7 shows the consolidated values of the coefficient of variation (CV) of $\breve{R}(t)$, $\hat{R}(t)$ and $\tilde{R}(t)$, for all the three case studies. Table 5.4.8 shows the consolidated values of the coefficient of variation (CV) of $\breve{R}(t)$, $\hat{R}(t)$ and $\tilde{R}(t)$, for all the three case studies.

Table 5.4.7 Gamma Models (Consolidated1)

Case study	CV($\hat{R}(t)$)	CV($\tilde{R}(t)$)	CV($\breve{R}(t)$)
1	0.6692	0.6714	**0.6632**
2	0.5305	0.5296	**0.5235**
3	0.6868	0.6914	**0.6781**

From Tables 5.4.7 and 5.4.8, it is observed that, the Improved Estimator ($\breve{R}(t)$) has the

Table 5.4.8 Gamma Models (Consolidated2)

Case study	QD($\hat{R}(t)$)	QD($\tilde{R}(t)$)	QD($\breve{R}(t)$)
1	0.6691	0.6754	**0.6632**
2	0.4386	0.4427	**0.4327**
3	0.6529	0.6615	**0.6447**

least values of the coefficient of variation and quartile coefficient of dispersion than those of MLE ($\hat{R}(t)$) and MVUE ($\tilde{R}(t)$) in all the three case studies, which means that $\breve{R}(t)$ is more efficient than $\hat{R}(t)$ and $\tilde{R}(t)$.

However, to choose the best estimator among the three estimates, the desirable properties of good estimators as mentioned in Section 1.1 of Chapter 1 are to be considered.

Unbiasedness of $\breve{R}(t)$: It has been shown above that $\hat{R}(t)$ is biased for $R(t)$, while $\tilde{R}(t)$ is unbiased for $R(t)$. Since the Improved Estimators are obtained from MLEs just by removing the bias present in the MLEs, they satisfy the unbiasedness property. Thus, $\breve{R}(t)$ is unbiased for $R(t)$.

Sufficiency of $\breve{R}(t)$: Improved Estimator is a function of MLE, which is sufficient. Since any function of sufficient estimator is also sufficient, Improved Estimator is sufficient.

Now, to compare the biased estimator $\hat{R}(t)$ with unbiased estimators $\tilde{R}(t)$ and $\breve{R}(t)$, the coefficient of variation and the quartile coefficient of dispersion are considered as measures of dispersion to check the efficiency property. The sample results of comparison of coefficients of variation and the quartile coefficients of dispersion for the three estimators indicate that Improved Estimator has least values of coefficient of variation and the quartile coefficient of dispersion as compared to those of MLE and MVUE of $R(t)$, which indicates that the Improved estimators are efficient compared to MLE and MVUE.

Thus, by referring to Table 1.1.1 of Chapter 1, Table 5.4.9 provides the statistical properties satisfied by MLE, MVUE and the Improved Estimator of reliability for Gamma class models.

Table 5.4.9 Gamma class models- Properties satisfied by estimators of reliability

	Unbiased	**Sufficient**	**Efficient**
MLE	No	Yes	No
MVUE	Yes	Yes	No
Improved Estimator	**Yes**	**Yes**	**Yes**

It can be seen from Table 5.4.9 that the Improved Estimator satisfies maximum number of properties of estimators as compared to MLE and MVUE of $R(t)$. Hence, it can be inferred that the estimate of reliability obtained using the Improved Estimator, is more efficient than those estimated using the methods of MLE and MVUE.
Hence, it is concluded that $\breve{R}(t)$ gives more accurate value of reliability than $\hat{R}(t)$ and $\tilde{R}(t)$, for Gamma class software reliability models.

Chapter 6

PARETO CLASS MODELS

In this class of models, the failure times (T) are assumed to have Pareto distribution with probability density function, given by

$$f(t) = \frac{\alpha}{\beta}\left(1 + \frac{t}{\beta}\right)^{-\alpha-1} = \frac{\alpha\beta^{\alpha}}{(t+\beta)^{\alpha+1}}, \qquad t > \beta \tag{6.0.1}$$

where α denotes the shape parameter and β, the scale parameter.
Eventhough not much work has been done in estimating the reliability of the software for this model, estimation of failure intensity function has been done by Kuo and Yang (Kuo and Yang (1995)). Baysian approach was used to estimate the software reliability. Some statistical tools, such as statistical usage testing have also been used in estimating the reliability of the software (Guen et al. (2004)). These are not specific to any model, but can be used for all class and types of model, including the Pareto model. Expectation maximization principle was used by Hiroyuki et. al (Okamura et al. (2003)) to estimate the parameters of Pareto model.

The random variable T, having the pdf as given in (6.0.1) is written as $T \sim P(\alpha, \beta)$. The reliability function at time t, denoted by$R(t)$, is obtained as

$$R(t) = P(T > t) = \int_{t}^{\infty} f(t)dt = \int_{t}^{\infty} \frac{\alpha}{\beta}\left(1 + \frac{t}{\beta}\right)^{-\alpha-1} dt \tag{6.0.2}$$

With suitable substitution and simplification, the above expression reduces to,

$$R(t) = \left(1 + \frac{t}{\beta}\right)^{-\alpha} \tag{6.0.3}$$

It is intended to obtain the estimates of this reliability using the methods of MLE and MVUE.

6.1 MLE OF $R(t)$

Since MLEs satisfy the invariance property, the MLE of $R(t)$, denoted by $\hat{R}(t)$ is obtained as

$$\hat{R}(t) = \left(1 + \frac{t}{\hat{\beta}}\right)^{-\hat{\alpha}} \tag{6.1.1}$$

where $\hat{\alpha}$ and $\hat{\beta}$ are the MLEs of α and β (if β is unknown) respectively.

To find the MLE of α: Let $(T_1, T_2, \ldots T_n)$ be a sample of size n from Pareto distribution as given in (6.0.1). Then, the likelihood function of this sample is given by

$$L = \prod_{i=1}^{n} f(t_i) = \alpha^n \beta^{n\alpha} \prod_{i=1}^{n} \frac{1}{(t_i + \beta)^{\alpha+1}} \tag{6.1.2}$$

Maximizing this likelihood function using the concept of differential calculus, the MLE of α, denoted by $\hat{\alpha}$, is obtained as the solution of $\dfrac{\partial \ln L}{\partial \alpha} = 0$ with $\dfrac{\partial^2 \ln L}{\partial \alpha^2} < 0$.

Now, $\dfrac{\partial \ln L}{\partial \alpha} = 0$ gives $\dfrac{n}{\alpha} + n \ln \beta - \displaystyle\sum_{i=1}^{n} \ln(t_i + \beta) = 0$, from which, the MLE of α is obtained as

$$\hat{\alpha} = \frac{n}{\sum_{i=1}^{n} \left(\ln(1 + \frac{t_i}{\hat{\beta}}) \right)} \tag{6.1.3}$$

To find the MLE of β: If the value of β is unknown, the MLE of β is the value of β that maximizes the likelihood function. It can be observed from the likelihood function as obtained in (6.1.2) that L is maximum, if β is maximum. Since $\beta < t$, the maximum value that β can take in the sample considered $(T_1, T_2, \ldots T_n)$, is T_1. Hence, the MLE of β is obtained as

$$\hat{\beta} = t_1 \tag{6.1.4}$$

Using (6.1.3) and (6.1.4) in (6.1.1), the MLE of $R(t)$, denoted by $\hat{R}(t)$, is obtained as

$$\hat{R}(t) = \left(1 + \frac{t}{t_1}\right)^{-\left(\frac{n}{\sum_{i=1}^{n}\left(\ln(1+\frac{t_i}{t_1})\right)}\right)} \tag{6.1.5}$$

6.2 MVUE OF $R(t)$

To find the MVUE of $R(t)$, define a function of the random variable T_1 as

$$U(t_1) = \begin{cases} 1 & \text{if } t_1 > t \\ 0 & \text{otherwise} \end{cases}$$

Then, $E[U(t_1)] = 1.P(T_1 > t) + 0.P(T_1 \leq t) = P(T_1 > t) = R(t)$.
Therefore, $U(t_1)$ is unbiased for $R(t)$.

Complete Sufficient Estimator of (α, β): Taking log on both sides of (6.1.2), the log likelihood function is given by $\ln L = n\ln\alpha + n\alpha\ln\beta - (\alpha+1)\sum_{i=1}^{n}\ln((t_i+\beta))$.

This can be written as $\ln L = n\ln\alpha + \alpha\sum_{i=1}^{n}\ln\beta - (\alpha+1)\sum_{i=1}^{n}\ln(t_i+\beta)$.

i.e., $\ln L = n\ln\alpha + \alpha\sum_{i=1}^{n}\ln\beta - \alpha\sum_{i=1}^{n}\ln(t_i+\beta) - \sum_{i=1}^{n}\ln(t_i+\beta)$.

i.e., $\ln L = n\ln\alpha - \alpha\left(\sum_{i=1}^{n}\ln(t_i+\beta) - \sum_{i=1}^{n}\ln\beta\right) - \sum_{i=1}^{n}\ln(t_i+\beta)$.

i.e., $\ln L = n\ln\alpha - \sum_{i=1}^{n}\ln(t_i+\beta) - \alpha\sum_{i=1}^{n}\ln\left(\frac{t_i+\beta}{\beta}\right)$.

i.e., $\ln L = n\ln\alpha - \sum_{i=1}^{n}\ln(t_i+\beta) - \alpha\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)$.

Now, applying the factorization theorem (as explained in Chapter 1), it can be seen that the log likelihood function (and hence L) depends on α, β and t_i, only through the value of $\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)$. Thus, the sufficient estimator of (α, β) is $\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)$.

Since each $T_i \sim P(\alpha,\beta)$, $\ln(1+\frac{t_i}{\beta}) \sim \mathscr{E}(\alpha)$ (Section 4 of Appendix A) and hence $\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right) \sim G(n,\alpha)$ (Section 1 of Appendix A). Further, by Result 3 of Appendix A, the estimator $\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)$ is also the complete statistic and hence it is the complete sufficient estimator of (α, β).

Also, $U(t_1)$ is an unbiased estimator of $R(t)$ and $R(t)$ is a function of α and β, as given in (6.0.3). Hence, by Theorem 1 of Chapter 1, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = E(U(t_1)|X) = \int_t^{\infty} f(t_1|x)dt_1$$

where $f(t_1|x)$ denotes the conditional pdf of T_1 given X and is given by
$f(t_1|x) = \dfrac{g(t_1,x)}{h(x)}$,
where $g(t_1,x)$ denotes the joint pdf of T_1 and X and $h(x)$ denotes the marginal pdf of X.
Hence, the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_t^{\infty} \frac{g(t_1,x)}{h(x)}dt_1 \qquad (6.2.1)$$

Since each $T_i \sim P(\alpha,\beta)$, $\ln(1+\dfrac{t_i}{\beta}) \sim \mathscr{E}(\alpha)$ and hence $X \sim G(n,\alpha)$ (Appendix A).
Hence, the pdf of X is given by

$$h(x) = \frac{1}{\Gamma(n)}\alpha^n e^{-\alpha x}x^{n-1}, \qquad x > 0 \qquad (6.2.2)$$

To find the pdf $g(t_1,x)$, split the sample $(T_1,T_2,T_3,\ldots T_n)$ into two samples as T_1 of size one and $(T_2,T_3,T_4,\ldots T_n)$ of size $(n-1)$.
Since T_1 and $\sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big)$ are independent, the joint pdf of T_1 and $\sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big)$ is obtained as
$g\bigg(t_1,\sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big)\bigg) = f(t_1)\cdot h\bigg(\sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big)\bigg)$,
where $f(t_1)$ and $h\bigg(\sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big)\bigg)$ denote the pdfs of T_1 and $\sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{T_i}{\beta})\Big)$ respectively.
Since $T_1 \sim P(\alpha,\beta)$, the pdf of T_1 is given by $f(t_1) = \dfrac{\alpha\beta^{\alpha}}{(t_1+\beta)^{\alpha+1}}$.
Also, since $X = \sum\limits_{i=1}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big) \sim G(n,\alpha)$, we have,
$Y = \sum\limits_{i=2}^{n}\Big(\ln(1+\dfrac{t_i}{\beta})\Big) \sim G(n-1,\alpha)$.
Hence, pdf of Y is given by $h(y) = \dfrac{1}{\Gamma(n-1)}\alpha^{n-1}e^{-\alpha y}y^{n-2}$.
Since T_1 and Y are independent, the joint pdf of T_1 and Y is obtained as

$$g(t_1, y) = f(t_1) \cdot h(y) = \frac{\alpha\beta^\alpha}{(t_1+\beta)^{\alpha+1}} \frac{1}{\Gamma(n-1)} \alpha^{n-1} e^{-\alpha y} y^{n-2}.$$

Considering the transformation $Y = X - \ln\left(1 + \frac{t_1}{\beta}\right)$ and noting that modulus of the Jacobian of the inverse transformation is one (Appendix B), the joint pdf of T_1 and X is obtained as

$$g(t_1, x) = \frac{\alpha\beta^\alpha}{(t_1+\beta)^{\alpha+1}} \frac{\alpha^{n-1}}{\Gamma(n-1)} e^{-\alpha y} y^{n-2} \tag{6.2.3}$$

Substituting (6.2.2) and (6.2.3) in (6.2.1), the MVUE of $R(t)$ is obtained as

$$\tilde{R}(t) = \int_t^\infty \frac{\frac{\alpha\beta^\alpha}{(t_1+\beta)^{\alpha+1}} \frac{\alpha^{n-1}}{\Gamma(n-1)} e^{-\alpha y} y^{n-2}}{\frac{1}{\Gamma(n)} \alpha^n e^{-\alpha x} x^{n-1}} dt_1$$

$$\text{i.e.,} \tilde{R}(t) = \int_t^\infty \frac{\Gamma(n)}{\Gamma(n-1)} \frac{\alpha^n \beta^\alpha}{\alpha^n (t_1+\beta)^{\alpha+1}} \frac{e^{-\alpha y}}{e^{-\alpha x}} \frac{y^{n-2}}{x^{n-1}} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_t^\infty \frac{(n-1)\Gamma(n-1)}{\Gamma(n-1)} \frac{\beta^\alpha}{(t_1+\beta)^{\alpha+1}} \frac{e^{-\alpha y}}{e^{-\alpha x}} \frac{y^{n-2}}{x^{n-1}} dt_1$$

which reduces to

$$\tilde{R}(t) = \int_t^\infty \frac{(n-1)}{(t_1+\beta)^{\alpha+1}} \beta^\alpha e^{-\alpha y + \alpha x} \frac{y^{n-2}}{x^{n-1}} dt_1.$$

On replacing y by $\left(x - \ln(1 + \frac{t_1}{\beta})\right)$, so that $x - y = \ln(1 + \frac{t_1}{\beta})$, the above equation reduces to

$$\tilde{R}(t) = \int_t^\infty \frac{(n-1)}{(t_1+\beta)^{\alpha+1}} \beta^\alpha e^{\alpha \ln(1+\frac{t_1}{\beta})} \frac{(x - \ln(1+\frac{t_1}{\beta}))^{n-2}}{x^{n-1}} dt_1,$$

$$\text{i.e., } \tilde{R}(t) = \int_t^\infty \frac{(n-1)}{(t_1+\beta)^{\alpha+1}} \beta^\alpha (e^{\ln(1+\frac{t_1}{\beta})})^\alpha \frac{(x - \ln(1+\frac{t_1}{\beta}))^{n-2}}{x^{n-2}} \frac{1}{x} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_t^\infty \frac{(n-1)}{(t_1+\beta)^{\alpha+1}} \beta^\alpha \left(1 + \frac{t_1}{\beta}\right)^\alpha \left(1 - \frac{\ln(1+\frac{t_1}{\beta})}{x}\right)^{n-2} \frac{1}{x} dt_1$$

$$\text{i.e., } \tilde{R}(t) = \int_t^\infty \frac{(n-1)}{(t_1+\beta)^{\alpha+1}} \beta^\alpha \frac{(t_1+\beta)^\alpha}{\beta^\alpha} \left(1 - \frac{\ln(1+\frac{t_1}{\beta})}{x}\right)^{n-2} \frac{1}{x} dt_1,$$

which can be written as

$$\tilde{R}(t) = \int_t^\infty \frac{(n-1)}{(t_1+\beta)}\left(1 - \frac{\ln(1+\frac{t_1}{\beta})}{x}\right)^{n-2}\frac{1}{x}dt_1$$

On replacing x by $\sum_{i=1}^{n} \ln\left(1+\frac{t_i}{\beta}\right)$, the above integral becomes

$$\tilde{R}(t) = \int_t^\infty \frac{(n-1)}{\beta}\frac{1}{(1+\frac{t_1}{\beta})}\frac{1}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\left(1 - \frac{\ln(1+\frac{t_1}{\beta})}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{n-2} dt_1$$

The above integral converges if $\ln\left(1+\frac{t_1}{\beta}\right) < \sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)$.

i.e., if $\left(1+\frac{t_1}{\beta}\right) < e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)} \implies \frac{t_1}{\beta} < e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)} - 1.$

Hence, $t_1 < \beta\left(e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)} - 1\right)$.

This is the upper limit of the above integral and hence, $\tilde{R}(t)$ is obtained as

$$\tilde{R}(t) = \frac{(n-1)}{\beta\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\int_t^{\beta\left(e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}-1\right)} \frac{1}{\left(1+\frac{t_1}{\beta}\right)}\left(1 - \frac{\ln(1+\frac{t_1}{\beta})}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-2)} dt_1$$

$$\text{i.e.,}\tilde{R}(t) = -\left(1 - \frac{\ln(1+\frac{t_1}{\beta})}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)}\Bigg|_t^{\beta\left(e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}-1\right)}$$

$$\text{i.e.,}\tilde{R}(t) = -\left(1 - \frac{\ln\left(1+\frac{\beta\left(e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}-1\right)}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)} + \left(1 - \frac{\ln\left(1+\frac{t}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)}$$

Further steps of simplification are provided below:

$$
\begin{aligned}
\tilde{R}(t) &= -\left(1 - \frac{\ln\left(e^{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)} + \left(1 - \frac{\ln\left(1+\frac{t}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)} \\
&= -\left(1 - \frac{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)} + \left(1 - \frac{\ln\left(1+\frac{t}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)} \\
&= -(1-1) + \left(1 - \frac{\ln\left(1+\frac{t}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)} \\
&= \left(1 - \frac{\ln\left(1+\frac{t}{\beta}\right)}{\sum_{i=1}^{n}\ln\left(1+\frac{t_i}{\beta}\right)}\right)^{(n-1)}
\end{aligned}
$$

Hence, the expression for MVUE of $R(t)$ is obtained as,

$$
\tilde{R}(t) = \begin{cases} \left(1 - \frac{\ln(1+\frac{t}{\beta})}{\sum_{i=1}^{n}\ln(1+\frac{t_i}{\beta})}\right)^{n-1} & \text{where } \ln(1+\dfrac{t}{\beta}) < \sum_{i=1}^{n}\ln\left(1+\dfrac{t_i}{\beta}\right) \\ 0 & \text{otherwise} \end{cases} \tag{6.2.4}
$$

6.3 IMPROVED ESTIMATOR OF $R(t)$

The reliability function for Pareto class models is given by

$$
\begin{aligned}
R(t) &= \left(1+\frac{t}{\beta}\right)^{-\alpha} \\
&= 1 - \frac{\alpha t}{\beta} + \frac{(-\alpha)(-\alpha-1)}{2!}\left(\frac{t}{\beta}\right)^2 + \frac{(-\alpha)(-\alpha-1)(-\alpha-2)}{3!}\left(\frac{t}{\beta}\right)^3 + \ldots
\end{aligned} \tag{6.3.1}
$$

$\hat{R}(t)$ and $\tilde{R}(t)$ are unbiased for $R(t)$, if (i) $E(\hat{R}(t)) = R(t)$ and (ii) $E(\tilde{R}(t)) = R(t)$ respectively.

To check whether $\hat{R}(t)$ is unbiased or not, consider

$$E(\hat{R}(t)) = E\left(\left(1+\frac{t}{\beta}\right)^{-\frac{n}{\sum_{i=1}^{n}\ln(1+\frac{t_i}{\beta})}}\right)$$, where β is taken as its estimated value t_1.

Taking $X = \sum_{i=1}^{n} \ln(1+\frac{t_i}{\beta})$, we have,

$$E(\hat{R}(t)) = E\left(\left(1+\frac{t}{\beta}\right)^{-\frac{n}{X}}\right) = E\left(1 - \frac{n}{X}\frac{t}{\beta} + \frac{(-\frac{n}{X})(-\frac{n}{X}-1)}{2!}(\frac{t}{\beta})^2 + \dots\right).$$

Since $X \sim G(n, \alpha)$, (Section 1 of Appendix A), it can be observed that,

$$\begin{aligned}
E\left(\frac{1}{X}\right) &= \int_0^\infty \frac{1}{x}\frac{\alpha^n}{\Gamma(n)} e^{-x\alpha} x^{n-1} dx \\
&= \frac{\alpha^n}{\Gamma(n)} \int_0^\infty e^{-\alpha x} x^{n-2} dx \\
&= \frac{\alpha^n}{\Gamma(n)} \frac{\Gamma(n-1)}{\alpha^{n-1}} \qquad \text{(Result 4 of Appendix A)} \\
&= \frac{\alpha\Gamma(n-1)}{(n-1)\Gamma(n-1)} = \frac{\alpha}{n-1}.
\end{aligned}$$

$$\begin{aligned}
\text{Similarly, } E\left(\frac{1}{X^2}\right) &= \int_0^\infty \frac{1}{x^2}\frac{\alpha^n}{\Gamma(n)} e^{-x\alpha} x^{n-1} dx \\
&= \frac{\alpha^n}{\Gamma(n)} \int_0^\infty e^{-\alpha x} x^{n-3} dx \\
&= \frac{\alpha^n}{\Gamma(n)} \frac{\Gamma(n-2)}{\alpha^{n-2}} \qquad \text{(Result 4 of Appendix A)} \\
&= \frac{\alpha^2\Gamma(n-2)}{(n-1)(n-2)\Gamma(n-2)} = \frac{\alpha^2}{(n-1)(n-2)}.
\end{aligned}$$

$$\begin{aligned}
E\left(\frac{1}{X^3}\right) &= \int_0^\infty \frac{1}{x^3}\frac{\alpha^n}{\Gamma(n)} e^{-x\alpha} x^{n-1} dx \\
&= \frac{\alpha^n}{\Gamma(n)} \int_0^\infty e^{-\alpha x} x^{n-4} dx \\
&= \frac{\alpha^n}{\Gamma(n)} \frac{\Gamma(n-3)}{\alpha^{n-3}} \qquad \text{(Result 4 of Appendix A)} \\
&= \frac{\alpha^3\Gamma(n-3)}{(n-1)(n-2)(n-3)\Gamma(n-3)} \\
&= \frac{\alpha^3}{(n-1)(n-2)(n-3)}
\end{aligned}$$

and so on.
Thus $E(\hat{R}(t)) = 1 - \Big(\frac{t}{\beta}\Big)\frac{n\alpha}{(n-1)} + \Big(\frac{t}{\beta}\Big)^2\frac{n^2\alpha^2 + n^2\alpha - 2n\alpha}{(n-1)(n-2)} + \ldots \neq R(t).$

Hence, $\hat{R}(t)$ is not unbiased for $R(t)$.

To verify that $\tilde{R}(t)$ is unbiased for $R(t)$, consider

$$E(\tilde{R}(t)) = E\left(1 - \frac{\ln(1+\frac{t}{\beta})}{\sum\limits_{i=1}^{n}\ln(1+\frac{t_i}{\beta})}\right)^{n-1} = E\left(1 - \frac{\ln(1+\frac{t}{\beta})}{X}\right)^{n-1}.$$

$$\text{i.e. } E(\tilde{R}(t)) = E\Bigg(1 - (n-1)\frac{\ln(1+\frac{t}{\beta})}{X} + \frac{(n-1)(n-2)}{2!}\frac{\ln(1+\frac{t}{\beta})^2}{X^2} - \frac{(n-1)(n-2)(n-3)}{3!}\frac{\ln(1+\frac{t}{\beta})^3}{X^3} + \ldots\Bigg)$$

$$\text{Thus, } E(\tilde{R}(t)) = 1 - \alpha\Big(\frac{t}{\beta}\Big) + \frac{(-\alpha)(-\alpha-1)}{2!}\Big(\frac{t}{\beta}\Big)^2 + \frac{(-\alpha)(-\alpha-1)(-\alpha-2)}{3!}\Big(\frac{t}{\beta}\Big)^3 + \ldots$$
$$= R(t)$$

Hence, $\tilde{R}(t)$ is unbiased for $R(t)$.

Since $\hat{R}(t)$ is biased for $R(t)$ and $\tilde{R}(t)$ is unbiased for $R(t)$, the bias of $\hat{R}(t)$ is given by

$$\begin{aligned}\text{Bias}(\hat{R}(t)) &= E(\hat{R}(t)) - R(t) = E(\hat{R}(t)) - E(\tilde{R}(t))\\ &= -\Big(\frac{\Phi t^2}{2}\Big)\Big(\frac{(2n)^2}{(2n-1)(2n-2)} - 1\Big)\\ &\quad + \Big(\frac{\Phi t^3}{3}\Big)\Big(\frac{(2n)^3}{(2n-1)(2n-2)(2n-3)} - 1\Big) + \ldots\end{aligned}$$

The bias can hence be found for the given sample failure data, by using the estimated values of $\hat{R}(t)$ and $\tilde{R}(t)$.
Hence, if $\mathscr{T} = \{t_1, t_2, \ldots t_n\}$ is the given sample failure data set of size n, then the bias is obtained by taking the difference in the means of $\hat{R}(t)$ and $\tilde{R}(t)$ and is obtained as

$$\text{Bias}(\hat{R}(t)) = \frac{\sum\limits_{t\epsilon\mathscr{T}}\hat{R}(t)}{n} - \frac{\sum\limits_{t\epsilon\mathscr{T}}\tilde{R}(t)}{n} \qquad (6.3.2)$$

Removing this bias from $\hat{R}(t)$, the Improved Estimator of $R(t)$, denoted by $\breve{R}(t)$ is obtained as

$$\breve{R}(t) = \hat{R}(t) - \text{Bias}(\hat{R}(t)) = \hat{R}(t) - \left(\frac{\sum_{t \epsilon \mathscr{T}} \hat{R}(t)}{n} - \frac{\sum_{t \epsilon \mathscr{T}} \tilde{R}(t)}{n} \right) \tag{6.3.3}$$

In all the above calculations, t is any time instance. For a sample failure time data set $\mathscr{T}$, as given above, t is a member of $\mathscr{T}$.

6.4 COMPARISON OF ESTIMATES

The three estimators of reliability are to be compared by comparing the properties satisfied by them. The Improved Estimator of $R(t)$ is unbiased and sufficient, as it is obtained from MLE of $R(t)$, by removing the bias present in it. The only property to be checked thus, is the efficiency property. Since MLE of $R(t)$ is biased as shown above, while MVUE of $R(t)$ and Improved Estimator of $R(t)$ are unbiased, coefficient of variation is used as a measure of dispersion instead of the variance, as mentioned in Section 1.1 of Chapter 1. The estimate with the least value of the coefficient of variation is considered as the efficient estimator. The comparison is also done by considering the quartile coefficient of dispersion, as mentioned in Section 1.1 of Chapter 1. Even with this measure, the estimate with the least value of the quartile coefficient of dispersion is considered as the efficient estimator. For this purpose, the following case studies have been considered and the three estimates have been found. The coefficients of variation and quartile coefficient of dispersion for these three estimates have also been obtained. For all the case studies, CV($\hat{R}(t)$), CV($\tilde{R}(t)$) and CV($\breve{R}(t)$) are respectively obtained using

$$\text{CV}(\hat{R}(t)) = \frac{S_{\hat{R}(t)}}{\overline{\hat{R}}(t)} \tag{6.4.1}$$

$$\text{CV}(\tilde{R}(t)) = \frac{S_{\tilde{R}(t)}}{\overline{\tilde{R}}(t)} \tag{6.4.2}$$

$$\text{CV}(\breve{R}(t)) = \frac{S_{\breve{R}(t)}}{\overline{\breve{R}}(t)} \tag{6.4.3}$$

Here, the sample variances $S^2_{\hat{R}(t)}$, $S^2_{\tilde{R}(t)}$ and $S^2_{\breve{R}(t)}$ are respectively obtained using

$$S^2_{\hat{R}(t)} = \sum_{t \epsilon \mathscr{T}} \frac{\left(\hat{R}(t) - \overline{\hat{R}}(t)\right)^2}{(n-1)} \tag{6.4.4}$$

$$S^2_{\tilde{R}(t)} = \sum_{t \epsilon \mathscr{T}} \frac{\left(\tilde{R}(t) - \overline{\tilde{R}}(t)\right)^2}{(n-1)} \quad (6.4.5)$$

$$S^2_{\breve{R}(t)} = \sum_{t \epsilon \mathscr{T}} \frac{\left(\breve{R}(t) - \overline{\breve{R}}(t)\right)^2}{(n-1)} \quad (6.4.6)$$

Further, the sample means $\overline{\hat{R}}(t)$, $\overline{\tilde{R}}(t)$, and $\overline{\breve{R}}(t)$ are respectively obtained using

$$\overline{\hat{R}}(t) = \frac{\sum\limits_{t \epsilon \mathscr{T}} \hat{R}(t)}{n} \quad (6.4.7)$$

$$\overline{\tilde{R}}(t) = \frac{\sum\limits_{t \epsilon \mathscr{T}} \tilde{R}(t)}{n} \quad (6.4.8)$$

$$\overline{\breve{R}}(t) = \frac{\sum\limits_{t \epsilon \mathscr{T}} \breve{R}(t)}{n} \quad (6.4.9)$$

Case study 1: On-Line Data Entry IBM Software Package

The data reported by Ohba (Ohba (1984)) are recorded from testing an on-line data entry software package developed at IBM. There are 15 failures, with failure times as indicated in Table 6.4.1.

Table 6.4.1 On-Line Data Entry IBM Software Package

Failure Number	1	2	3	4	5	6	7	8	9	10
Failure Time	10	19	32	43	58	70	88	103	125	150
Failure Number	11	12	13	14	15					
Failure Time	169	199	231	256	296					

For this case study, β is estimated as $\hat{\beta} = t_1 = 10$.

Table 6.4.2 denotes the MLE, MVUE and the Improved Estimator of reliability functions for Pareto class models. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in Table 6.4.2 in equations (6.4.4) to (6.4.9), we get,

$\overline{\hat{R}}(t) = 0.391703$; $S^2_{\hat{R}(t)} = 0.0231$; $\overline{\tilde{R}}(t) = 0.403189$; $S^2_{\tilde{R}(t)} = 0.0237$;

$\overline{\breve{R}}(t) = 0.403189$; $S^2_{\breve{R}(t)} = 0.0229$.

Table 6.4.2 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Pareto Model (Case study 1)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	SD$_{\hat{R}(t)}$	SD$_{\tilde{R}(t)}$	SD$_{\breve{R}(t)}$
1	10	0.7414	0.7542	0.75289	0.12229	0.12321	0.12229
2	19	0.63153	0.64682	0.64301	0.05752	0.05936	0.05752
3	32	0.53822	0.55407	0.5497	0.02147	0.02276	0.02147
4	43	0.4868	0.50229	0.49828	0.00904	0.00982	0.00904
5	58	0.43714	0.45183	0.44863	0.00206	0.00237	0.00206
6	70	0.40753	0.42151	0.41901	0.00025	0.00034	0.00025
7	88	0.37334	0.38629	0.38483	0.00034	0.00029	0.00034
8	103	0.35108	0.36323	0.36257	0.00165	0.0016	0.00165
9	125	0.32513	0.33623	0.33662	0.00443	0.00448	0.00443
10	150	0.30214	0.31218	0.31363	0.00802	0.00828	0.00802
11	169	0.28785	0.29719	0.29934	0.01078	0.01124	0.01078
12	199	0.26923	0.27759	0.28072	0.015	0.01578	0.015
13	231	0.25317	0.26063	0.26466	0.01919	0.02032	0.01919
14	256	0.24261	0.24945	0.2541	0.02223	0.02364	0.02223
15	296	0.22837	0.23434	0.23986	0.02668	0.02851	0.02668

Hence, using (6.3.2), the bias in $\hat{R}(t)$ is obtained as
Bias$(R(t)) = 0.391703 - 0.403189 = -0.01148$.
Thus, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.01148) = \hat{R}(t) + 0.01148$.
Also, using equations (6.4.1) to (6.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.3880, CV($\tilde{R}(t)$)=0.3818 and **CV($\breve{R}(t)$)=0.3755**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.2692 and Q_3=0.4868. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.2878. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2776 and Q_3=0.5023. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.2881. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2807 and Q_3=0.4983. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.2793**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 6.4.1.

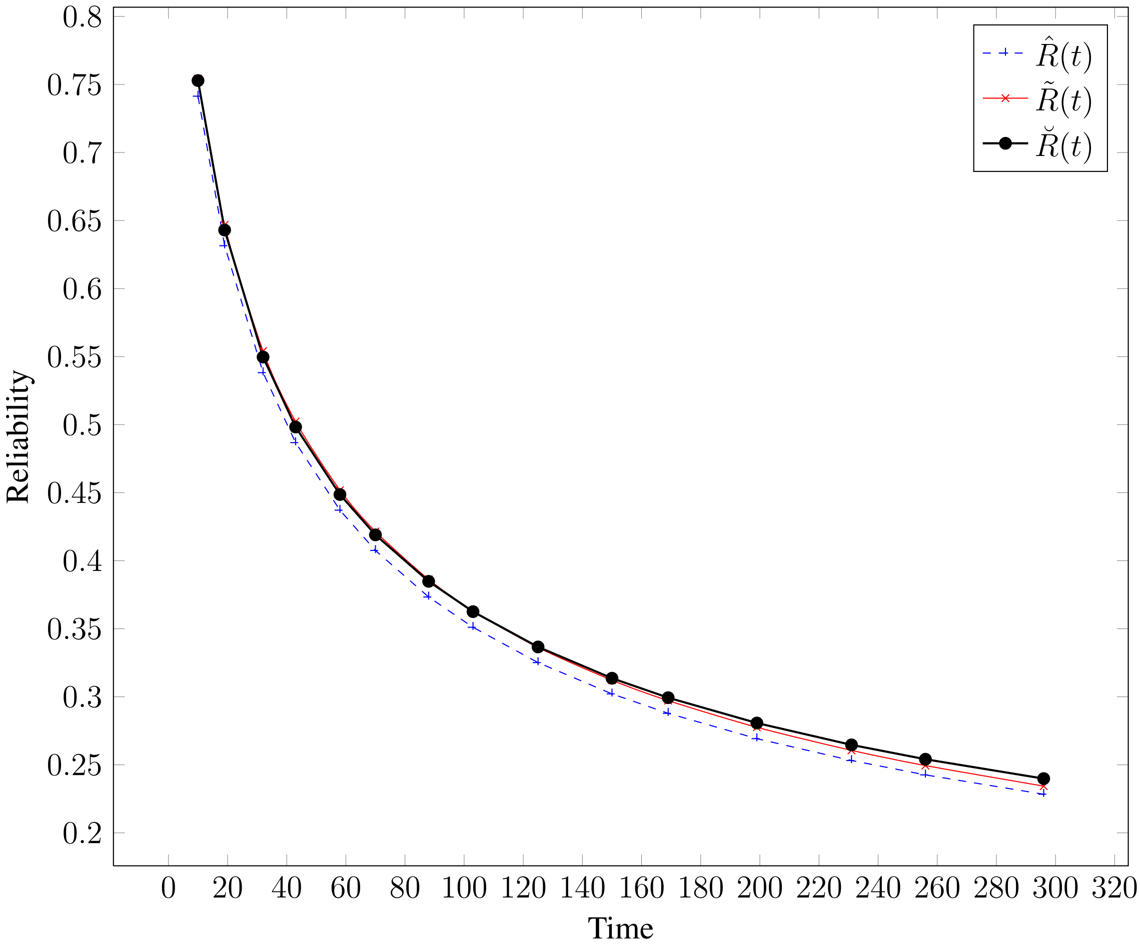

Figure 6.4.1 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Pareto Model (Case study 1)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 2: Nuclear Power Agency

A nuclear power agency uses a computer-based monitoring system for its reactors. The operating system for the computer is employed for this and other applications in an estimated 5000 installations throughout the world. A total of 17 failures have occurred with failure times as listed in Table 6.4.3 (Musa et al. (1991)).

Table 6.4.3 Nuclear Power Agency

Failure number	1	2	3	4	5	6
Failure time	932	4035	4696	4893	6369	6524
Failure number	7	8	9	10	11	12
Failure time	7882	8170	9339	10400	10542	11036
Failure number	13	14	15	16	17	
Failure time	11696	11905	12266	12954	14000	

For this case study, β is estimated as $\hat{\beta} = t_1 = 932$.
Table 6.4.4 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, $\text{SD}_{\hat{R}(t)}$, $\text{SD}_{\tilde{R}(t)}$ and $\text{SD}_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively. Using the values obtained in

Table 6.4.4 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Pareto Model (Case study 2)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	$\text{SD}_{\hat{R}(t)}$	$\text{SD}_{\tilde{R}(t)}$	$\text{SD}_{\breve{R}(t)}$
1	932	0.73324	0.74473	0.74384	0.12572	0.12633	0.12572
2	4035	0.47283	0.48629	0.48343	0.00887	0.00941	0.00887
3	4696	0.44711	0.46019	0.45771	0.00468	0.00503	0.00468
4	4893	0.44028	0.45324	0.45088	0.0038	0.00409	0.0038
5	6369	0.39794	0.40998	0.40854	0.00037	0.00043	0.00037
6	6524	0.39422	0.40617	0.40482	0.00024	0.00028	0.00024
7	7882	0.36577	0.37691	0.37637	0.00017	0.00015	0.00017
8	8170	0.36054	0.37152	0.37114	0.00033	0.00032	0.00033
9	9339	0.34156	0.3519	0.35216	0.00138	0.0014	0.00138
10	10400	0.32685	0.33666	0.33745	0.00269	0.00277	0.00269
11	10542	0.32504	0.33478	0.33564	0.00288	0.00297	0.00288
12	11036	0.31896	0.32847	0.32956	0.00357	0.0037	0.00357
13	11696	0.31139	0.32059	0.32199	0.00453	0.00472	0.00453
14	11905	0.30911	0.31822	0.31971	0.00484	0.00505	0.00484
15	12266	0.30529	0.31425	0.31589	0.00538	0.00563	0.00538
16	12954	0.29843	0.30709	0.30903	0.00644	0.00676	0.00644
17	14000	0.28888	0.29713	0.29948	0.00806	0.00849	0.00806

Table 6.4.4 in equations (6.4.4) to (6.4.9), we get,
$\overline{\hat{R}}(t) = 0.37867$; $S^2_{\hat{R}(t)} = 0.0116$; $\overline{\tilde{R}}(t) = 0.38930$; $S^2_{\tilde{R}(t)} = 0.0117$;
$\overline{\breve{R}}(t) = 0.38927$; $S^2_{\breve{R}(t)} = 0.0115$.

Hence, using (6.3.2), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = 0.37867 - 0.38930 = -0.01063$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.01063) = \hat{R}(t) + 0.01063$.

Using equations (6.4.1) to (6.4.3), the coefficient of variation of the three estimators, are respectively obtained as

CV($\hat{R}(t)$)=0.2844, CV($\tilde{R}(t)$)=0.2778 and **CV($\breve{\boldsymbol{R}}(\boldsymbol{t})$)=0.2754**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.3103 and Q_3=0.4191. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.1492. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.3194 and Q_3=0.4316. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.1494. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.3209 and Q_3=0.4297. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{\boldsymbol{R}}(\boldsymbol{t})$)=0.1450**.

It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Figure 6.4.2.

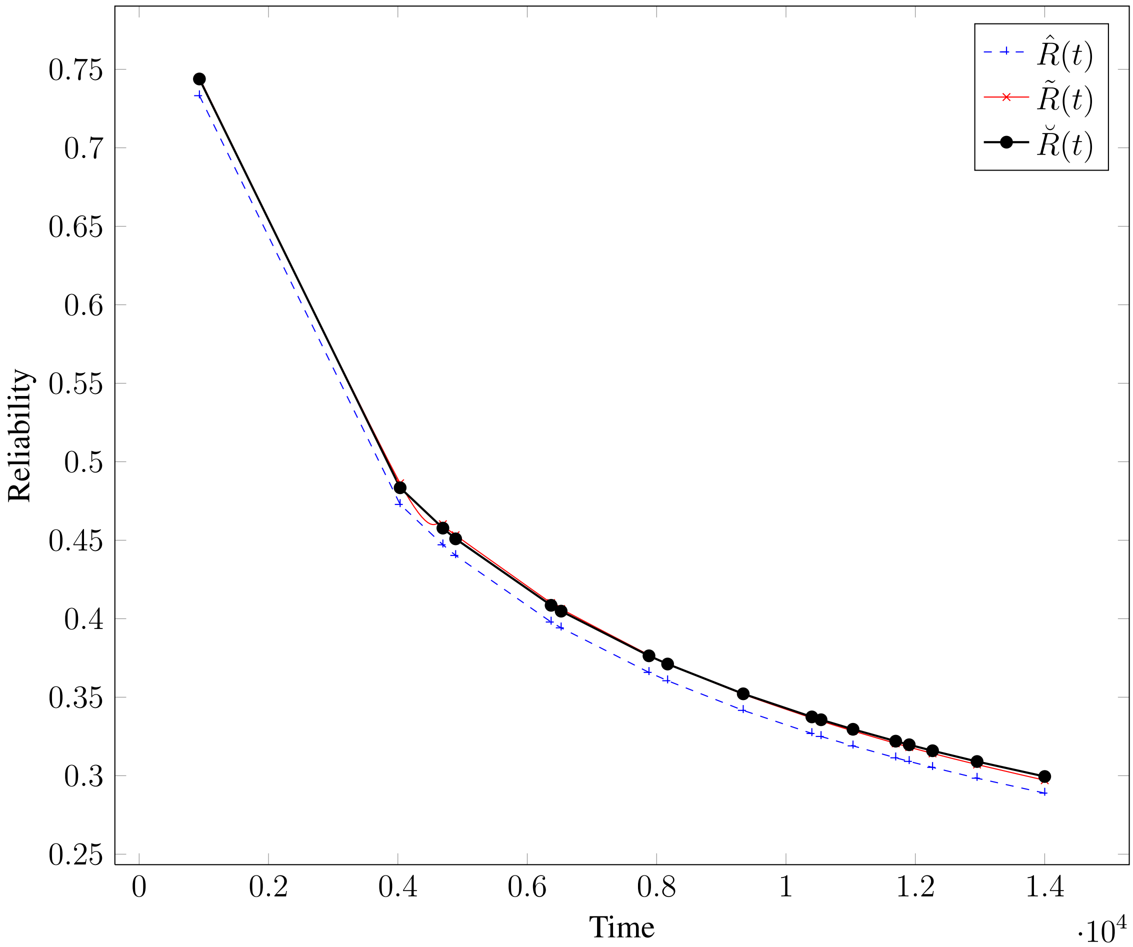

Figure 6.4.2 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Pareto Model (Case study 2)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Case study 3: Failure data set of Lyu

The failure time data for 10 failures obtained by Lyu (Lyu (2004)) are given in Table 6.4.5.

Table 6.4.5 Failure data set of Lyu

Failure number	1	2	3	4	5	6	7	8	9	10
Failure time	7	18	26	36	51	73	93	118	146	181

For this case study, β is estimated as $\hat{\beta} = t_1 = 7$.
Table 6.4.6 denotes the values of MLE, MVUE and the Improved Estimator of reliability. In this table, SD$_{\hat{R}(t)}$, SD$_{\tilde{R}(t)}$ and SD$_{\breve{R}(t)}$ denote the squares of deviations of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ from their corresponding means respectively.
Using the values obtained in Table 6.4.6 in equations (6.4.4) to (6.4.9), we get,

Table 6.4.6 $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Pareto Model (Case study 3)

Failure Number	Failure Time(t)	$\hat{R}(t)$	$\tilde{R}(t)$	$\breve{R}(t)$	SD$_{\hat{R}(t)}$	SD$_{\tilde{R}(t)}$	SD$_{\breve{R}(t)}$
1	7	0.72763	0.74765	0.74498	0.1111	0.11288	0.1111
2	18	0.55769	0.58186	0.57505	0.02669	0.02897	0.02669
3	26	0.491	0.51475	0.50836	0.00935	0.01062	0.00935
4	36	0.43486	0.45726	0.45222	0.00164	0.00208	0.00164
5	51	0.37908	0.3992	0.39644	0.00023	0.00016	0.00023
6	73	0.32709	0.3442	0.34445	0.00452	0.00455	0.00452
7	93	0.29527	0.3101	0.31262	0.00981	0.01032	0.00981
8	118	0.26654	0.27904	0.28389	0.01633	0.01759	0.01633
9	146	0.24293	0.25333	0.26029	0.02292	0.02507	0.02292
10	181	0.22103	0.22931	0.23838	0.03003	0.03326	0.03003

$\overline{\hat{R}}(t) = 0.39431$; $S^2_{\hat{R}(t)} = 0.0258$; $\overline{\tilde{R}}(t) = 0.41167$; $S^2_{\tilde{R}(t)} = 0.0273$;
$\overline{\breve{R}}(t) = 0.41167$; $S^2_{\breve{R}(t)} = 0.0258$.
Hence, using (6.3.2), the bias in $\hat{R}(t)$ is obtained as
$\text{Bias}(\hat{R}(t)) = 0.39431 - 0.41167 = -0.01736$.
Removing this bias from $\hat{R}(t)$, the Improved Estimator of $R(t)$ is obtained as
$\breve{R}(t) = \hat{R}(t) - (-0.01736) = \hat{R}(t) + 0.01736$.
Using equations (6.4.1) to (6.4.3), the coefficient of variation of the three estimators, are respectively obtained as
CV($\hat{R}(t)$)=0.4077, CV($\tilde{R}(t)$)=0.4012 and **CV($\breve{R}(t)$)=0.3905**.
It can be observed that the Improved Estimator $(\breve{R}(t))$ has the least value of the coefficient of variation as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.
Further, the first and third quartiles of $\hat{R}(t)$ are respectively obtained as Q_1=0.2665 and Q_3=0.4910. Thus, the quartile coefficient of dispersion of $\hat{R}(t)$ is obtained as QD($\hat{R}(t)$)=0.2964. The first and third quartiles of $\tilde{R}(t)$ are respectively obtained as Q_1=0.2790 and Q_3=0.5148. Hence, the quartile coefficient of dispersion of $\tilde{R}(t)$ is obtained as QD($\tilde{R}(t)$)=0.2969. Also, the first and third quartiles of $\breve{R}(t)$ are respectively obtained as Q_1=0.2839 and Q_3=0.5084. Thus, the quartile coefficient of dispersion of $\breve{R}(t)$ is obtained as **QD($\breve{R}(t)$)=0.2833**.
It can be observed that the Improved Estimator $\breve{R}(t)$ has the least value of the quartile

coefficient of dispersion as compared to those of $\hat{R}(t)$ and $\tilde{R}(t)$.

The reliability curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ are shown in Fig. 6.4.3.

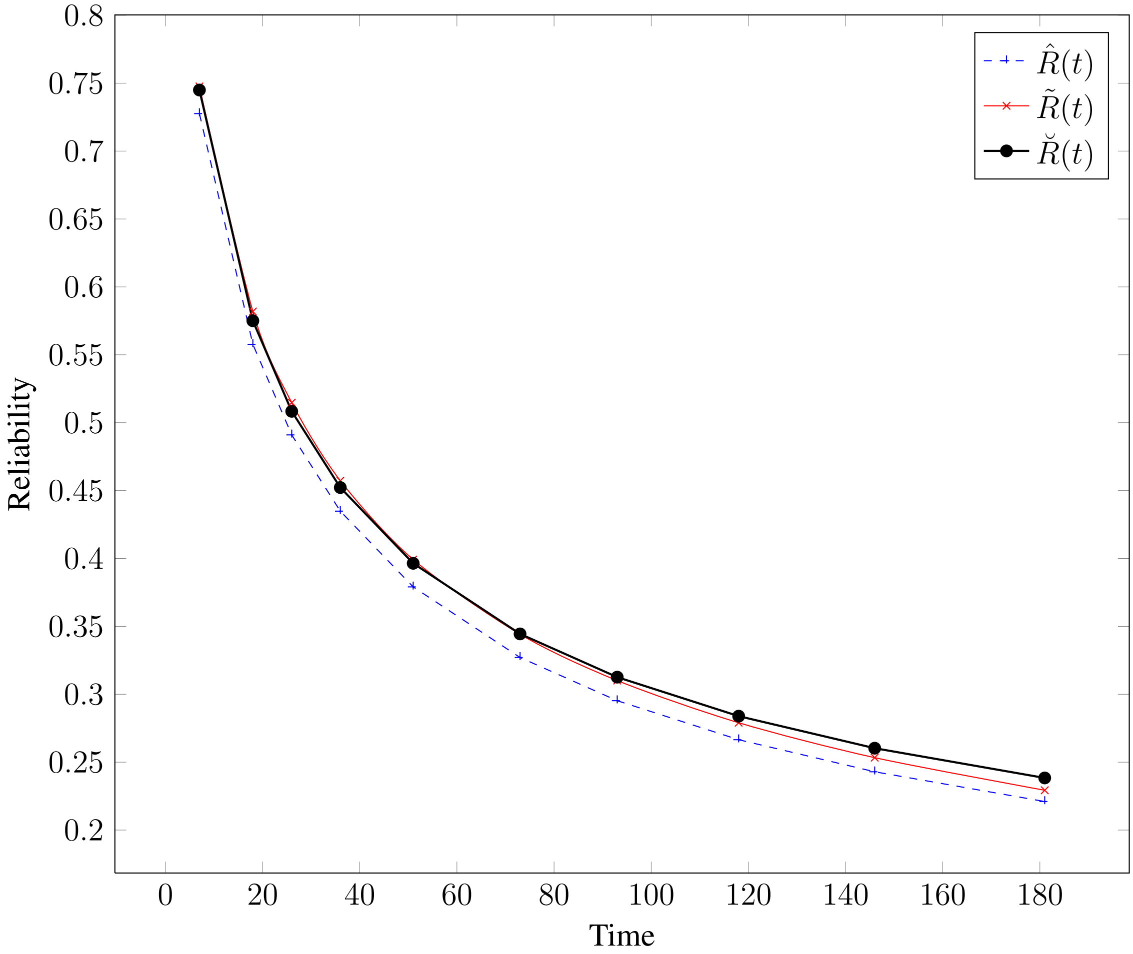

Figure 6.4.3 Curves of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for Pareto Model (Case study 3)

From the three reliability curves, it can be observed that the value of the reliability in the early stages obtained from $\breve{R}(t)$ is more closer to one than the values of reliability obtained from $\hat{R}(t)$ and $\tilde{R}(t)$. Further, it is observed that the estimated values of reliability corresponding to $\breve{R}(t)$ are slightly higher than those of $\hat{R}(t)$ and $\tilde{R}(t)$ for most of the time instances.

Table 6.4.7 shows the consolidated values of the coefficient of variation (CV) of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for all the three case studies. Table 6.4.8 shows the consolidated values of the quartile coefficient of dispersion (QD) of $\hat{R}(t)$, $\tilde{R}(t)$ and $\breve{R}(t)$ for all the three case

Table 6.4.7 Pareto Models (Consolidated1)

Case study	CV($\hat{R}(t)$)	CV($\tilde{R}(t)$)	CV($\breve{R}(t)$)
1	0.388	0.3818	**0.3755**
2	0.2844	0.2778	**0.2754**
3	0.4077	0.4012	**0.3905**

Table 6.4.8 Pareto Models (Consolidated2)

Case study	QD($\hat{R}(t)$)	QD($\tilde{R}(t)$)	QD($\breve{R}(t)$)
1	0.2878	0.2881	**0.2793**
2	0.1492	0.1494	**0.1450**
3	0.2964	0.2969	**0.2833**

studies. From Tables 6.4.7 and 6.4.8, it is observed that, the Improved Estimator ($\breve{R}(t)$) has the least values of the coefficient of variation and quartile coefficient of dispersion than those of MLE ($\hat{R}(t)$) and MVUE ($\tilde{R}(t)$) in all the three case studies, which means that $\breve{R}(t)$ is more efficient than $\hat{R}(t)$ and $\tilde{R}(t)$.
However, to choose the best estimator among the three estimates, the desirable properties of good estimators as mentioned in Section 1.1 of Chapter 1 are to be considered.
Unbiasedness of $\breve{R}(t)$: It has been shown above that $\hat{R}(t)$ is biased for $R(t)$, while $\tilde{R}(t)$ is unbiased for $R(t)$. Since the Improved Estimators are obtained from MLEs just by removing the bias present in the MLEs, they satisfy the unbiasedness property. Thus, $\breve{R}(t)$ is unbiased for $R(t)$.
Sufficiency of $\breve{R}(t)$: Improved Estimator is a function of MLE, which is sufficient. Since any function of sufficient estimator is also sufficient, Improved Estimator is sufficient.
Now, to compare the biased estimator $\hat{R}(t)$ with unbiased estimators $\tilde{R}(t)$ and $\breve{R}(t)$, the coefficient of variation and the quartile coefficient of dispersion are considered as a measure of dispersion to check the efficiency property. The sample results of comparison of coefficients of variation and the quartile coefficients of dispersion for the three estimators indicate that Improved Estimator has least values of coefficient of variation and the quartile coefficient of dispersion as compared to those of MLE and MVUE of $R(t)$, which indicates that the Improved estimators are efficient compared to MLE and MVUE.
Thus, by referring to Table 1.1.1 of Chapter 1, Table 6.4.9 provides the statistical properties satisfied by MLE, MVUE and the Improved Estimator of reliability for exponential class models.

Table 6.4.9 Pareto class models- Properties satisfied by estimators of reliability

	Unbiased	**Sufficient**	**Efficient**
MLE	No	Yes	No
MVUE	Yes	Yes	No
Improved Estimator	**Yes**	**Yes**	**Yes**

It can be seen from Table 6.4.9 that the Improved Estimator satisfies maximum number of properties of estimators as compared to MLE and MVUE of $R(t)$. Hence, it can be inferred that the estimate of reliability obtained using the Improved Estimator, is more efficient than those estimated using the methods of MLE and MVUE.
Hence, it is concluded that $\breve{R}(t)$ gives more accurate value of reliability than $\hat{R}(t)$ and $\tilde{R}(t)$, for Pareto class software reliability models.

Chapter 7

CONCLUSION AND FUTURE DIRECTIONS

Software quality has become a major concern of all software developers. One such measure of software quality is the software reliability, which is the probability of failure-free operation of a computer program in a specified environment for a specified period of time. Failures, which are random in nature, are described by software reliability models. Usually, these models are described in terms of the failure time distributions. Depending on the environment and various other factors, several software reliability models have been developed. Even though there is no single model that fits in all the situations, a proper model depending on the user requirements and specific environment can always be selected. Each such model has certain number of parameters. These parameters play vital role in determining the reliability measures. Reliability of software can be measured using various quantities such as - the Mean Time To Failure (MTTF), the Failure intensity function, the Mean value function, the Failure rate, the Reliability function etc. However, estimating the reliability of a given software for various models through the reliability function is a good mode of assessment of reliability, since it helps the software developers to ensure that user requirements are met and also helps users to decide whether or not to accept the software.

Herein, the finite failures category software reliability models have been considered. The reliability of these models have been estimated by considering their failure time distributions. Exponential, Weibull, Gamma and Pareto class models are considered. The work is aimed at first obtaining the estimates of reliability for these four classes of models using the methods of Maximum Likelihood Estimation (MLE) and Minimum Variance Unbiased Estimation (MVUE). The MVUEs of reliability are always unbiased, while the MLEs are not always unbiased. It is observed that for the four classes of models mentioned above, MLEs are biased estimators. The bias in MLEs are hence obtained using the MVUEs. The MLEs are thus improved by removing the bias present in them, thus obtaining the Improved Estimators for all class of finite failures category

models.
The three estimators, viz, the MLE, the MVUE and the Improved Estimators of reliability are then compared to choose the best estimator amongst them. The desirable properties of the estimators are considered for this purpose. Few case studies from different fields have been considered for comparing the estimators by checking the properties satisfied by them. It is found that the Improved Estimator of reliability satisfies maximum number of desirable properties of good estimators, while MLE and MVUE satisfy only few of them, for all finite failures category models. Hence it is concluded that Improved Estimator gives more accurate value of reliability than MLE and MVUE.
Chapter one gives an overview of the concepts of software reliability, software reliability models and estimation.
Chapter two deals with a detailed literature review in the filed of estimation of parameters and reliability for various classes of software reliability models, the outcome of the review and scope for further research thereby. It also states the objectives and methodology of the proposed research work.
In chapter three, the exponential class software reliability models have been considered and the reliability was estimated using the methods of MLE and MVUE. The Improved Estimator was also obtained. The three estimators were compared using the properties satisfied by them and the Improved Estimator was shown to be the best estimator among the three estimators, by using sample failure data. This Improved Estimator of reliability can be used to estimate the reliability more accurately, for any model, which has exponential failure time distribution.
In chapter four, the Weibull class models were considered and the estimates of reliability were obtained using the methods of MLE and MVUE. The Improved Estimator was also obtained. The three estimators were compared using the properties satisfied by them and the Improved Estimator was shown to be the best estimator among the three estimators, by using sample failure data. This Improved Estimator of reliability can be used to estimate the reliability more accurately, for any model, which has Weibull failure time distribution.
Chapter five deals with the Gamma class models and the reliability estimates were obtained using the methods of MLE and MVUE. The MLE was improved by removing the bias present in it to get the Improved Estimator. By checking the properties satisfied by the three estimates using sample failure data, Improved Estimator was found to be the best estimator as compared to MLE and MVUE. This Improved Estimator is applicable for all software reliability models that have Gamma failure time distribution.
In chapter six, the Pareto class software reliability model has been considered and the estimates of reliability using the methods of MLE and MVUE were obtained. The

estimators are then used to obtain the Improved Estimator. The three estimators are compared using the desirable properties satisfied by them. A set of sample failure data from various fields were considered for this purpose. It was found that the Improved Estimator provides a better estimate of reliability as compared to MLE and MVUE of reliability. This Improved Estimator is applicable for all software reliability models, which have failure time distribution as Pareto.

Though the coefficients of variation and the quartile coefficients of dispersion are considered here as a measures of dispersion to check the efficiency property, the comparison can further be enhanced by using other measures of dispersion. Further, statistical testing methods can also be used for validating the above claim. Also, here it is assumed for Weibull model that only one of the parameters in the model is unknown, by considering the model due to Schick-Wolverton. This assumption may be relaxed when both the parameters in the model are unknown. Hence, the work can be extended in this direction; but in such a case, the complexity may increase. However, a similar procedure may be adopted to obtain the unbiased, sufficient and efficient estimator of the reliability.

Appendix A

Appendix1

Moment Generating function: The moment generating function (mgf) of a random variable X, having the probability function $f(x)$, denoted by $M_X(t)$, is a function that generates moments and is defined by $M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx$,

where t is any real parameter>0.

Result 1: The mgf of the sum of a number of independently and identically distributed random variables is equal to the product of their respective mgfs.
Symbolically, if $X_1, X_2, \ldots X_n$ are independently and identically distributed random variables, then the mgf of their sum $X_1 + X_2 + \ldots + X_n$ is given by
$M_{X_1+X_2+\ldots X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot M_{X_3}(t) \ldots \cdot M_{X_n}(t)$.

Result 2: The mgf of a distribution, if it exists, uniquely determines the distribution. This implies that corresponding to a given probability distribution, there is only one mgf and corresponding to a given mgf, there is only one probability distribution. Hence, $M_X(t) = M_Y(t) \implies X$ and Y are identically distributed.

MGF of some standard distributions:
(i) Exponential distribution: Let $X \sim \mathscr{E}(\Phi)$. Then, its pdf is $f(x) = \Phi e^{-\Phi x}, \quad x > 0$.
Hence, its mgf is given by
$M_X(t) = E(e^{tX}) = \int_0^{\infty} e^{tx} \Phi e^{-\Phi x} dx = \int_0^{\infty} \Phi e^{-x(\Phi - t)} dx = \Phi\Big(\frac{e^{-x(\Phi-t)}}{-(\Phi - t)}\Big)\Big|_0^{\infty} = \frac{\Phi}{\Phi - t}$.
Hence,

$$M_X(t) = \frac{\Phi}{\Phi - t}, \quad \Phi > t. \tag{A.0.1}$$

(ii) Gamma distribution: Let $X \sim G(\alpha, \beta)$. Then, its pdf is
$f(x) = \dfrac{\beta^\alpha}{\Gamma(\alpha)} e^{-x\beta} x^{\alpha-1} \quad x > 0, \alpha, \beta > 0.$
Hence, its mgf is given by
$$M_X(t) = E(e^{tX}) = \int_0^\infty e^{tx} \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-x\beta} x^{\alpha-1} dx = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{(\beta-t)x} x^{\alpha-1} dx$$
$$= \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\beta-t)^\alpha} = \left(\frac{\beta}{\beta-t}\right)^\alpha.$$

Hence,
$$M_X(t) = \left(\frac{\beta}{\beta-t}\right)^\alpha, \quad \beta > t. \tag{A.0.2}$$

Section 1

Sum of independent exponentially distributed random variables is distributed as Gamma:

The concept of mgf can be used to prove this result. Let $X_1, X_2, \ldots X_n$ be independently distributed as $\mathscr{E}(\Phi)$. Then, from (A.0.1), $M_{X_i}(t) = \dfrac{\Phi}{\Phi - t}$.
Hence, the mgf of the sum $X_1 + X_2 + \ldots + X_n$ is obtained using Result 1 as
$$M_{X_1+X_2+\ldots X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot M_{X_3}(t) \ldots \cdot M_{X_n}(t) = \left(\frac{\Phi}{\Phi-t}\right)^n.$$
Comparing it with (A.0.2), it can be seen that $\sum_{i=1}^{n} X_i$ has $G(n, \Phi)$. This result is true for any number of independent random variables. Hence, $\sum_{i=1}^{n-1} X_i$ has $G(n-1, \Phi)$.

Section 2

Power of Weibull distributed random variable is exponential:

Let the r. v. $X \sim W(\Phi, \beta)$. Then, the pdf of X is given by $f(x) = \alpha\beta x^{\beta-1} e^{-\Phi x^\beta}, x > 0$. Consider the transformation $Y = x^\beta$. Then, the inverse transformation is $X = Y^{\frac{1}{\beta}}$. Hence, $\dfrac{dy}{dx} = \beta x^{\beta-1}$. Therefore, the pdf of Y, denoted by $g(y)$ is obtained as $g(y) = f(x) \cdot |\dfrac{dx}{dy}|$, where x is expressed in terms of y. Now, differentiating y with respect to x, $\dfrac{dy}{dx} = \beta x^{\beta-1}$. Hence, $\dfrac{dx}{dy} = \dfrac{1}{\beta x^{\beta-1}}$. Thus, the pdf of Y is obtained as
$$g(y) = \alpha\beta \left(y^{\frac{1}{\beta}}\right)^{\beta-1} e^{-\Phi\left(y^{\frac{1}{\beta}}\right)^\beta} \cdot \frac{1}{\beta\left(y^{\frac{1}{\beta}}\right)^{\beta-1}} = \Phi e^{-\Phi y}, \quad y > 0.$$
Hence $Y \sim \mathscr{E}(\Phi)$.

Section 3

Sum of independent Gamma variates is a Gamma variate:

Let $X \sim G(\alpha, \beta)$. Let $X_1, X_2, \ldots X_n$ be independently distributed as $G(\alpha, \beta)$. Then, from (A.0.2), $M_X(t) = \left(\frac{\beta}{\beta - t}\right)^{\alpha}, \quad \beta > t.$.

Hence, the mgf of the sum $X_1 + X_2 + \ldots + X_n$ is obtained using Result 1 as

$M_{X_1+X_2+\ldots X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot M_{X_3}(t) \ldots \cdot M_{X_n}(t) = \left(\frac{\beta}{\beta - t}\right)^{n\alpha}.$

Comparing it with (A.0.2), it can be seen that $\sum_{i=1}^{n} X_i$ has $G(n\alpha, \beta)$. This result is true for any number of independent random variables. Hence, $\sum_{i=1}^{n-1} X_i$ has $G(\alpha(n-1), \beta)$.

Replacing α by 2 and β by Φ, we have, $X \sim G(2, \Phi)$.

$\therefore M_{X_1+X_2+\ldots X_n}(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdot M_{X_3}(t) \ldots \cdot M_{X_n}(t) = \left(\frac{\Phi}{\Phi - t}\right)^{2n}$, which is the mgf of Gamma random variable with parameters $2n$ and Φ.

Hence, if $X \sim G(2, \Phi)$, then $\sum_{i=1}^{n} X_i \sim G(2n, \Phi)$.

Section 4

Log of Pareto r.v. is exponential r.v.:

Let $X \sim P(\alpha, \beta)$. Then its pdf is given by $f(x) = \frac{\alpha\beta^{\alpha}}{(x+\beta)^{\alpha+1}}$.

Consider the transformation $Y = \ln\left(1 + \frac{X}{\beta}\right)$. Hence, $1 + \frac{X}{\beta} = e^Y$, so that $\frac{X}{\beta} = e^Y - 1$ and hence, $X = \beta(e^Y - 1)$. i.e., $X + \beta = \beta e^Y$. Therefore, $\frac{dx}{dy} = \beta e^y$.

Hence, the pdf of Y is given by $g(y) = f(x)|\frac{dx}{dy}|$, where x is expressed in terms of y.

i.e., $g(y) = \frac{\alpha\beta^{\alpha}}{(\beta e^y)^{\alpha+1}}\beta e^y = \frac{\alpha\beta^{\alpha}}{\beta e^y(\beta e^y)^{\alpha}}\beta e^y = \alpha e^{-\alpha y}$, which is the pdf of exponential distribution with parameter α. Hence, $Y \sim \mathscr{E}(\alpha)$.

Also, since sum of n independent exponential r.v.s is a Gamma r.v. as established in Section 1 above,

$\sum_{i=1}^{n} Y_i \sim G(n, \alpha)$.

Result 3: Let the statistic T have Gamma distribution $G(n, \alpha)$ with pdf $f(t) = \frac{\Phi^n}{\Gamma(n)} e^{-\Phi t} t^{n-1}$. Then, T is a complete statistic.

i.e., $E(h(T)) = 0$ for every Φ implies $h(t) = 0$, where $h(T)$ is any function of T.

Proof: Consider $E(h(T)) = 0$.

i.e., $\int_0^\infty h(t)\frac{\Phi^n}{\Gamma(n)}e^{-\Phi t}t^{n-1}dt = 0$, which implies $\frac{\Phi^n}{\Gamma(n)}\int_0^\infty e^{-\Phi t}\{h(t)\cdot t^{n-1}\}dt = 0$.

i.e., $\mathrm{L}\{h(t)\cdot t^{n-1}\} = 0$, which implies $h(t)\cdot t^{n-1} = \mathrm{L}^{-1}(0)$, where L denotes the Laplace transform.

$\therefore h(t)\cdot t^{n-1} = 0$. This implies either $h(t) = 0$ or $t^{n-1} = 0$.

But, $t^{n-1} \neq 0$. Hence, $h(t) = 0$.

Result 4: For any variable x, $\int_0^\infty e^{-ax}x^{n-1}dx = \frac{\Gamma(n)}{a^n}$.

Appendix B

Appendix2

(i) Consider the transformation $U = X + Y$ and $V = X$, where $U = \sum_{i=1}^{n} T_i$ and $V = T_1$, so that $U = X + Y$ implies $\sum_{i=1}^{n} T_i = T_1 + \sum_{i=2}^{n} T_i$. Thus, the inverse transformation is $X = V$ and $Y = U - V$. i.e., $X = T_1$ and $Y = \sum_{i=1}^{n} T_i - T_1 = \sum_{i=2}^{n} T_i$.
Hence, the Jacobian of this inverse transformation is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial t_1}{\partial \sum_{i=1}^{n} t_i} & \frac{\partial t_1}{\partial t_1} \\ \frac{\partial \sum_{i=2}^{n} t_i}{\partial \sum_{i=1}^{n} t_i} & \frac{\partial \sum_{i=2}^{n} t_i}{\partial t_1} \end{vmatrix} = \begin{vmatrix} \frac{1}{\left(\frac{\partial \sum_{i=1}^{n} t_i}{\partial t_1}\right)} & 1 \\ \frac{1}{\left(\frac{\partial \sum_{i=1}^{n} t_i}{\partial \sum_{i=2}^{n} t_i}\right)} & 0 \end{vmatrix}$$

$$\text{i.e.,} J = \begin{vmatrix} \frac{1}{\left(\frac{\partial (t_1 + \sum_{i=2}^{n} t_i)}{\partial t_1}\right)} & 1 \\ \frac{1}{\left(\frac{\partial (t_1 + \sum_{i=2}^{n} t_i)}{\partial \sum_{i=2}^{n} t_i}\right)} & 0 \end{vmatrix} = \begin{vmatrix} \frac{1}{1+0} & 1 \\ \frac{1}{0+1} & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

Hence, $\mid J \mid = 1$.

(ii) Consider the transformation $U = X + Y$ and $V = X$, where $U = \sum_{i=1}^{n} T_i^{\beta}$ and $V = T_1^{\beta}$, so that $U = X + Y$ implies $\sum_{i=1}^{n} T_i^{\beta} = T_1^{\beta} + \sum_{i=2}^{n} T_i^{\beta}$. Thus, the inverse transformation is $X = V$ and $Y = U - V$. i.e., $X = T_1^{\beta}$ and $Y = \sum_{i=1}^{n} T_i^{\beta} - T_1^{\beta} = \sum_{i=2}^{n} T_i^{\beta}$.

Hence, the Jacobian of this inverse transformation is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial t_1^\beta}{\partial \sum_{i=1}^n t_i^\beta} & \frac{\partial t_1^\beta}{\partial t_1^\beta} \\ \frac{\partial \sum_{i=2}^n t_i^\beta}{\partial \sum_{i=1}^n t_i^\beta} & \frac{\partial \sum_{i=2}^n t_i^\beta}{\partial t_1^\beta} \end{vmatrix} = \begin{vmatrix} \frac{1}{\left(\frac{\partial \sum_{i=1}^n t_i^\beta}{\partial t_1^\beta}\right)} & 1 \\ \frac{1}{\left(\frac{\partial \sum_{i=1}^n t_i^\beta}{\partial \sum_{i=2}^n t_i^\beta}\right)} & 0 \end{vmatrix}$$

$$\text{i.e.,} J = \begin{vmatrix} \frac{1}{\left(\frac{\partial (t_1^\beta + \sum_{i=2}^n t_i^\beta)}{\partial t_1^\beta}\right)} & 1 \\ \frac{1}{\left(\frac{\partial (t_1^\beta + \sum_{i=2}^n t_i^\beta)}{\partial \sum_{i=2}^n t_i^\beta}\right)} & 0 \end{vmatrix} = \begin{vmatrix} \frac{1}{1+0} & 1 \\ \frac{1}{0+1} & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

Hence, $\mid J \mid = 1$.

(iii) Consider the transformation $U = X + Y$ and $V = X$, where $U = \sum_{i=1}^n \ln(1 + \frac{T_i}{\beta})$ and $V = \ln(1 + \frac{T_1}{\beta})$, so that $U = X + Y$ implies $\sum_{i=1}^n \ln(1+\frac{T_i}{\beta}) = \ln(1+\frac{T_1}{\beta}) + \sum_{i=2}^n \ln(1+\frac{T_i}{\beta})$.

Thus, the inverse transformation is $X = V$ and $Y = U - V$. i.e., $X = \ln(1 + \frac{T_1}{\beta})$ and $Y = \sum_{i=2}^n \ln(1 + \frac{T_i}{\beta})$.

Hence, the Jacobian of this inverse transformation is given by

$$J = \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix} = \begin{vmatrix} \frac{\partial \ln(1+\frac{t_1}{\beta})}{\partial \sum_{i=1}^n \ln(1+\frac{t_i}{\beta})} & \frac{\partial \ln(1+\frac{t_1}{\beta})}{\partial \ln(1+\frac{t_1}{\beta})} \\ \frac{\partial \sum_{i=2}^n \ln(1+\frac{t_i}{\beta})}{\partial \sum_{i=1}^n \ln(1+\frac{t_i}{\beta})} & \frac{\partial \sum_{i=2}^n \ln(1+\frac{t_i}{\beta})}{\partial \ln(1+\frac{t_1}{\beta})} \end{vmatrix} = \begin{vmatrix} \frac{1}{1+0} & 1 \\ \frac{1}{0+1} & 0 \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

Hence, $\mid J \mid = 1$.

Bibliography

Chauhan, S. K., Mishra, R., and Singh, R. B. (2016). Estimation of software reliability on the basis of bits for embedded system. *International Journal of Engineering Research and General Science*, 4(2):131–135.

Dale, C. (1991). The assessment of software reliability. *Reliability Engineering and System Safety*, 34(1):91–103.

Denton, J. A. (1999). *Accurate Software Reliability Estimation*. PhD thesis, Colorado State University Fort Collins, Fort Collins, Colorado.

Frenkel, I., Gertsbakh, I. B., and Khvatskin, L. V. (2003). Parameter estimation and hypothesis testing for nonhomogeneous poisson process. *Transport and telecommunication*, 4(2):9–17.

Garg, R. and Sharma, K. (2010). Performance analysis of software reliability models using matrix method. *International Journal of Computer and Information Engineering*, 4(11):1646–1653.

Goseva-Popstojanova, K. and Trivedi, K. S. (2001). Architecture-based approach to reliability assessment of software systems. *Elsevier journal of performance evaluation*, 45(2-3):179 – 204.

Guen, H. L., Marie, R., and Thelin, T. (2004). Reliability estimation for statistical usage testing using markov chains. In *Proc., 15th international symposium on software reliability engineering*, pages 54–65. IEEE.

Gupta, S. C. and Kapoor, V. K. (1996). *Fundamentals of Mathematical Statistics*. Sultan Chand & Sons, New Delhi.

Guure, C. B., Ibrahim, N. A., Adam, M. B., Bosomprah, S., and Ahmed, A. O. M. (2013). Bayesian parameter and reliability estimate of weibull failure time distribution. *Bulletin of the Malaysian Mathematical science society*, 37(3):611–632.

Hishitani, J., Yamada, S., and Osaki, S. (1990). Comparison of two estimation methods of the mean time interval between software failures. In *Proc., IEEE Ninth Annual International Phoenix Conference on Computers and Communications*, pages 418–424. IEEE.

Hossain, S. A. and Dahiya, R. C. (1993). Estimating the parameters of a non homogeneous poisson process model for software reliability. *IEEE transactions on reliability*, 42(4):604–612.

Jacoby, R. and Tohma, Y. (1991). Parameter value computation by least square method and evaluation of software availability and reliability at service operation by the hyper-geometric distribution software reliability growth model. In *Proc., IEEE 13th international conference on software engineering*, pages 226–237. IEEE.

Joe, H. and Reid, N. (1985). On the software reliability models of jelinki-moranda and littlewood. *IEEE Transactions on Reliability*, R-34(3):216 – 218.

Joseph, S. (2014). *Reliability Estimation of Open Source Software based Computational Systems*. PhD thesis, Cochin University of Science and Technology, Cochin.

Kantham, R. R. L. and Rao, R. S. (2009). Pareto distribution: A software reliability growth model. *International Journal of Performability Engineering*, 5(3):275–281.

Kaur, G. and Bahl, K. (2014). Software reliability, metrics, reliability improvement using agile process. *International Journal of Innovative Science, Engineering and Technology*, 1(3).

Kaur, M., Singh, S., and Rakshit, M. (2013). A review of various metrics used in software reliability. *International Journal of Computer Science and Engineering Technology*, 4(7):874–876.

Keiller, P. A., Littlewood, B., and Sofer, D. R. M. A. (1983). *On the quality of software reliability prediction*. Springer, Berlin, Heidelberg.

Khalaf, K. and Mustafa, T. (2009). Software reliability modeling using soft computing technique. *European journal of scientific research*, 26(1):154–160.

Kim, T., Lee, K., and Baik, J. (2015). An effective approach to estimating the parameters of software reliability growth models using a real-valued genetic algorithm. *Journal of Systems and Software*, 102:134–144.

Kiran, N. R. and Ravi, V. (2008). Software reliability prediction by soft computing techniques. *Journal of Systems and Software*, 81(4):576–583.

Kuo, L. and Yang, T. (1995). Baysian computation of software reliability. *Journal of Computational and Graphical Statistics*, 4(1):65–82.

Lavanya, G., Neeraja, K., Basha, S. A., and Sangeetha, Y. (2017). Parameter estimation of goel-okumoto model by comparing aco with mle method. *International Research Journal of Engineering and Technology*, 4(3):1605–1615.

Ledoux, J. (2002). Littlewood reliability model for modular software and poisson approximation.

Littlewood, B. (1981). Stochastic reliability growth: A model for fault-removal in computer programs and hardware design. *IEEE Transactions on Reliability*, R-30(4):313–320.

Littlewood, B. (1991). Software reliability modeling: Achievements and limitations. In *Proc., Advanced Computer Technology, Reliable Systems and Applications*, pages 336–344. IEEE.

Lyu, M. R. (2004). *Hand book of Software Reliability Engineering*. IEEE Computer Society Press, McGraw Hill, California, U.S.A.

Mohan, K. K., Verma, A., and Srividya, A. (2010). Software reliability estimation through black box and white box testing at prototype level. In *Proc., IEEE 2nd International Conference on Reliability, Safety and Hazard - Risk-Based Technologies and Physics-of-Failure Methods*. IEEE.

Musa, J. D., Iannino, A., and Okumoto, K. (1991). *Software Reliability Measurement, Prediction, Application*. International Edition, Mc-Graw Hill, New York, USA.

Nagar, P. and Thankachan, B. (2012). Applications of goel-okumoto model in software reliability measurement. *Special issue of International Journal of Computer Applications on issues and challenges in networking, intelligence and computing technologies*, pages 1–3.

Ohba, M. (1984). *Inflection S-shaped Software Reliability Growth Models*, volume 235. Springer Verlag, Merlin.

Okamura, H., Ando, M., and Dohi, T. (2007). A generalized gamma software reliability model. *Systems and Computers in Japan*, 38(2):81–90.

Okamura, H., Watanabe, Y., Dohi, T., and Osaki, S. (2003). An estimation of software reliability models based on em algorithm. *Electronics and Communication in Japan (Part III; Fundamental electronic Science*, 86(6):29–37.

Padberg, F. (2001). A fast algorithm to compute maximum likelihood estimates for the hypergeometric software reliability model. In *Proc., 2nd Asia-Pacific conference on quality software*, pages 40–49. IEEE.

Park, G. Y. and Jang, S. C. (2014). A software reliability estimation method to nuclear safety software. *Nuclear Engineering and Technology*, 46(1):55–62.

Quereshi, M. A. and Jeske, D. R. (1997). Using proxy failure times with the jelinski-moranda software reliability model. In *Proc., IEEE 8th international symposium on software reliability engineering*, pages 358–365, Albuquerque, NM, USA,. IEEE.

Quyoum, A., din Dar, M. U., and Quadri, S. M. K. (2010). Improving software reliability using software engineering approach- a review. *International Journal of Computer Applications*, 10(5):41–47.

Ramani, S., Gokhale, S. S., and Trivedi, K. S. (2000). Software reliability estimation and prediction tool. *Performance Evaluation*, 39(1-4):37–60.

Ramasamy, S. and Lakshmanan, I. (2017). Machine learning approach for software reliability growth modeling with infinite testing effort function. *Mathematical Problems in Engineering*, 2017(1):1–6.

Rebello, S. and Goyal, N. K. (2010). Software system reliability and safety assessment: An extended FMEA approach. *International Journal of Reliability and Safety*, 4(4):366–380.

Schick, G. J. and Wolverton, R. W. (1973). Predicting the reliability of software systems using fuzzy logic. In *proceedings of the operations Research, Physica-Verlag, Wurzburg-Wein*, pages 395–422.

Sehgal, P. and Meenal (2016). Software reliability estimation using artificial neural networks. *International Journal of Research in Management, Science and Technology*, 4(2):102–107.

Shanmugam, L. and Florence, L. (2012). A comparison of parameter best estimation method for software reliability models. *International Journal of Software Engineering and Applications (IJSEA)*, 3(5):91–102.

Singh, B., Viveros, R., and Parnas, D. L. (1997). *Estimating Software Reliability Using Inverse Sampling*. Communications Research Laboratory, McMaster University.

Singpurwalla, N. D. and Soyer, R. (1996). *Assessing the reliability of software; An overview in Reliability and Maintenance of complex systems*, volume 154 of *NATO ASI*. Springer Berlin Heidelberg.

Sinha, S. K. and Kale, B. K. (1980). *Life Testing and Reliability Estimation*. Wiley Eastern Limited, New Delhi.

Sohnlein, S., Saglietti, F., Bitzer, F., Meitner, M., and Baryschew, S. (2010). Software reliability assessment based on the evaluation of operational experience. *Lecture Notes in Computer Science*, 5787(1):24–38.

Trivedi, K. S. (2012). *Probability and Statistics with Reliability, Queuing and Computer Science Applications*. Wiley India Edition, Navi Mumbai, second edition.

Turk, A., Lutfiah, Alsolami, G., and Eftekhar (2016). Jelinski-moranda software reliablity growth model : A brief literature and modification. *International Journal of Software Engineering and Applications*, 7:33–44.

Turk, L. I. A. and Alsolami, E. G. (2016). On generalized littlewood- verrall model for software reliability with applications. *International Journal of Development Research*, 6(11):10313–10330.

Xu, J. and Yao, S. (2016). Software reliability growth model with partial differential equations for various debugging processes. *Mathematical Problems in Engineering*, 2016:1–13.

Yamada, S., Ohba, M., and Osaki, S. (1983). S-shaped reliability growth modeling for software error detection. *IEEE Transactions on Reliability*, R-32(5):475–478.

Yang, M. C. K. and Chao, A. (1995). Reliability -estimation and stopping-rules for software testing, based on repeated appearances of bugs. *IEEE Transactions on Reliability*, 44(2):315–321.

List of Publications

Journals

1. B. Roopashri Tantri and Murulidhar N. N. (2014). “An efficient estimator of reliability for exponential class software reliability models ”, *Lecture Notes on Software Engineering*, 2(3), 201-204.

2. B. Roopashri Tantri and Murulidhar N. N. (2014). “Convergence of MLE to MVUE of reliability for exponential class software reliability models", *International Journal on Recent and Innovation Trends in Computing and Communication*, 2(8), 2133-2136.

3. B. Roopashri Tantri and Murulidhar N. N. (2015). “MVUE of failure rate for exponential class software reliability models ”, *International Journal of Advanced Research in Computer Science and Software Engineering*, 5(4), 347-350.

4. B. Roopashri Tantri and Murulidhar N. N. (2017). “Estimation of Software Reliability for Littlewood Pareto Failure Time Model", *International Journal of Software Engineering and Its Applications*, 11(6), 25-34.

5. B. Roopashri Tantri and Murulidhar N. N. “Comparison of Reliability Estimates of Gamma Failure Time Software Reliability Model", Accepted for publication in *ARPN Journal of Engineering & Applied Sciences*.

Conference Proceedings / Book Chapters

1. Murulidhar N. N. and B. Roopashri Tantri (2013). "Comparison of Parameters of Jelinski-Moranda Model with the Least Square Estimated Values", *Proc., 3rd International Engineering Symposium - IES 2013*, Kumamoto University, Japan, 1-4.

2. B. Roopashri Tantri and Murulidhar N. N. (2015). "Reliability Analysis of Exponential Models Based on Skewness and Kurtosis", *Emerging Research in Computing, Information, Communication and Applications : ERCICA 2015*, **Springer** (Book Chapter), Vol 1, 53- 59.

3. B. Roopashri Tantri and Murulidhar N. N. (2016). "Software Reliability Estimation of Gamma Failure Time Models", *Proc., International Conference on System Reliability and Science (ICSRS 2016)*, Paris, France, **IEEE**, 105-109.

4. B. Roopashri Tantri and Murulidhar N. N. (2018) "Comparison of Reliability Estimates of Gamma Failure Time Software Reliability Model", *Proc., IOSRD 73rd International Conference on Future Trends in Engineering and Business*, Chennai, India, Vol. 73, 58-61.

5. B. Roopashri Tantri and Murulidhar N. N. "Software Reliability Estimation of Schick-Wolverton Rayleigh Failure Time Model", Communicated to *26th ISSAT International Conference on Reliability and Quality in Design*.

BIO DATA

Name : B. Roopashri Tantri

Email Id : roopatantri@gmail.com

Date of Birth : 29-05-1970

Permanent Address : B. Roopashri Tantri,
Flat No. 310, Jana Jeeva Castle,
8^{th} Cross, Canara Bank Layout,
Post: Vidyaranyapura
Bengaluru - 560097, Karnataka.

Educational Qualifications:

Degree	Year of Passing	Institute
B.Sc. (Physics, Mathematics, Statistics)	1991	M. G. M. College, Udupi. Mangaluru University.
M.Sc. (Statistics)	1993	Mangaluru University.
M.Tech. (Systems Analysis & Computer Applications)	2000	NITK, Surathkal.