CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Qatar University Institutional Repository

QATAR UNIVERSITY

COLLEGE OF ARTS AND SCIENCES

INFERENCE ABOUT THE GENERALIZED EXPONENTIAL QUANTILES BASED ON PROGRESSIVELY TYPE-II CENSORED DATA

BY

RASHA A. RIHAN

A Thesis Submitted to

the Faculty of the College of Arts and Sciences

in Partial Fulfillment of the Requirements for the Degree of

Masters of Science in Applied Statistics

June 2019

© 2019 Rasha. All Rights Reserved.

COMMITTEE PAGE

The members of the Committee approve the Thesis of Rasha defended on 09/04/2019.

Prof. Dr. Ayman Suleiman Bakleezi
Thesis/Dissertation Supervisor

Dr. Faiz Ahmed Elfaki
Committee Member

Dr. Saddam Akber Abbasi
Committee Member

Approved:

Rashid Al-Kuwari, Dean, College of College of Arts and Sciences

ABSTRACT

RIHAN, RASHA, ADNAN, Masters : June : 2019, Applied Statistics

Title: INFERENCE ABOUT THE GENERALIZED EXPONENTIAL QUANTILES BASED on PROGRESSIVELY CENSORED DATA

Supervisor of Thesis: Prof. Dr. Ayman Suleiman Bakleezi.

In this study, we are interested in investigating the performance of likelihood inference procedures for the p^{th} quantile of the Generalized Exponential distribution based on progressively censored data. The maximum likelihood estimator and three types of classical confidence intervals have been considered, namely asymptotic, percentile, and bootstrap-t confidence intervals. We considered Bayesian inference too. The Bayes estimator based on the squared error loss function and two types of Bayesian intervals were considered, namely the equal tailed interval and the highest posterior density interval. We conducted simulation studies to investigate and compare the point estimators in terms of their biases and mean squared errors. We compared the various types of intervals using their coverage probability and expected lengths. The simulations and comparisons were made under various types of censoring schemes and sample sizes. We presented two examples for data analysis, one of them is based on simulated data set and the other one based on a real lifetime data. Finally, we compared the classical inference and the Bayesian inference procedures. We concluded that Bias and MSE for classical statistics estimators show bitter results than the Bayesian estimators. Also, Bayesian intervals which attain the nominal error rate have the best average widths. We presented our conclusions and discussed ideas for possible future research.

DEDICATION

This thesis is dedicated to my family.

For their infinite love, care, and inspiration throughout the years.

ACKNOWLEDGMENTS

"All praises and thanks to Allah on what we got and without his gaudiness, we wouldn't get to this!" I would like to express my thankful to all my family and friends for their patient and encouragement they offered me throughout my entire study. I would like to express my appreciation to my supervisor Prof. Dr. Ayman Bakleezi for his patient and support, as well as his useful comments, observations, and follow-up throughout the duration of learning and working on this master thesis. Furthermore, I would like to thank Dr. Faiz Elfaki and Dr. Saddam Abbasi not for only serving on my thesis dissertation committee, but also for teaching me statistical courses during my years of study. And all thanks and appreciation to all the doctors who taught me the science during my studies for the master and bachelor coursework. Last, but not least I would like to express my thanks and gratitude to all my teachers who contributed to my education throughout my years of study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

*Note to Student: Please note that the capitalization of titles on List of Tables and List of Figures are both acceptable. For example, Page 7 uses Title Case, while page 8 uses Sentence Case. We only ask that you be consistent with your choice throughout the document.

CHAPTER 1: INTRODUCTION

In statistical analysis a lifetime or failure time data is wildly used in many areas. Then the lifetime can be defined by having time scale, time origin and an event, which noted as failure or death. In this study we are interested in censored lifetime data especially a progressively type-II censored data. The lifetime data is called censored when an information about an individual survival time is available, but the survival time is not known exactly. The progressively type-II censored data will be approached if the deletions are carried out at an observed failure time. The analysis of this type of data is important in many sciences like the biomedical, engineering, and social sciences. For more explanation, lifetime distribution methodology applications are mainly used to investigate the manufactured items' durability or to study human diseases and their treatment. The interest in analyzing such data is not new. In about 1970, dealing with this type of data had been expanded rapidly depending on methodology, theory, and fields of application. Since about 1980, software packages for lifetime data had been developed widely with a lot of new features and packages.

The lifetime in general is a positive random variable T assumed to be continuous with probability density function (pdf) $f(T)$. Some examples of common lifetime distributions are the exponential distribution, Weibull distribution, log normal distribution, log logistic distribution, and generalized exponential distribution. Censored data occurs when failure times for some units have not been completely observed. There are many causes for censoring. One of these causes is the absence of an event during the study time. If an observation has been lost to follow-up from the study because of death or any other reasons, then this can be a reason for censoring.

In this thesis, we consider the likelihood inference of the quantiles of the Generalized Exponential distribution based on progressively type II censored data. Then, our main research problems are:

1- Investigating and studying the performance of two types of statistical inference, namely; classical inference and Bayesian inference.
2- Considering point estimation as well as interval estimation. Three types of classical confidence intervals have been constructed, namely; the asymptotic interval, the percentile interval, and bootstrap-t interval. For the Bayesian intervals, we considered equal tail intervals as well as the highest probability density (HPD) interval.
3- Comparing between the classical statistical inference point estimation and the Bayesian inference point estimation.

Problems 1, 2, and 3 are presented in chapters 2, 3, and 5, respectively. The rest of this chapter gives a brief explanation and a review of some literature related to our study.

1.1 The Generalized Exponential Distribution

The Generalized Exponential distribution denoted by GE, is a relatively new distribution applied on the life time data. It is introduced by (Gupta & Kundu, Generalized Exponential Distributions, 1999) as a possible alternative for Weibull and Gamma distributions. The main idea of using GE distribution instead of Weibull or gamma distributions is that has the many properties are quite like the gamma distribution, but the distribution function is like the one of Weibull distribution, which will let the computation simpler. The GE distribution is skewed to the right and its monotone hazard function is like the monotone hazard functions of gamma and Weibull distributions. (Khan, 1987) assumed that the GE distribution has two parameters and in case of having an additional parameter which called the location parameter. The GE distribution with an additional parameter fits many situations of life and reliability test results whereas the coefficient of variation of the data is significantly greater than the unity. The GE distribution can be used as an alternative to the Weibull and gamma families for analyzing lifetime data. (Gupta & Kundu, Generalized Exponential Distributions, 1999) presented the distribution function, the probability density function and properties of the distribution. They considered statistical inference techniques for GE distribution.

On the other hand, (Gupta & Kundu, Generalized Exponential Distribution: Statistical Inferences, 2002) derived the maximum likelihood estimation of the unknown parameters of a generalized exponential distribution for both, complete sample and censored sample. They presented the MLEs for both types of censoring, type I and type II. The consistency and the asymptotic normality results of the MLE's of the GE distribution had been provided by the researchers, in case of complete data. On the other hand, in case of type I censored data and if the data were in grouped form, the Fisher Information matrix had been provided. Also, (Gupta & Kundu, Generalized Exponential Distribution: Existing Results and Some Recent Developments, 2007) assumed that the GE distribution is more useful for analyzing lifetime data than gamma distribution, Weibull distribution or log-normal distribution. They presented the source of this model. Also, some properties and different estimation procedures had been presented. And (Gupta & Kundu, Generalized Exponential Distribution: Bayesian Estimations, 2008) derived the Bayesian estimators for the two unknown parameters of the GE distribution. They assumed gamma distribution as prior distributions for both shape and scale parameters.

This research is based on the generalized exponential distribution. First, we will present the GE distribution with three parameters. A random variable X said to be generalized exponential distributed if X has the following distribution function

$$F(x;\alpha,\lambda,\mu)=(1-e^{-(x-\mu)/\lambda})^{\alpha} \qquad (x>\mu,\alpha>0,\lambda>0). \tag{1}$$

The corresponding density function is

$$f(x;\alpha,\lambda,\mu)=\frac{\alpha}{\lambda}(1-e^{-(x-\mu)/\lambda})^{\alpha-1}e^{-(x-\mu)/\lambda} \quad (x>\mu,\alpha>0,\lambda>0), \tag{2}$$

where, the shape parameter is α, the scale parameter is λ and the location parameter is μ. The GE distribution can be denoted as $GE(\alpha,\lambda,\mu)$. The behavior of the hazard function for different values of the shape parameter α and its relation with the Gamma and Weibull distributions is explained in the table below.

Table 1. The shape parameter α different values.

Parameter:	**Gamma**	**Weibull**	**GE**
$\alpha=1$	$1/\lambda$	$1/\lambda$	$1/\lambda$
$\alpha>1$	Increasing from 0 to $1/\lambda$	Increasing from 0 to ∞	Increasing from 0 to $1/\lambda$
$\alpha<1$	Decreasing from ∞ to $1/\lambda$	Decreasing from ∞ to 0	Decreasing from ∞ to $1/\lambda$

This research will consider only two parameters. If the location parameter is zero as in most applications of lifetime data models. The shape and scale parameters denoted by θ and σ respectively. Therefore, our density and cumulative functions will be defined as follows

$$f(x;\theta,\sigma)=\frac{\theta}{\sigma}e^{-x/\sigma}(1-e^{-x/\sigma})^{\theta-1} \qquad (x>0,\theta>0,\sigma>0). \qquad (3)$$

$$F(x;\theta,\sigma)=(1-e^{-x/\sigma})^{\theta} \qquad (x>0,\theta>0,\sigma>0). \qquad (4)$$

The probability density function proposed in equation (3), has been plotted in the figure below with indicating different parameters sets. These parameters are $(\theta_1,\sigma_1)=(2,1.2)$, $(\theta_2,\sigma_2)=(1.2,0.5)$, $(\theta_3,\sigma_3)=(1.5,0.7)$, $(\theta_4,\sigma_4)=(1.7,0.9)$ and $(\theta_5,\sigma_5)=(2.3,1.5)$.

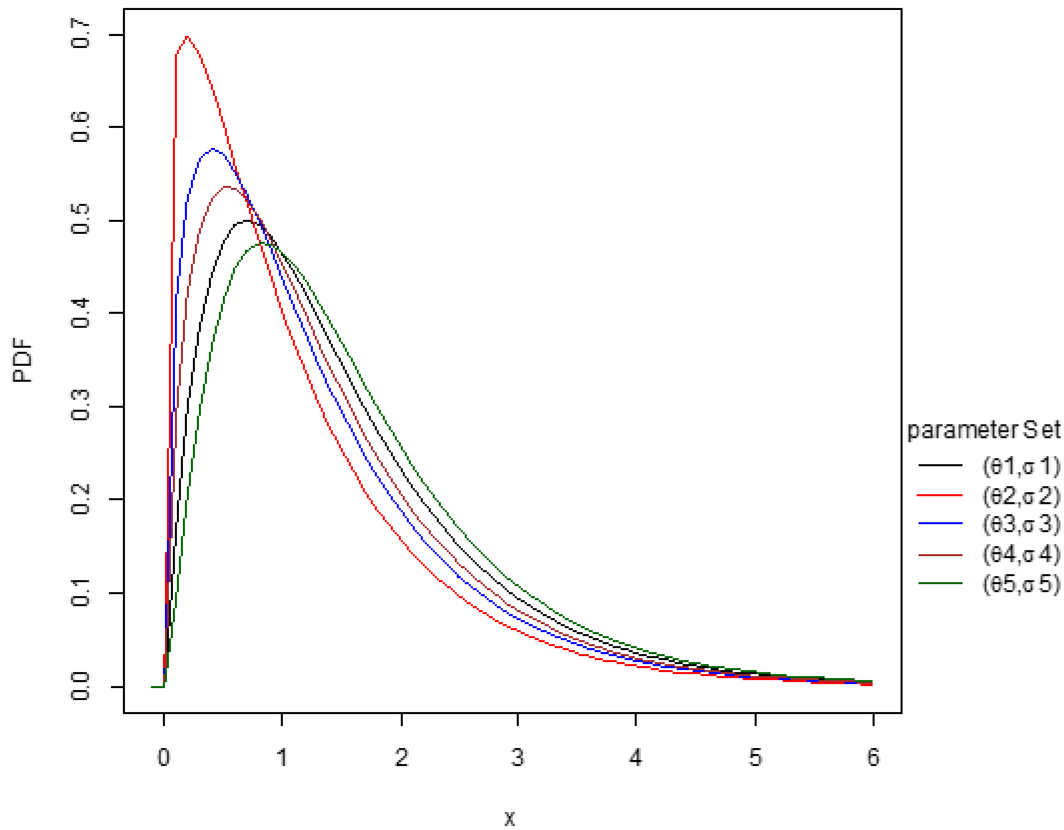

Figure 1. PDF plot for GE distribution.

1.2 Quantiles

Quantiles can be found in many different areas of statistics. It can be applied in many fields like finance, investment, economics, engineering and medicine. The range of a probability distribution can be divided into continuous intervals with equal probabilities by cut points called quantiles. A finite set can be partitioned into q subsets of equal sizes by values called *q*-quantiles. Historically, there are many uses of sample quantiles in statistics. (Eubank, 1984). In 1846, Quetelet used the probable error of a distribution estimator based on the semi-interquartile range. Also, quantiles like the median had been discussed by (Galton, 1889) , (Edgeworth, Progressive Means, 1886) and (Edgeworth, Review of Fisher's Mathematical Investigations, 1893). Two studies had been made in 1899 by (Sheppard, 1899) and then by (Pearson, 1920) were interested in the problem of choosing the optimal quantile for the mean estimation and the normal distribution's standard deviation by the subsets of the sample quantiles' linear functions. The asymptotic distribution of the sample quantiles derivation had been discussed in detail in (Pearson, 1920). Simirnoff (1935) explained the behavior of a sample quantile in case of having large sample. Also, he introduced the limiting distribution's strict derivation. These results had been generalized in 1946 by (Mosteller, 1946), which were supported by (Ogawa, 1951), was highly intentioned with the idea of having quantiles to be an estimation tools in location and scale parameter models. Later, quantiles have been used widely in problems of the classical and the robust statistical inference. Also, quantiles were very important in (Tukey, 1977) and (Parzen, 1979a) works. Which were on the exploratory data analysis and the nonparametric data modeling.

The p^{th} quantile for any variable T is the value t_p such that

$$\Pr\left(T \leq t_p\right) = p.$$

Where, $t_p = F^{-1}(p)$. Noticed that, the pth quantile can be referred as the 100 pth percentile of the distribution. (Pfeiffer, 1990) introduced some properties for the p^{th} quantile, which are shown below:

1- Suppose that the cumulative distribution function $F(x)$ is continuous and strictly increasing on a closed interval $[a, b]$, then the p^{th} quantile function t_p is also continuous and strictly increasing on the closed interval $[F(a), F(b)]$.

2- If the cumulative distribution function $F(x)$ has a jump at $x = a$, then the p^{th} quantile function $t_p(p) = a$ for $p \in (F(a-0), F(a)]$.

3- If $F(x) = F(a)$ for $x \in [a, b)$ for $F(a-c) < F(a)$ and $F(b+c) > F(a)$, $\forall\ c > 0$, then $t_p\big(F(a)\big) = a$ and $p > F(a)$ implies that $t_p(p) \geq b$.

4- The p^{th} quantile function t_p is left continuous. For more explanation, the t_p function is the inverse function of the cumulative distribution function $F(x)$, then to get the graph of t_p, we may reflect the graph of $F(x)$ in the main diagonal. By considering the jump of $F(x)$ to be a horizontal line. So, the jump of $F(x)$ will be a horizontal interval for t_p, and a horizontal interval for $F(x)$ becomes a jump for t_p. Then $F(x)$ is right continuous and t_p is left continuous.

5- Suppose that $F\big(t_p(p) - 0\big) \leq p \leq F(t_p(p))$, therefore if $F(x)$ is continuous at $t_p(p)$, then $F\left(t_p(p)\right) = p$.

1.3 Progressively Censored Data

In statistics, economics, engineering, and medical researches, censoring is a condition where the value of measurement or observation is only partially known. Generally, if the exact survival time is not known, but we have some information about the survival time, in this case the resulting data is said to be censored. Therefore, in survival analysis, time until an event occurs is the variable of interest. The event describes death, disease, or some other individual experience. Survival data has many applications in biomedical science, industrial reliability (for example, reliability engineering, such as lifetime of electronic devices, components, or systems), sociology (for example, period of first marriage), marketing (for example, length of newspaper or magazine contribution), and so on. Now, we review some examples of the survival analysis. Suppose our experiment is about leukemia patients, where the event of interest (failure) is "going out of the reduction" and the outcome is "time in a week until a person goes out of the reduction". The next example is in sociology and it about repetition, where the event is "getting rearrested" and the outcome is "time in weeks until rearrested". As another example, suppose we got a data about transplant patients, the event in this situation is "death" and the outcome is "time in months from receiving a transplant until death".

(Balakrishnan & Cramer, The Art of Progressive Censoring, 2014) As mentioned before, the censored data problem is that the observed data of a variable is partially known. The problem is related to the missing data, where the observed value of some variable is unknown. Therefore, it can be said that the following causes are the reasons of why censoring occurs in the data:

1- There are many reasons that would make a person withdraw the study.
2- During a period of the study, an individual lost to follow-up.
3- No events occur from a person until the study ends.

Theoretically, censoring can be defined as follows. In a life test we usually got n units to test it and a progressive censoring scheme or censoring plan denoted by $(R_1, \ldots, R_r)$. The units removed from the test within the test period, that process is called progressive censoring in general. There are several models of progressively censored data, but the most popular are the progressive Type-II censoring or the progressive Type-I censoring. The following notations are for progressively censored data.

1- $n, m, R_1, R_2, \dots \in \mathbb{N}_0$ are all integers.

2- The sample size is m; and for some models, it could be random.

3- The total number of units in the experiment is n.

4- The effectively employed removals number at the j^{th} censoring time is R_{j}.

5- $\mathcal{R} = (R_1, \dots, R_r)$ denoted as the censoring scheme, where r is the number of censoring times.

The progressive Type-II censoring is occurred when a surviving unit will be selected randomly to be removed from the experiment when observing a failure to reduce the time and the cost of the experiment.

The set of allowable Type-II censoring schemes can be defined as follows:

$$\mathcal{G}_{m,n}^{m} = \{(r_1, \dots, r_m) \in \mathbb{N}_0^m : \textstyle\sum_{i=1}^{m} r_i = n - m\} \ .$$

Suppose we have k successive zeros, then the notation for that situation will be 0^{*k}. In general $(a_1, 0^{*n_1}, a_2, a_3, 0^{*n_2}, a_4, 0^{*n_3}) = \left(a_1, \underbrace{0, \dots, 0}_{n_1\ times}, a_2, a_3, \underbrace{0, \dots, 0}_{n_2\ times}, a_4, \underbrace{0, \dots, 0}_{n_3\ times}\right)$. For example, (1^{*m}) means that the censoring scheme $(1, \dots, 1) \in \mathcal{G}_{m,n}^{m}$. The following table explains some schemes.

Table 2. Different kinds of censoring schemes.

Scheme $\boldsymbol{\mathcal{R}} = (\boldsymbol{R_1}, \dots, \boldsymbol{R_r})$	**Meaning**
$\sigma_m = (0^{*m-1}, n-m)$	Right censoring, i.e., the sample size is n and the first order statistics are m.
(0^{*m})	Complete sample size $(m = n)$
$\sigma_1 = (n-m, 0^{*m-1})$	First-step censoring plan (FSP), i.e., after failure the exclusion takes place.
$\sigma_k = (0^{*k-1}, n-m, 0^{*m-k})$	One-step censoring plan (OSP), i.e., after the k^{th} failure the removal takes place, $2 \le k \le m-1$

1.4 Literature on Inference Based on Progressively Censored Data

In this section we shall mention related literature reviews to our study. (Balakrishnan, Progressive Censoring Methodology: An Appraisal (with Discussions), 2007) considered the progressively censored order statistics properties and provided the progressively censored samples' procedures. Basically, he focused in his study on many developments related to this topic. Also, he suggested some problems which further research in future can be. Researcher focused on the progressive Type-II right censoring situation, but he presented a brief idea about Type-I right censoring. He explained the basic distribution theory of the progressively Type-II right censoring. Then he talked about the development of that type of censoring. Three main distributions were interested to apply it for progressively Type-II censored order statistics distribution.

(Sarhan & Abuammoh, 2008) derived the inference procedures for the Generalized Exponential distribution based on a progressively Type-II censored data. They applied Monte Carlo simulation to estimate based on point estimation and the interval estimates.

(Ng, Kundu, & Chan, 2009) considered the adaptive Type-II progressive censoring scheme. The maximum likelihood estimation (MLE) has been derived based on the exponential distribution. Also, they constructed the confidence intervals based on diverse methods and applied Monte Carlo simulation to compare their coverage probabilities and expected widths.

(Krishna & Kumar, 2011) obtained the maximum likelihood and Bayesian estimates for one parameter Lindley distribution based on a progressively Type II censored sample. For applications, they used Monte Carlo simulation for calculating interval estimation and coverage probability for their parameter.

(Ye, Chan, Xie, & Ng, 2014) introduced some properties of an adaptive type II progressive censoring. They reduced the bias of the maximum likelihood estimators by using bias correction. After that, they derived the Fisher information matrix for the maximum likelihood estimators of the extreme value distributed lifetimes, considering these properties. To construct confidence intervals for extreme value distribution parameters, they proposed four different approaches. They applied Monte Carlo simulation to compare between these methods. To correct the bias, they used the bootstrap method. The confidence intervals in this study were based on observed information matrix, Fisher information matrix, parametric percentile bootstrap and studentized bootstrap.

In Chapter 2, we derive the likelihood inference about the quantiles of the GE distribution based on a progressively type-II censored data by using invariance property of MLE after considering the likelihood inference of the GE distribution based on a progressively type-II censored data due to (Sarhan & Abuammoh, 2008). Also, we construct an MLE intervals such as asymptotic, percentile and bootstrap-t confidence intervals due to (Baklizi A. , 2008) and (Baklizi A. , 2009). The new method in this thesis is considering these intervals for the quantiles of the GE distribution based on a progressively censored data. In Chapter 3, we derive the Bayesian inference about the quantiles of the GE distribution based on a progressively type-II censored data due to (Gupta & Kundu, Generalized Exponential Distribution: Bayesian Estimations, 2008) and (Krishna & Kumar, 2011). For simulation study, we apply an importance sampling method. Of course, after constructing confidence intervals for both statistical methods (classical and Bayesian), we calculate the coverage probability and expected lengths of these intervals. Chapter 5 presents a comparison between these two methods considering the bias, mean squared error, coverage probability and expected lengths. Our contributions are that bias and MSE results for Bayesian method are closer to zero more than classical methods. Also, Bayesian intervals have the best expected lengths, especially the equal tail intervals.

CHAPTRE 2: LIKELIHOOD INFERENCE

The main idea of the maximum likelihood estimation (MLE) is estimating the parameters of statistical models given observations. (Aldrich, 1997) In 1912, R. A. Fisher derived the “absolute criterion” from the “principle of inverse probability”. The “optimum”, in 1921, was related to the notation of “likelihood” and it was known as a quantity of evaluating hypothetical quantities based on the data given. In 1922, the “maximum likelihood gave estimates which satisfied “sufficiency” and “efficiency”. In that days there were two ways of estimating the likelihood, based on the distribution of the entire sample or sometimes on the distribution of a statistic. Therefore, it could be said that the “Mathematical foundations of theoretical statistics” appeared in 1922 to express the “Maximum likelihood”.

2.1 An Overview of The Likelihood Inference

The maximum likelihood method is based on the likelihood function, which is known as the joint probability distribution or the joint probability density of the random variables $X_1, X_2, \dots, X_n$ at $X_1 = x_1, X_2 = x_2, \dots, X_n = x_n$. The likelihood function is denoted by $L(\underline{x}; \theta) = f(x_1, x_2, \dots, x_n; \theta)$, where $x_1, x_2, \dots, x_n$ are the values of a random sample from a population with parameter θ. Therefore, the maximum likelihood estimator is found by maximizing the likelihood function with respect to θ, and then we call the value of θ that maximizes the likelihood function as the maximum likelihood estimate of θ.

We need to find the first partial derivative of the natural logarithm of the likelihood function with respect to θ to estimate the parameter θ. That first partial derivative is usually called the "score". The Fisher information is given by

$$I(\theta) = E\left[\frac{\partial^2}{\partial\theta^2} ln\, L(\underline{x};\theta)\right]. \qquad (5)$$

Noted that, $0 \leq I(\theta) < \infty$. In this study, we may more interested in using the observed Fisher information matrix, which defined as the negative of the second derivative of the log-likelihood function. Therefore, we can say that the Fisher information $I(\theta)$ is the expected value of the observed Fisher information matrix.

The invariance property of maximum likelihood estimators is a very useful property. Generally, if a specific distribution has a parameter θ, but suppose the interested estimator is for some function of θ, say $\tau(\theta)$. Formally, a theorem below can express the invariance property of MLEs:

Theorem 2.1.1 (Invariance property of MLEs) (Casella & Berger, 2002) . If the MLE of θ is $\hat{\theta}$, then for any function $\tau(\theta)$, the MLE of $\tau(\theta)$ is $\tau(\hat{\theta})$.

Now, we shall define the Asymptotic Normality. Say we have $\hat{\theta}$ is asymptotically normal if

$$\sqrt{n}(\hat{\theta} - \theta_0) \xrightarrow{D} N(0, \sigma^2_{\theta_0}),$$

where the parameter $\sigma^2_{\theta_0}$ is known to be the asymptotic variance of our estimator $\hat{\theta}$, while θ_0 is known as a true value of parameter θ . Suppose that $\hat{\theta}_{MLE}$ converges in probability to θ_0. Then we say that $\hat{\theta}_{MLE}$ is consistent.

Generally, let a statistical model $\{f(.,\theta): \theta \in \Theta\}$ of probability density function (pdf) or probability mass function (pmf) on $X \subseteq R^d$ satisfied and in addition to the consistency assumption. Then we have:

1- The true value of $\theta_0 \in \Theta$.

2- There exists $U \subseteq \Theta$, such that function $\theta \rightarrow f(x, \theta)$, $\forall\, x \in X$ twice continuously differentiable with respect to $\theta \in$ U.

3- A $p \times p$ non-singular Fisher information matrix $I(\theta_0)$ and $E_{\theta_0}\left[\left|\left|\nabla_\theta log f(x, \theta_0)\right|\right|\right] < \infty$ satisfies.

4- A compact ball $K \subseteq U$ of a non-empty interior, exists, which is centered at θ_0, such that $E_{\theta_0} sup_{\theta \in K}\left[\left|\left|\nabla_\theta^2 log f(x, \theta)\right|\right|\right] < \infty$,

$$\int_x sup_{\theta \in K} |\nabla_\theta log f(x, \theta_0)| d\theta < \infty ,$$

$$\int_x sup_{\theta \in K} \left|\nabla_\theta^2 log f(x, \theta)\right| d\theta < \infty ,$$

Related to the assumptions above with their properties, then asymptotic normality of the MLE must hold.

Theorem 2.1.2 (Asymptotic normality of the MLE.) Let $X_1, X_2, X_3, \ldots,$ be identically independent distributed (iid) for $f(x|\theta)$ and $\hat{\theta}$ be the MLE of θ . Then,

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \rightarrow N\left(0, \frac{1}{I(\theta_0)}\right). \qquad (6)$$

Noticed that, when the Fisher information is larger, then the asymptotic variance of the estimator will be smaller.

Assume that we have a sequence of random variables T_n such that

$$\sqrt{n}(T_n - \theta) \xrightarrow{D} N(0, \sigma^2) \text{ as } n \rightarrow \infty ,$$

Let $g(x)$ be a cntinous function such that $g(x)^{`} \neq 0$ then

$$\sqrt{n}(g(T_n) - g(\theta)) \xrightarrow{D} N(0, \sigma^2 \left[g(\grave{\theta})\right]^2).$$

This result is called the delta method. In the multivariate case, the delta method will be defined as in the following theorem.

Theorem 2.1.3 (Multivariate Delta method) (Gugushvili, 2014) Suppose we have a multiparameter vector of differentiated parameters $\tau = g(\theta_1, \dots, \theta_k)$ and it can be differentiated as

$$\nabla g = \begin{pmatrix} \frac{\partial g}{\partial \theta_1} \\ \vdots \\ \frac{\partial g}{\partial \theta_k} \end{pmatrix},$$

and let $\hat{\tau} = g(\hat{\theta})$. Then $\hat{\tau} - \tau \to N\left(0, \left(\hat{\nabla} g\right)^T \widehat{I^{-1}}_n \left(\hat{\nabla} g\right)\right)$,

where $I_n^{-1}(\theta)$ be the inverse of the Fisher information matrix $I_n(\theta)$, with $\widehat{I^{-1}}_n = I_n^{-1}\left(\hat{\theta}\right)$ and $\hat{\nabla} g = \nabla g\left(\hat{\theta}\right)$.

2.2 The Likelihood Inference

Generally, a data is said to be progressively censored when n items entered in a life time testing experiment and observing m failures. While the first failure is observed an R_1 of the surviving units will be selected randomly and removed. Also, when the second failure is observed, the same thing will be done with an R_2 of the surviving units. Therefore, the experiment will be terminated when mth failures will be observed and all remaining surviving units are removed (i.e, $R_m = n - R_1 - R_2 - \cdots - R_{m-1} - m$). The progressively censored sample is denoted as $X_{1:m:n} < X_{2:m:n} < \cdots < X_{m:m:n}$.

In this research our sample will be progressively type II censored data, therefore the joint density function of $\underline{X} = (X_{1,m,n}, \dots, X_{m,m,n})$ with censoring scheme $\underline{R} = (R_1, \dots, R_m)$ is given by:

$$f_{1,2,\dots,m:m:n}(x_1, x_2, \dots, x_m) = d_J \left(\prod_{i=1}^{m} f(x_{i:m:n})\right)\left(\prod_{i=1}^{m} \left(1 - F(x_{i:m:n})\right)^{R_i}\right)$$

$$, \quad 0 < x_{1:m:n} < x_{2:m:n} < \cdots < x_{m:m:n} < \infty , \qquad (7)$$

where, $d_J = \prod_{i=1}^{m}\left[n - i + 1 - \sum_{k=1}^{\max\{i-1,J\}} R_k\right]$.

Now, for the generalized exponential distribution, to find the likelihood function of θ and σ based on the progressively type II censored data we will substitute equations (3) and (4) into equation (7) to get:

$$L(\underline{x};\theta,\sigma) = d_J\left(\prod_{i=1}^{m}\frac{\theta}{\sigma}e^{-x_i/\sigma}\left(1-e^{-x_i/\sigma}\right)^{\theta-1}\right)\left(\prod_{i=1}^{m}\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)^{R_i}\right), \quad (8)$$

Taking *ln* for equation (6) to get:

$$lnL(\underline{x};\theta,\sigma) = const + mln\theta - mln\sigma - \frac{1}{\sigma}\left[\sum_{i=1}^{m}x_i\right] + (\theta-1)\left[\sum_{i=1}^{m}\ln\left(1-e^{-x_i/\sigma}\right)\right]$$

$$+\left[\sum_{i=1}^{m}R_i\ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)\right]. \qquad (9)$$

Now, we need to find the maximum likelihood function estimators for θ and σ. To do that we must take the first derivative of equation (9) firstly with respect to θ and then with respect to σ, and then equating each derivative to zero.

$$\frac{\partial lnL(\underline{x};\theta,\sigma)}{\partial\theta} = \frac{m}{\theta} + \sum_{i=1}^{m}\ln\left(1-e^{-x_i/\sigma}\right) - \sum_{i=1}^{J}R_i\frac{\left(1-e^{-x_i/\sigma}\right)^{\theta}ln\left(1-e^{-x_i/\sigma}\right)}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)}$$

$$= 0 , \qquad (10)$$

$$\frac{\partial lnL(\underline{x};\theta,\sigma)}{\partial\sigma} = \frac{-m}{\sigma} + \frac{1}{\sigma^2}\sum_{i=1}^{m}x_i - \frac{(\theta-1)}{\sigma^2}\sum_{i=1}^{m}\frac{x_i e^{-x_i/\sigma}}{\left(1-e^{-x_i/\sigma}\right)}$$

$$+\frac{\theta}{\sigma^2}\sum_{i=1}^{J}R_i\frac{x_i e^{-x_i/\sigma}\left(1-e^{-x_i/\sigma}\right)^{\theta-1}}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)} = 0 , \quad (11)$$

The MLE of $\hat{\theta}$ and $\hat{\sigma}$ can be found by solving the system simultaneous non-linear equation by Newton-Raphson method.

The p^{th} quantile of GE distribution can be found by finding the inverse function of equation (4) as follows.

$$t_p = F^{-1}(p) = -\sigma \log\left(1 - p^{1/\theta}\right). \qquad (12)$$

In our research we are interested in finding the maximum likelihood estimator for the p^{th} quantile of GE distribution, which has been found by the invariance property of MLE by the following equation:

$$t_p\left(\hat{\theta}, \hat{\sigma}\right) = \hat{x}_p = -\hat{\sigma} \log\left(1 - p^{1/\hat{\theta}}\right). \qquad (13)$$

To find the approximate confidence intervals for x_p for large m, we need to find the observed Fisher Information matrix of the parameters θ and σ, denoted as follows:

$$J(\theta, \sigma) = \begin{bmatrix} -\dfrac{\partial^2 lnL\left(\underline{x}; \theta, \sigma\right)}{\partial \theta^2} & -\dfrac{\partial^2 lnL\left(\underline{x}; \theta, \sigma\right)}{\partial \theta \partial \sigma} \\ -\dfrac{\partial^2 lnL\left(\underline{x}; \theta, \sigma\right)}{\partial \sigma \partial \theta} & -\dfrac{\partial^2 lnL\left(\underline{x}; \theta, \sigma\right)}{\partial \sigma^2} \end{bmatrix}, \qquad (14)$$

$$\frac{\partial^2 lnL(\underline{x};\theta,\sigma)}{\partial \theta^2} = \frac{-m}{\theta^2} - \sum_{i=1}^{J} R_i \left[\frac{2\left(1-e^{-x_i/\sigma}\right)^{\theta} ln\left(1-e^{-x_i/\sigma}\right)}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)} + \frac{2\left(1-e^{-x_i/\sigma}\right)^{2\theta} ln\left(1-e^{-x_i/\sigma}\right)}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)^2} \right],$$

$$\frac{\partial^2 lnL\left(\underline{x}; \theta, \sigma\right)}{\partial \theta \partial \sigma} = -\frac{1}{\sigma^2} \sum_{i=1}^{m} \frac{x_i e^{-x_i/\sigma}}{\left(1 - e^{-x_i/\sigma}\right)} + \frac{\sum_{i=1}^{J} R_i x_i e^{-x_i/\sigma}}{\sigma^2}$$

$$\left[\frac{\left(1-e^{-x_i/\sigma}\right)^{\theta-1}}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)} + \frac{\theta\left(1-e^{-x_i/\sigma}\right)^{\theta} ln\left(1-e^{-x_i/\sigma}\right)}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)} + \frac{\theta\left(1-e^{-x_i/\sigma}\right)^{2\theta-1} ln\left(1-e^{-x_i/\sigma}\right)}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)^2} \right],$$

$$\frac{\partial^2 lnL(\underline{x};\theta,\sigma)}{\partial\sigma^2} = \frac{m}{\sigma^2} - \frac{2}{\sigma^3}\sum_{i=1}^{m} x_i - \frac{(\theta-1)}{\sigma^4}\sum_{i=1}^{m}\frac{x_i^2 e^{-x_i/\sigma}}{(1-e^{-x_i/\sigma})}$$

$$-\frac{(\theta-1)}{\sigma^4}\sum_{i=1}^{m}\frac{x_i^2 e^{-2x_i/\sigma}}{(1-e^{-x_i/\sigma})^2} + \frac{2(\theta-1)}{\sigma^3}\sum_{i=1}^{m}\frac{x_i e^{-x_i/\sigma}}{(1-e^{-x_i/\sigma})}$$

$$-\frac{\theta(\theta-1)}{\sigma^4}\sum_{i=1}^{J} R_i \frac{x_i^2 e^{-2x_i/\sigma}\theta\left(1-e^{-x_i/\sigma}\right)^{\theta-2}}{(1-(1-e^{-x_i/\sigma})^{\theta})}$$

$$+\frac{\theta}{\sigma^4}\sum_{i=1}^{J} R_i \frac{x_i^2 e^{-2x_i/\sigma}\theta\left(1-e^{-x_i/\sigma}\right)^{\theta-1}}{(1-(1-e^{-x_i/\sigma})^{\theta})}$$

$$-\frac{\theta^2}{\sigma^4}\sum_{i=1}^{J} R_i \frac{x_i^2 e^{-2x_i/\sigma}\theta\left(1-e^{-x_i/\sigma}\right)^{2\theta-2}}{(1-(1-e^{-x_i/\sigma})^{\theta})^2}$$

$$-\frac{2\theta}{\sigma^3}\sum_{i=1}^{J} R_i \frac{x_i e^{-x_i/\sigma}\left(1-e^{-x_i/\sigma}\right)^{\theta-1}}{\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)},$$

The inverse of the observed Fisher Information matrix $J^{-1}(\hat{\theta},\hat{\sigma})$ would be the asymptotic variance covariance matrix. Which denoted as follows

$$J_n(\hat{\theta},\hat{\sigma}) = J^{-1}(\theta,\sigma)|_{\theta=\hat{\theta},\sigma=\hat{\sigma}} = \frac{1}{|J(\theta,\sigma)|}\begin{bmatrix} -\frac{\partial^2 lnL(\underline{x};\theta,\sigma)}{\partial\sigma^2} & \frac{\partial^2 lnL(\underline{x};\theta,\sigma)}{\partial\theta\partial\sigma} \\ \frac{\partial^2 lnL(\underline{x};\theta,\sigma)}{\partial\sigma\partial\theta} & -\frac{\partial^2 lnL(\underline{x};\theta,\sigma)}{\partial\theta^2} \end{bmatrix}. \quad (15)$$

Since, equation (15) is a variance covariance matrix then it is symmetric.

Then the asymptotic distribution is as follows

$$\sqrt{n}\begin{pmatrix}\hat{\theta}-\theta \\ \hat{\sigma}-\sigma\end{pmatrix} \xrightarrow{d} N\left(0, J_n(\hat{\theta},\hat{\sigma})\right), \quad (16)$$

Now, by applying multivariate delta method, we have:

$$\widehat{\nabla} t_p = \begin{bmatrix} \frac{\partial t_p(\hat{\theta},\hat{\sigma})}{\partial \hat{\theta}} \\ \frac{\partial t_p(\hat{\theta},\hat{\sigma})}{\partial \hat{\sigma}} \end{bmatrix} = \begin{bmatrix} \frac{\hat{\sigma} p^{1/\hat{\theta}} lnp}{\hat{\theta}^2 (1 - p^{1/\hat{\theta}})} \\ -\log(1 - p^{1/\hat{\theta}}) \end{bmatrix}, \qquad (17)$$

Then,

$$\sqrt{n}\left(-\hat{\sigma}\log(1 - p^{1/\hat{\theta}}) + \sigma \log(1 - p^{1/\theta})\right) \to N\left(0, \widehat{\nabla} t_p{}^T J_n(\hat{\theta},\hat{\sigma}) \widehat{\nabla} t_p\right), \quad (18)$$

where,

$$\widehat{\nabla} t_p{}^T J_n(\hat{\theta},\hat{\sigma}) \widehat{\nabla} t_p =$$

$$\begin{bmatrix} \frac{\hat{\sigma} p^{1/\hat{\theta}} lnp}{\hat{\theta}^2 (1-p^{1/\hat{\theta}})} & -\log(1 - p^{1/\hat{\theta}}) \end{bmatrix} \begin{bmatrix} var(\hat{\sigma}) & cov(\hat{\theta},\hat{\sigma}) \\ cov(\hat{\theta},\hat{\sigma}) & var(\hat{\theta}) \end{bmatrix} \begin{bmatrix} \frac{\hat{\sigma} p^{1/\hat{\theta}} lnp}{\hat{\theta}^2 (1-p^{1/\hat{\theta}})} \\ -\log(1 - p^{1/\hat{\theta}}) \end{bmatrix},$$

$$= var(\hat{\sigma}) \frac{\hat{\sigma}^2 p^{2/\hat{\theta}} (lnp)^2}{\hat{\theta}^4 (1 - p^{1/\hat{\theta}})^2} + 2cov(\hat{\theta},\hat{\sigma}) \log(1 - p^{1/\hat{\theta}}) \frac{\hat{\sigma} p^{1/\hat{\theta}} lnp}{\hat{\theta}^2 (1 - p^{1/\hat{\theta}})}$$

$$+ var(\hat{\theta}) \left(\log(1 - p^{1/\hat{\theta}})\right)^2. \quad (19)$$

We are interested in this study to find the bootstrap estimator of the standard deviation of the maximum likelihood of the p^{th} quantile. Generally, the standard deviation is the square root of the variance. Therefore, the square root of equation (19) will be our bootstrap estimator.

$$\widehat{SD}_{xp}$$

$$= \sqrt{var(\hat{\sigma}) \frac{\hat{\sigma}^2 p^{2/\hat{\theta}} (lnp)^2}{\hat{\theta}^4 (1 - p^{1/\hat{\theta}})^2} + 2cov(\hat{\theta},\hat{\sigma}) \log(1 - p^{1/\hat{\theta}}) \frac{\hat{\sigma} p^{1/\hat{\theta}} lnp}{\hat{\theta}^2 (1 - p^{1/\hat{\theta}})} + var(\hat{\theta}) \left(\log(1 - p^{1/\hat{\theta}})\right)^2}, (20)$$

$\widehat{\widehat{SD}}_{xp} =$

$$\sqrt{var(\hat{\sigma})\frac{\hat{\hat{\sigma}}^2 p^{2/\hat{\hat{\theta}}}(lnp)^2}{\hat{\hat{\theta}}^4\hat{\hat{\theta}}\left(1-p^{1/\hat{\hat{\theta}}}\right)^2} + 2cov\left(\hat{\hat{\theta}},\hat{\sigma}\right)\log\left(1-p^{1/\hat{\hat{\theta}}}\right)\frac{\hat{\hat{\sigma}}p^{1/\hat{\hat{\theta}}}lnp}{\hat{\hat{\theta}}^2\left(1-p^{1/\hat{\hat{\theta}}}\right)} + var(\hat{\hat{\theta}})\left(\log\left(1-p^{1/\hat{\hat{\theta}}}\right)\right)^2}.$$

The a $(1-\alpha)\%$ confidence interval for x_p, which based on the asymptotic results that we got is denoted by

$$\hat{x}_p \pm z_{\alpha/2}\widehat{SD}_{xp} \ . \qquad (21)$$

where $\widehat{SD}_{xp}$ is the asymptotic standard deviation of equation (19) obtained by substituting $\hat{\theta}$ and $\hat{\sigma}$.

2.3 Bootstrap Methods

The bootstrap can be considered as an example of modern science in statistics. In 1969, the idea of bootstrap was first proposed by (Simon, 1969). After that, (Efron, 1979a) inspired by the earlier work on the jackknife to publish the bootstrap in "Bootstrap methods: another look at the jackknife". In 1981, a Bayesian extension had been developed.

Bootstrap methods were applied extensively in the literature, (Li, 2011) estimated the interval for the quantiles of two-parameter exponential distributions. He used two methods, bootstrap and fiducial inferences. In his study, he was interested in calculating the coverage probabilities and expected lengths of both methods. He used numerical simulation study for comparing between these two methods. The results showed that fiducial inference method had well performance under all the examined conditions. He applied the Monte Carlo simulation to get coverage probabilities and expected length estimation. The study ended with that coverage probabilities of fiducial intervals were close to $1-\propto$ and it was larger than the bootstrap intervals for small p. but for large p the coverage probabilities for both methods were close to each other. For that reason, fiducial intervals showed better performance. On the other hand, the expected lengths for bootstrap method showed better results than fiducial method. (Bang & Zhao, 2012) suggested to use censored data to construct confidence intervals by applying some statistical methods such as bootstrap. They did simulation to study the properties of these methods. (Panichkitkosolkul & Saothayanun, 2012) introduced the structure of the bootstrap confidence intervals based on the half-logistic distribution. They applied many types of bootstrap confidence intervals, such as, standard bootstrap, percentile bootstrap and bias-corrected percentile bootstrap confidence intervals. They compared between coverage probabilities and average lengths of bootstrap confidence intervals by using Monte Carlo simulations. The study showed that the coverage probabilities of the standard bootstrap confidence intervals were getting closer to the confidence level than other types of bootstrap confidence intervals. (Baklizi A. , 2008) developed confidence intervals for different quantiles based on one and two independent samples. He considered the maximum likelihood estimator based on record values from the Weibull distribution. He constructed bootstrap-t, bootstrap-t with bootstrap estimated variance and bootstrap percentile intervals and compared between them. This study ended with the length of intervals increased as p^{th} quantile values increased. On the other hand, the larger the sample size the shorter intervals. The error rates were appeared larger for small sample sizes than the nominals. But, for intervals based on the asymptotic normality of the MLE and the observed information matrix or the Fisher information matrix the error rates seems to be moderate. While for all three types of bootstrap intervals, the error rates were very large. (Baklizi A. , 2009) considered the quantiles of the generalized exponential distribution. This study concluded that intervals lengths seems better for higher P values and smaller sample sizes. Intervals' error rates were larger than for nominal ones especially for small sample sizes. While error rates for intervals based on the asymptotic normality of the MLE seemed to be small. Bootstrap intervals, error rates were the largest especially for the percentile interval and the bootstrap-t interval with variance estimates from the Fisher information matrix.

Many journal articles have stressed that bootstrap has great practical value. Also, that journal articles emphasized to consider bootstrapping in applied work. On the other hand, it is not surprising that extremely precise results can be found when combined bootstrapping with modern insights, while traditional methods fail badly. Generally, all bootstrap methods depend on the data has been gotten from a study to detect the sampling distributions, which used to calculate the confidence intervals and test hypotheses. The basic idea of bootstrapping depends on random sampling with replacement. Bootstrapping helps to define sampling distribution of sample means, without considering the normality assumption. It also could assign measures of accuracy (such as variance, bias, confidence intervals, prediction error or some other such measures) to sample estimate. This technique allows by using random sampling method, to estimate the sampling distribution of any statistic. Bootstrap has advantages and disadvantages. The most important advantage is the simplicity of bootstrapping. A standard errors and confidence intervals estimates can be derived easily for complex estimators of complex parameters of the distribution. These complex estimators could be percentile points, odds ratio, proportions, and correlation coefficients. Also, to check the stability and control the results, we can be bootstrapping.

On the other hand, bootstrap confidence intervals are asymptotically more precise than the standard intervals, which depend on the sample variance and the normality assumptions. But the simplicity of bootstrapping hides aside of the disadvantage. Which is the fact that important assumptions need to take care of it when bootstrapping, for example independence of samples. There are many methods for bootstrapping. In this study we are interested in bootstrap methods for means, namely the percentile method and the bootstrap t method.

The bootstrap t method arises when we are interested to compute a confidence interval for μ. Suppose the T statistics which is given by

$$T = \frac{\bar{X}-\mu}{s/\sqrt{n}},$$

and it has a Student's t distribution. Therefore, the confidence interval for the population mean is given by $\left(\bar{X} \pm T\frac{s}{\sqrt{n}}\right)$ when sampling from a Normal distribution. The bootstrap t method or percentile t bootstrap, as it called sometimes, detects the distribution of T. In this situation we get the bootstrap samples same as percentile method, but for this bootstrap sample we must calculate the sample mean and standard deviation, then label them to be $\bar{X}^*$ and s^*.

2.3.1 Bootstrap-t Confidence Interval

To compute the bootstrap-t confidence interval. First, we need to calculate a vector which is given by $Z^* = \frac{(\hat{\hat{x}}_p - \hat{x}_p)}{\widehat{\widehat{SD}}_{xp}}$, where $\widehat{\widehat{SD}}_{xp}$ is the estimated asymptotic standard deviation of $\hat{x}_p$ and it is defined in equation (20). After that, we need to find z_α^* which is the α quantile of the bootstrap distribution Z^*. Then, the bootstrap-t interval is given by

$$\left(\hat{x}_p - z^*_{1-\alpha/2}\widehat{SD}_{xp}, \hat{x}_p - z^*_{\alpha/2}\widehat{SD}_{xp}\right) . \quad (22)$$

2.3.2 Percentile Confidence Interval

The maximum likelihood estimator of x_p has been introduced in chapter 2 to be $\hat{x}_p$ as shown in equation (13). Now, we are interested to find the maximum likelihood estimator for the bootstrap samples which generated from the Generalized Exponential distribution with parameters $\hat{\theta}$ and $\hat{\sigma}$. Therefore, by applying the invariance property on equation (13), then $\hat{\hat{x}}_p$ will be the maximum likelihood estimator for the bootstrap samples. The percentile confidence interval can be constructed by finding the cumulative distribution function of $\hat{\hat{x}}_p$ which would be denoted by $\hat{H}$. The $1-\alpha$ percentile interval for $\hat{x}_p$ is defined by

$$\left(\hat{H}^{-1}\left(\frac{\alpha}{2}\right), \hat{H}^{-1}\left(1-\frac{\alpha}{2}\right)\right) . \quad (23)$$

2.4 Simulation Study

We will investigate the performance of the point and interval estimators. We will consider the bias and MSE for point estimators, the coverage probability and the expected length for confidence intervals.

We did simulation for N=2000. Also, we chose different values of P, which cover the whole range of P $(0 < P < 1)$. Our parameters values fixed to be $\theta = 2$ and $\sigma = 1.2$ for table 5. It is important to mention that we choose B=500, which indicates the bootstrap samples for each scheme. Note that the best choice of B to get better results is extremely large. But in this study, we have been chosen the number of bootstrap samples to be B=500. It is noticed that the choice of B is small as (Davidson & MacKinnon, 2001) did. After that we will take other different values of θ and σ to check the stability of our results.

This study interested in calculating the bias, the mean-squared error (MSE), and the asymptotic variance. Table 5 shows the results of different estimates replicated 2000 times. For that replication we calculated the following for $x_p, p = 0.1, 0.25, 0.5, 0.75, 0.9$;

1- Bias: The expected value of the difference between the estimator's $(\hat{x}_p)$ value and the true value parameter (x_p) is called the bias function of an estimator and it is denoted by:

$$bias = E(\hat{x}_p - x_p).$$

2- MSE: Or it called the risk function. It is the expected value of the squared of the difference between the estimator and the true value parameter. MSE measures the quality of an estimator, whenever it is closer to zero, the better. MSE values are always non-negative and it is denoted by:

$$MSE = E(\hat{x}_p - x_p)^2.$$

After we calculated our parameters, we used it to calculate our confidence intervals. First, we substituted the five different values that we got for $\hat{x}_p$ into equation (21) to find the asymptotic intervals for all values of P's. We sorted the five vectors of $\hat{\hat{x}}_p$'s before calculating the percentile confidence intervals which noticed in equation (23). To find out the Bootstrap-t confidence interval we calculated first the vector Z^*, which is explained in section 2.3.1, before calculating equation (22) we sorted Z^* vectors. For both percentile and Bootstrap-t confidence intervals we used "quantile" equation in R software. In this research we are interested in finding the average length and the error rates for each interval we have got. The average length for any confidence interval can be calculated by subtracting the lower bound of the confidence interval from the upper bound. To calculate the error rate for each interval type, we counted how many times that the values of parameter (x_p) can be higher than the upper bound or less than the lower bound and then we take the proportion of them to get the error rate. Shortly, we called the confidence intervals in tables 6 and 7 as follows:

1- A I: Asymptotic confidence interval.
2- P R C: Percentile confidence interval.
3- Boot-t: Bootstrap-t confidence interval.

To evaluate the performance of the p^{th}quantile estimator $\hat{x}_p$ and the bootstrap estimator $\hat{\hat{x}}_p$, we did a simulation study by using R software. To do so, following Mohi El-Din et al. (2016) we chose different censoring schemes with different sample sizes n and different choices of m, where n and m are the total number of units and the sample size, respectively. Table 3 below shows the censoring schemes used in the simulation study.

Table 3. Censoring schemes.

Scheme	**N**	**M**	**R**
1	50	30	$(0^{11}, 19, 0^9, 1, 0^8)$
2	50	40	$(0^{15}, 1,3, 0^2, 1, 0^{10}, 1, 0^2, 2, 0^4, 1,1)$
3	70	50	$(0^{20}, 5,5,4,3, 0^6, 1,1, 0^{10}, 1, 0^7)$
4	90	60	$(0^{15}, 7,3,0,4,2, 0^{10}, 2,3,5,1, 0^{13}, 1,1,1, 0^{10})$
5	100	70	$(0^{35}, 13,1, 0^{10}, 3,0,3, 0^7, 2,2,0,5, 0^8, 1)$
6	100	80	$(0^{42}, 6,1,4,1, 0^{14}, 1,2, 0^{10}, 2,2,1, 0^5)$
7	130	100	$(0^{60}, 1,2,4,0,2, 0^{10}, 2,1,0,5,7,0,3,0,2,1, 0^{15})$

For generating a progressively type II censored data we used simple simulations steps which had been presented by (Balakrishnan & Sandhu, A Simple Simulational Algorithm for Generating Progressive Type II Censored Samples, 1995) . The following simulation algorithm steps explain the way of generating a progressively censored type II data:

1- Generate m independent observations, such that $m \sim Uniform(0,1)$. These observations are called $W_1, W_2, \dots, W_m$.

2- Calculate $V_i = W_i^{1/(i+R_m+R_{m-1}+\cdots+R_{m-i+1})}$, $\forall\ i = 1, 2, \dots, m$.

3- Compute $U_i = 1 - V_m V_{m-1} \dots V_{m-i+1}, \forall\ i = 1, 2, \dots, m$. Noticed that, $U_1, U_2, \dots, U_m$ are required for progressive type II censored sample from Uniform (0,1) distribution.

4- Finally, set $X_i = F^{-1}(U_i)$, $\forall\ i = 1, 2, \dots, m$. Where the inverse cdf of the distribution under consideration is known as $F^{-1}(.)$. Then the required progressive type II censored sample from the distribution $F(.)$ is $X_1, X_2, \dots, X_m$.

Note that, the simulation above needs exactly m uniform observations and doesn't need any sorting.

Before we started our simulation in R software, we downloaded some specific packages in R to make sure that our simulation done perfectly. We will mention some of these packages such as "optimization" and "optimx" to apply the optim function. At the end of the simulation we transferred our results tables to a word document, to do so we downloaded "rtf" and "Rcpp" packages.

In R software, to find the maximum likelihood estimators we applied a function called optim in R software, but the optim function couldn't find the MLE values directly, so we wrote a command at the end of the function "return(-log_L)", which multiply equation (9) by mines to get our results. Of course, after we got the values of $\hat{\theta}$ and $\hat{\sigma}$, we substituted them into equation (13) to get the values of $\hat{x}_p$'s. Also, to find the Fisher Information matrix, which is defined in equation (5), we included in the optim function a command called "hessian = TRUE". After that we found the inverse of the observed fisher information matrix, noted in equation (15). Then we substituted it in equation (19), which has been calculated directly in R software. Then to find the bootstrap estimator $\hat{\hat{x}}_p$, we repeated the optim function using the values of $\hat{\theta}$ and $\hat{\sigma}$ that we got for B=500 repeating this step for N=2000. So, we got the values of $\hat{\hat{\theta}}$ and $\hat{\hat{\sigma}}$, and similarly we repeated the steps above to get the values of $\hat{\hat{x}}_p$'s and to find the inverse of the observed fisher information matrix. After we calculated $\hat{x}_p$ and $\hat{\hat{x}}_p$ values, we used it to find what we interested in (Bias and MSE) as explained above.

We faced a problem in finding a suitable initial guess. For that reason, we applied an optim function in R software two times. First, we used $\theta = 2$ and $\sigma = 1.2$ as an initial guess for the first optim function. Also, we applied the Taylor expansion in equation (9). We expanded the following term in equation (9) to be as follows:

$$\sum_{i=1}^{m} R_i \ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)$$

$$\approx -\sum_{i=1}^{m} R_i\left(1-e^{-x_i/\sigma}\right)^{\theta} - \frac{1}{2}\sum_{i=1}^{m} R_i\left(1-e^{-x_i/\sigma}\right)^{2\theta}$$

$$-\frac{1}{3}\sum_{i=1}^{m} R_i\left(1-e^{-x_i/\sigma}\right)^{3\theta} - \frac{1}{4}\sum_{i=1}^{m} R_i\left(1-e^{-x_i/\sigma}\right)^{4\theta}\ . \quad (24)$$

Note that we used the first fourth terms of the expansion as an approximation because Taylor expansion likelihood is easy to maximize. After that we used its solution parameters for θ and σ as an initial guess for our original likelihood function. Note that, second time we applied the optim function, a Taylor expansion hasn't been used. We directly used equation (9) for optimizing our parameters.

Table 4. Bias and MSE results for classical statistics methods.

Scheme		**P=0.1**	**P=0.25**	**P=0.5**	**P=0.75**	**P=0.9**
1	Bias	0.019	0.011	-0.006	-0.031	-0.062
	MSE	0.01	0.016	0.037	0.118	0.315
2	Bias	0.017	0.01	-0.004	-0.025	-0.052
	MSE	0.01	0.014	0.028	0.083	0.227
3	Bias	0.012	0.007	-0.002	-0.016	-0.033
	MSE	0.007	0.011	0.023	0.069	0.182
4	Bias	0.011	0.009	0.003	-0.006	-0.018
	MSE	0.006	0.008	0.018	0.057	0.154
5	Bias	0.008	0.004	-0.002	-0.012	-0.022
	MSE	0.004	0.007	0.017	0.059	0.165
6	Bias	0.012	0.01	0.004	-0.006	-0.018
	MSE	0.005	0.007	0.014	0.039	0.105
7	Bias	0.007	0.005	0.001	-0.003	-0.008
	MSE	0.004	0.006	0.013	0.039	0.11

In General, the results in table 4 shows that results seem to be slightly similar from one experiment or another, but it is important to note that bias and the MSE, values are lower when we choose larger values of sample sizes m and n defined in table 3.

On the other hand, it is very clear that values of bias in each scheme is decreasing when increasing the p^{th} quantile values. In contrast, mean square error are increased when increasing the p^{th} quantile values.

For confidence intervals, we are interested in calculating the interval length and the error rate for each interval. Tables 5 and 6 show the results for intervals lengths and error rates respectively for 2000 replications. On the other hand, confidence intervals are calculated for both $\alpha = 0.1$ and $\alpha = 0.05$ respectively, in tables 5 and 6.

The lengths of all types of intervals can be found by the difference between the upper bound and the lower bound of the intervals. The error rates can be calculated by checking whether the estimator x_p belongs to the confidence intervals or not.

Table 5. Coverage probability and expected lengths results for classical statistics methods when $\alpha = 0.1$

$\boldsymbol{\alpha = 0.1}$ & $\boldsymbol{N = 2000}$											
Scheme	**Interval Type**	**P=0.1**		**P=0.25**		**P=0.5**		**P=0.75**		**P=0.9**	
		A L	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**
1	A I	0.312	0.114	0.395	0.115	0.612	0.12	1.091	0.131	1.788	0.135
	P R C	0.315	0.129	0.395	0.114	0.605	0.126	1.071	0.144	1.748	0.15
	Boot-t	0.326	0.093	0.413	0.101	0.649	0.112	1.176	0.11	1.952	0.116
2	A I	0.31	0.119	0.381	0.118	0.543	0.107	0.926	0.117	1.509	0.125
	P R C	0.327	0.155	0.39	0.126	0.524	0.118	0.84	0.161	1.331	0.192
	Boot-t	0.338	0.088	0.417	0.079	0.59	0.096	1.012	0.134	1.674	0.15
3	A I	0.263	0.121	0.328	0.121	0.485	0.118	0.844	0.119	1.382	0.124
	P R C	0.265	0.121	0.328	0.121	0.481	0.119	0.834	0.125	1.361	0.131
	Boot-t	0.271	0.102	0.337	0.107	0.5	0.109	0.881	0.112	1.452	0.109
4	A I	0.234	0.111	0.294	0.098	0.442	0.099	0.777	0.106	1.273	0.111
	P R C	0.235	0.119	0.293	0.104	0.438	0.099	0.769	0.111	1.259	0.115
	Boot-t	0.239	0.1	0.3	0.094	0.453	0.099	0.805	0.104	1.328	0.104
5	A I	0.222	0.106	0.275	0.108	0.407	0.107	0.719	0.111	1.186	0.117
	P R C	0.229	0.124	0.276	0.113	0.396	0.112	0.678	0.13	1.107	0.145
	Boot-t	0.233	0.082	0.287	0.086	0.42	0.104	0.745	0.116	1.241	0.119
6	A I	0.222	0.106	0.273	0.105	0.389	0.108	0.663	0.102	1.082	0.103
	P R C	0.227	0.122	0.275	0.118	0.383	0.111	0.643	0.108	1.044	0.111
	Boot-t	0.23	0.081	0.2809	0.096	0.396	0.099	0.679	0.099	1.116	0.1
7	A I	0.203	0.109	0.25	0.107	0.363	0.111	0.631	0.116	1.037	0.123
	P R C	0.21	0.129	0.254	0.11	0.352	0.116	0.586	0.14	0.946	0.157
	Boot-t	0.214	0.086	0.263	0.086	0.375	0.105	0.649	0.138	1.073	0.149

Table 6. Coverage probability and expected lengths results for classical statistics methods when $\alpha = 0.05$

		$\alpha = 0.05$ & $N = 2000$									
Scheme	**Interval**	**P=0.1**		**P=0.25**		**P=0.5**		**P=0.75**		**P=0.9**	
	Type	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**
1	A I	0.373	0.061	0.474	0.058	0.734	0.069	1.308	0.079	2.144	0.084
	P R C	0.376	0.081	0.472	0.069	0.723	0.069	1.28	0.083	2.096	0.086
	Boot-t	0.393	0.045	0.5	0.049	0.785	0.058	1.426	0.056	2.369	0.062
2	A I	0.369	0.057	0.455	0.06	0.649	0.063	1.105	0.08	1.803	0.084
	P R C	0.388	0.092	0.464	0.072	0.625	0.069	1.001	0.103	1.586	0.124
	Boot-t	0.404	0.04	0.501	0.041	0.709	0.051	1.218	0.077	2.014	0.094
3	A I	0.314	0.065	0.392	0.061	0.579	0.062	1.009	0.067	1.652	0.069
	P R C	0.316	0.083	0.391	0.072	0.573	0.066	0.995	0.065	1.626	0.068
	Boot-t	0.325	0.052	0.404	0.057	0.601	0.06	1.059	0.052	1.75	0.044
4	A I	0.279	0.061	0.351	0.063	0.527	0.063	0.927	0.066	1.519	0.066
	P R C	0.28	0.066	0.348	0.065	0.52	0.065	0.912	0.068	1.496	0.066
	Boot-t	0.286	0.044	0.358	0.055	0.542	0.055	0.962	0.058	1.59	0.057
5	A I	0.263	0.061	0.325	0.068	0.481	0.057	0.849	0.054	1.401	0.056
	P R C	0.271	0.084	0.328	0.075	0.467	0.06	0.801	0.067	1.306	0.072
	Boot-t	0.278	0.046	0.34	0.05	0.498	0.056	0.885	0.058	1.476	0.061
6	A I	0.264	0.068	0.326	0.062	0.465	0.06	0.794	0.056	1.297	0.057
	P R C	0.27	0.076	0.327	0.067	0.457	0.064	0.767	0.062	1.246	0.064
	Boot-t	0.274	0.045	0.336	0.055	0.475	0.056	0.815	0.062	1.342	0.062
7	A I	0.241	0.064	0.298	0.064	0.432	0.058	0.751	0.06	1.233	0.061
	P R C	0.25	0.082	0.302	0.068	0.418	0.065	0.696	0.081	1.125	0.09
	Boot-t	0.255	0.048	0.314	0.043	0.447	0.058	0.775	0.078	1.283	0.081

Before commenting on tables 5 and 6, we shall describe the coverage probabilities and indicate whether it reach the nominal coverage probability or not? Where the nominal error for $\alpha = 0.1$ and $\alpha = 0.05$ are between 0.08 and 0.12, and between 0.04 and 0.06, respectively. For $\alpha = 0.1$, schemes 1, 2, 3, and 7 don't attain the nominal coverage probability for some confidence intervals, especially when $p = 0.75, 0.9$. On the other hand, some confidence intervals in table 6 show more problems about attaining the coverage probability, which is clear in all schemes, except scheme 4, and for all p^{th} quantiles, except for $p = 0.5$.

From tables 5 and 6, we noted that the length of the three types of intervals are getting smaller while taking larger samples (m and n are larger which are defined in table 3).

It is very clear that the bootstrap-t interval's lengths are larger than the other intervals, but the smallest one is the percentile interval of the other intervals especially when $p = 0.5, 0.75, 0.9$. That result is more pronounced when m and n , which are defined in table 3, are larger and in addition when $p = 0.5, 0.75, 0.9$. Also, the average lengths seem to be smaller when $\alpha = 0.1$.

Now, the results for the error rates for each type of intervals. Generally, we can conclude that percentile confidence interval shows more problems in attaining the coverage probability all over the schemes and for all p^{th} quantiles. On the other hand, bootstrap-t interval is more likely to attain the coverage probability for all schemes especially when $p = 0.1$ & 0.25.

It is interesting to note that error rates for the three types of intervals are similar from scheme to another and get closer to the nominal probabilities when m and n are larger.

To clarify our results more, we have chosen only four schemes results to plot it by using R software again. Figure 1 presents the plot of the bias and MSE results for schemes 1, 2, 3, and 6. While figures 2 and 4 present the plots of the expected lengths of confidence intervals for the same schemes. Finally, figures 3 and 5 present the plots of these schemes' coverage probability. It is noted that values of p are plotted on the x-access and all the results plotted on the y-access.

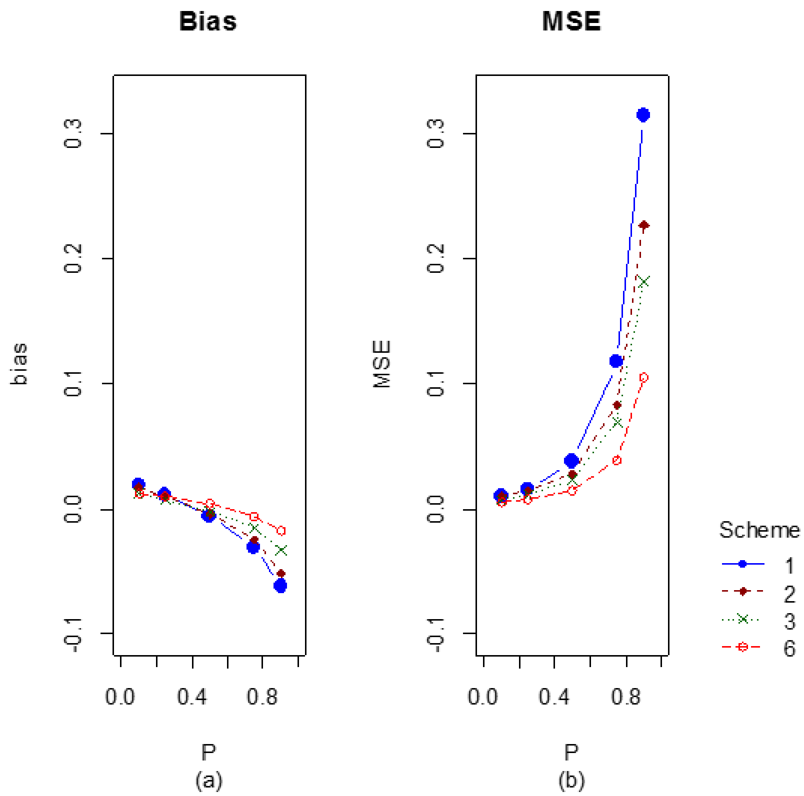

Figure 2. Bias and MSE plots for classical statistics methods.

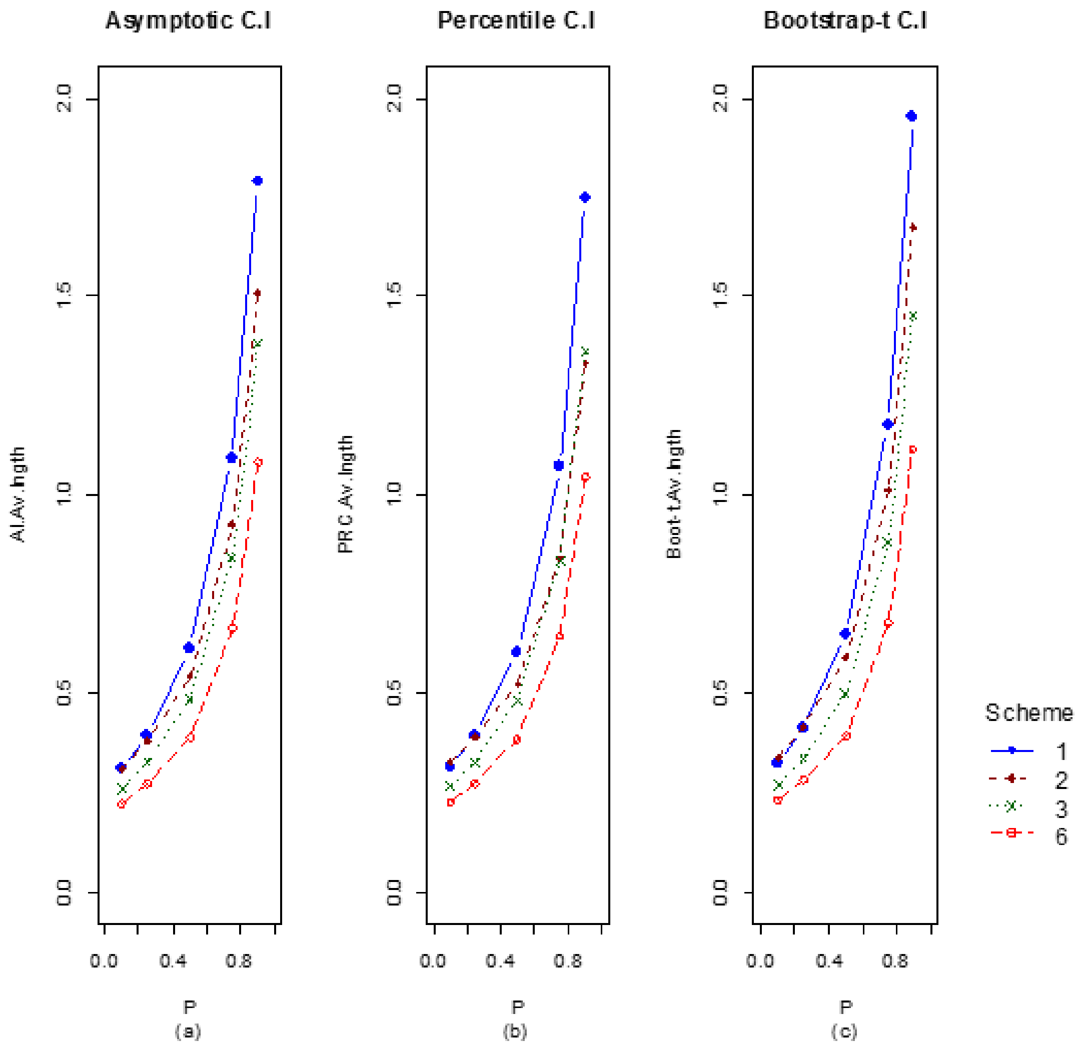

Figure 3. Expected lengths plots for classical statistics methods when $\alpha = 0.1$

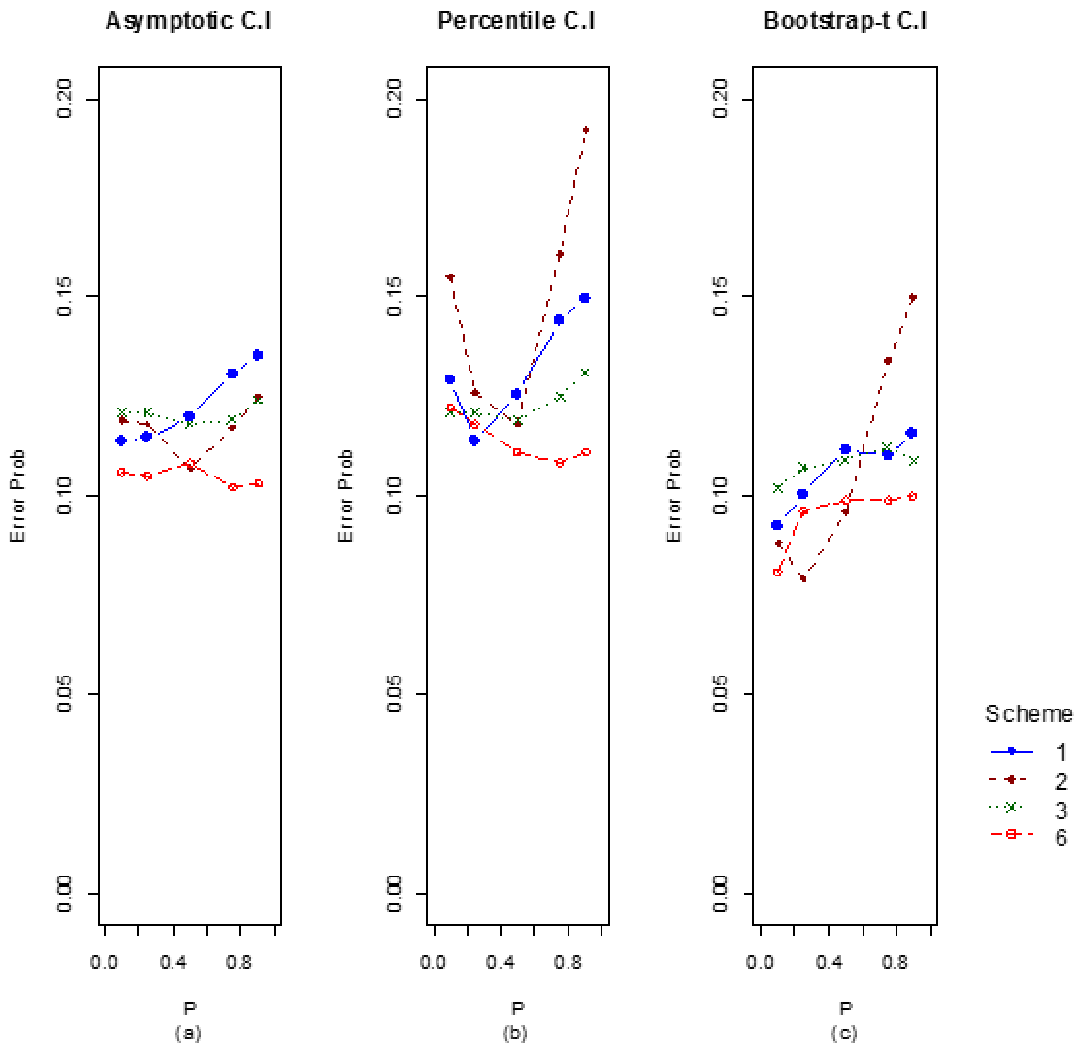

Figure 4. Coverage probability plots for classical statistics methods when $\alpha = 0.1$

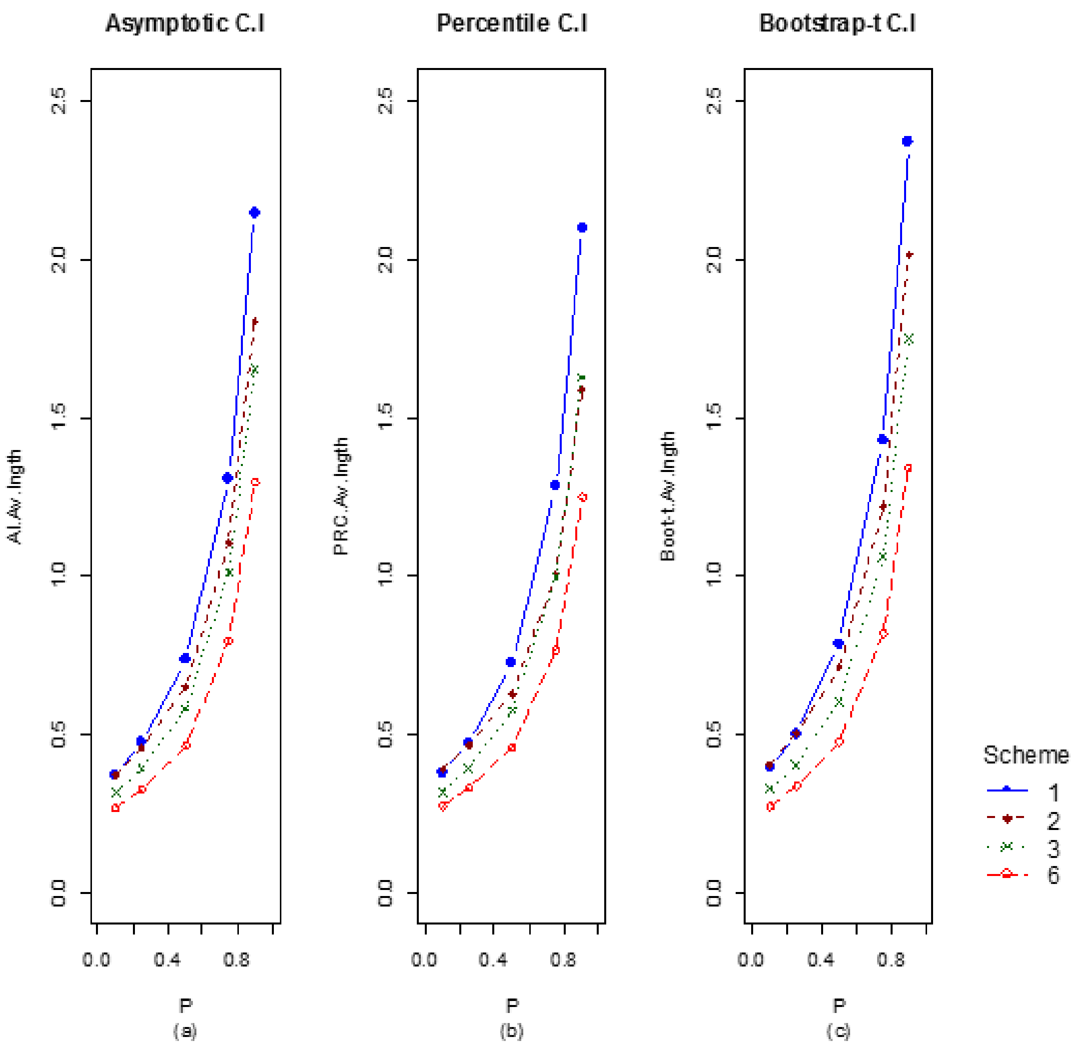

Figure 5. Expected lengths plots when for classical statistics methods $\alpha = 0.05$

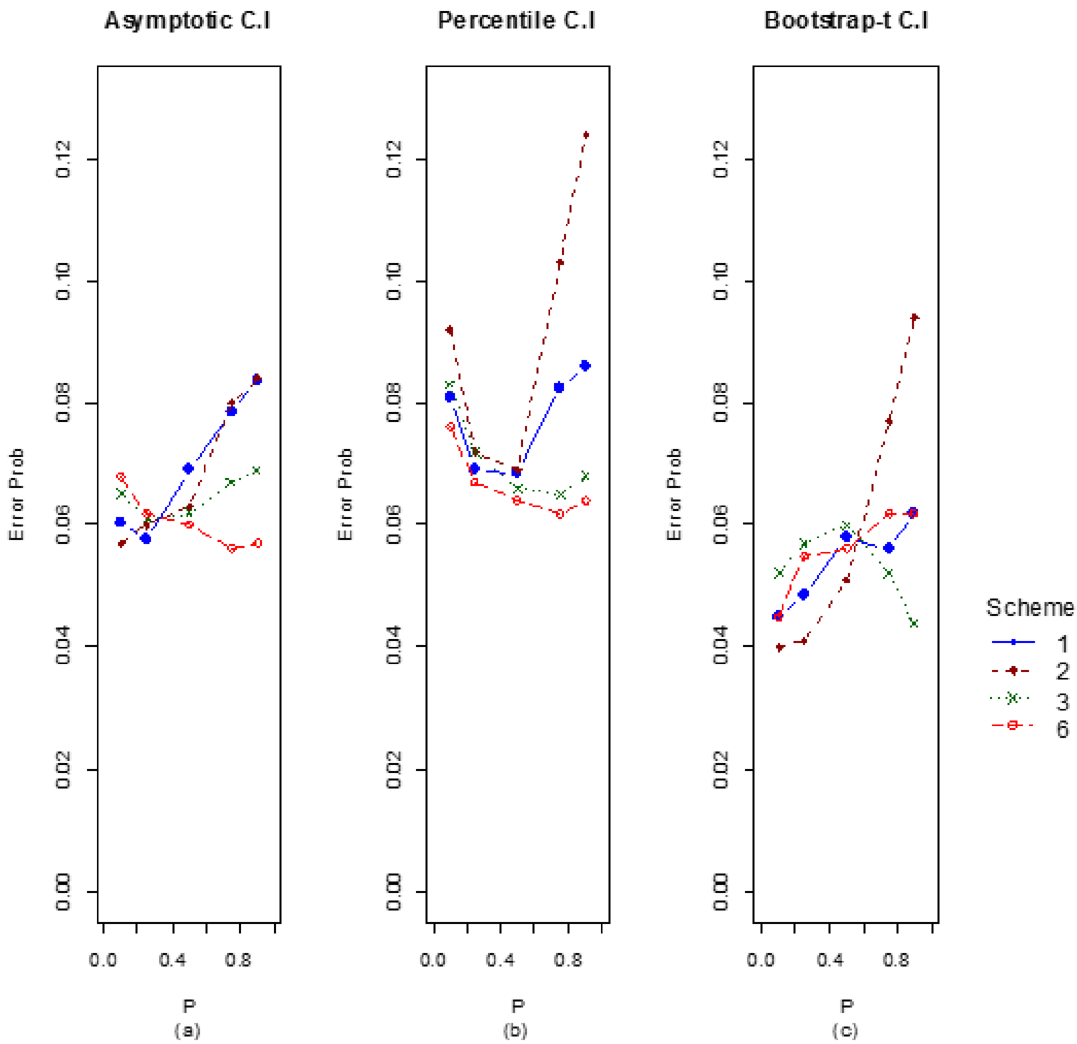

Figure 6. Coverage probability plots for classical statistics methods when $\alpha = 0.05$

Also, we applied our simulation again, but with different values of θ and σ. This process has been made to check the stability of our results and to compare between them. Therefore, we apply it for only three schemes (1, 2, and 3) for our purposes. We take another different value of θ and σ as shown in tables 7 and 8. Also, we concentrate on only one value of p^{th} quantile, i.e; $p = 0.5$.

The table below displays the results of bias and MSE of our parameter (x_p).

Table 7. Bias and MSE results of schemes 1, 2, 3 for different values of θ and σ.

P=0.5				
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Bias**	**MSE**
1	$\theta = 1.2$	$\sigma = 0.5$	0.001	0.005
2			-0.001	0.004
3			0	0.003
1	$\theta = 1.5$	$\sigma = 0.7$	-0.004	0.011
2			0.002	0.009
3			0.0004	0.007
1	$\theta = 1.7$	$\sigma = 0.9$	0.001	0.021
2			-0.009	0.014
3			0.003	0.011

Now, table 8 and table 9 present the results for the average length and error rates for all three types of intervals that we have. Of course, these tables are presenting the all cases of different values of θ and σ and for both values of $\alpha = 0.1$ & $\alpha = 0.05$.

Table 8. Coverage probability and expected lengths results for different values of θ and σ for $\alpha = 0.1$ and $p = 0.5$.

$\boldsymbol{\alpha = 0.1}$ & $\boldsymbol{p = 0.5}$					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Interval Type**	**A L**	**E R**
1	$\theta = 1.2$	$\sigma = 0.5$	A I	0.222	0.124
			P R C	0.219	0.122
			Boot-t	0.238	0.106
2			A I	0.197	0.118
			P R C	0.195	0.124
			Boot-t	0.210	0.103
3			A I	0.175	0.120
			P R C	0.173	0.118
			Boot-t	0.181	0.109
1	$\theta = 1.5$	$\sigma = 0.7$	A I	0.333	0.12
			P R C	0.329	0.124
			Boot-t	0.355	0.106
2			A I	0.297	0.115
			P R C	0.291	0.115
			Boot-t	0.318	0.105
3			A I	0.264	0.103
			P R C	0.263	0.104
			Boot-t	0.274	0.102
1	$\theta = 1.7$	$\sigma = 0.9$	A I	0.445	0.128
			P R C	0.440	0.128
			Boot-t	0.472	0.125
2			A I	0.389	0.120
			P R C	0.378	0.118
			Boot-t	0.418	0.105
3			A I	0.352	0.096
			P R C	0.349	0.104
			Boot-t	0.363	0.092

Table 9. Coverage probability and expected lengths results for different values of θ and σ for $\alpha = 0.05$ and $p = 0.5$.

$\boldsymbol{\alpha = 0.05}$ & $\boldsymbol{p = 0.5}$					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Interval Type**	**A L**	**E R**
1	$\theta = 1.2$	$\sigma = 0.5$	A I	0.264	0.072
			P R C	0.261	0.071
			Boot-t	0.287	0.05
2			A I	0.235	0.072
			P R C	0.232	0.069
			Boot-t	0.254	0.048
3			A I	0.208	0.069
			P R C	0.206	0.063
			Boot-t	0.218	0.056
1	$\theta = 1.5$	$\sigma = 0.7$	A I	0.397	0.068
			P R C	0.391	0.072
			Boot-t	0.428	0.053
2			A I	0.354	0.066
			P R C	0.346	0.067
			Boot-t	0.382	0.054
3			A I	0.315	0.06
			P R C	0.312	0.059
			Boot-t	0.328	0.051
1	$\theta = 1.7$	$\sigma = 0.9$	A I	0.530	0.078
			P R C	0.523	0.078
			Boot-t	0.569	0.062
2			A I	0.463	0.072
			P R C	0.449	0.070
			Boot-t	0.501	0.053
3			A I	0.419	0.053
			P R C	0.415	0.056
			Boot-t	0.436	0.043

From table 7, bias results for all different values of θ and σ are getting smaller, especially for $\theta = 1.2$ and $\sigma = 0.5$. Similarly, for the MSE results. It is worth to mention that these results' behavior is the same as results behavior when $\theta = 2$ and $\sigma = 1.2$.

Comparing between the results for different values of θ and σ when $p = 0.5$, bias and MSE results are the smallest when $\theta = 1.2$ and $\sigma = 0.5$. While for $\theta = 2$ and $\sigma = 1.2$ results are the largest. Therefore, we can say that the smaller values of θ and σ the better bias and MSE results got.

As we discussed before, that table 8 presents the results of average lengths and error rates for different values of θ and σ when $\alpha = 0.1$ and $p = 0.5$. In general, attaining the coverage probability for some types of confidence intervals, is behaving the same as when $\theta = 2$ and $\sigma = 1.2$. Also, it is very clear that all coverage probabilities for all schemes for different values of θ and σ in table 8 attain the nominal error rate, therefore the average lengths for all these types of confidence intervals are comparable. Then we can say that the smallest average lengths are for $\theta = 1.2$ and $\sigma = 0.5$ and the best average lengths is for the percentile confidence interval. It is noted that for each value of θ and σ , the expected lengths of confidence intervals are getting smaller when m and n are larger.

From table 9, attaining the nominal error rates for some confidence intervals for different values of θ and σ show some problems, especially when $\theta = 1.2$ and $\sigma = 0.5$ and $\theta = 1.7$ and $\sigma = 0.9$. It is noted that, this problem appears in asymptotic and percentile confidence intervals for schemes 1 and 2. Therefore, we can only compare between scheme 3 average lengths of confidence intervals. Similarly, the comparable average lengths in table 9, behave the same as average lengths presented in table 8.

Finally, we can say that the smaller values of θ and σ, the better average lengths that we get, and this is true for $\alpha = 0.1$ and $\alpha = 0.05$.

CHAPTER 3: BAYESIAN INFERENCE

Bayesian statistics is an important field in statistics. This field depends on the Bayesian interpretation of probability. Generally, a degree of belief of an event can be expressed by probability. An event may have a prior knowledge, then we can say that the degree of belief may base on that prior knowledge. Where these prior knowledges could be the results of previous experiments or personal beliefs about the event. To describe how Bayesian statistics works, at the beginning of any problem, we shall start with some probabilities, which called prior probabilities Bayesian statistics to get more information or updated probabilities. These updated probabilities are called posterior probabilities. Bayesian statistics depends fundamentally on Bayes theorem. The main idea of Bayes theorem is describing the conditional probability of an event based on an observed data or a prior information or beliefs about the event.

(Gelman, et al., 2013) and (Fienberg, 2006) gave a brief introduction about Bayesian statistics. In 18^{th} century, the Bayes theorem had been introduced firstly by a mathematician and theologian Thomas Bayes. And he published his paper in 1763 which described the formulation of a specific case of Bayes theorem. Between the end of the 18^{th} century and the 19^{th} century, many research papers were published on Bayes theory. Pierre-Simon Laplace was the first one developed it to a modern formulation in his "Théorie analytique des probabilités." Laplace developed the Bayesian interpretation of probabilities. The Bayesian methods used by Laplace are still used to solve many statistical problems. Also, other later authors developed many Bayesian methods. But this method wasn't commonly used until the 1950s. Bayesian methods weren't favored during the 20^{th} century because of their philosophy. Also, Bayesian methods need a lot of computing a programing system to complete them, which weren't available for much of the 20th century. While most of the methods used in that period were for frequentist interpretations. After developing computers and powerful computers showed up with new algorithms such as the Markov chain Monte Carlo, therefore, Bayesian methods has not been widely used by statistician until the 21st century. Bayesian statistics currently is an important active branch of statistics.

3.1 An Overview on Bayesian Inference

The posterior distribution is considered as the most important quantity in Bayesian inference. All the information about the unknown parameter θ are available in the posterior distribution after getting an observed data $X = x$. The general definition can be defined as the following. Suppose we have an observed data $X = x$ of a random variable X with density function $f(x|\theta)$ and the prior distribution has a density function $f(\theta)$. Then the posterior distribution can be defined based on Bayes' theorem as follows

$$f(\theta|x) = \frac{f(x|\theta)f(\theta)}{\int f(x|\theta)f(\theta)d\theta} \quad , \qquad (26)$$

where $f(x|\theta)$ is simply can be known as the likelihood function $L(\theta)$. We write $L(\theta) = f(x|\theta)$ because θ is random, so we explicit condition on a specific value θ.

Now, based on Bayes theorem, the density of the posterior distribution is proportional to the numerator of equation (26), i.e:

$$f(\theta|x) \propto f(x|\theta)f(\theta) \qquad or \qquad f(\theta|x) \propto L(\theta)f(\theta) \ , \qquad (27)$$

where " $\propto$ " is known as "is proportional to". Therefore, generally the density function of the posterior distribution can be computed by multiplying the likelihood function and the prior density function.

In Bayesian statistics, the prior and posterior distribution are said to be conjugate prior distributions, if they belong to the same family distribution and the prior distribution is called a conjugate prior for the likelihood function. For more explanation, suppose we have a Gaussian likelihood function and choose a Gaussian prior distribution. Then the posterior distribution is also Gaussian. Therefore, the Gaussian family is said to be conjugate to itself or called self-conjugate. More generally, assume that the likelihood function ($L(\theta) = f(x|\theta)$) which based on the observation $X = x$. A class $\mathcal{G}$ of distributions is called conjugate with respect to $L(\theta)$ if the posterior distribution $f(\theta|x)$ is in $\mathcal{G}$ for all x whenever the prior distribution $f(\theta)$ is in $\mathcal{G}$ (Held & Sabanés Bové, 2014).

In Bayesian statistics to estimate the unknown parameter θ, at least three possible Bayesian point estimates are offered, such as the posterior mean, mode, and median. The question is which one should we choose for our application? To answer this question, we should define a loss function which indicated to be a theoretic way to take a decision. A loss function $l(a, \theta) \in \mathbb{R}$ computes the loss encountered when estimating the true parameter θ by a. For more explanation, suppose $a = \theta$, then the related loss function is set to zero: $l(a, \theta) = 0$. The common used loss function is the quadratic loss function $l(a, \theta) = (a - \theta)^2$. Another choice for the loss function, is the linear loss function $l(a, \theta) = |a - \theta|$ or the zero-one loss function

$$l_\varepsilon(a, \theta) = \begin{cases} 0, & |a - \theta| \leq \varepsilon\,, \\ 1, & |a - \theta| > \varepsilon, \end{cases}$$

In this situation we need to choose a suitable added parameter $\varepsilon > 0$. Now, to indicate the point estimate a to minimize the posteriori expected loss with respect to $f(\theta|x)$. That point estimate is called a Bayes estimate. Formally, a Bayes estimate of θ with respect to a loss function $l(a,\theta)$ minimizes the expected loss with respect to the posterior distribution $f(\theta|x)$. It minimizes

$$E\{l(a,\theta)|x\} = \int_{\Theta} l(a,\theta)f(\theta|x)d\theta \,. \quad (28)$$

Now, let us introduce the credible region's definition. A subset $C \subseteq \Theta$ with $\int_C f(\theta|x)d\theta = \gamma$ is called γ.100% credible region for θ with respect to $f(\theta|x)$. If C is a real interval.

As any confidence interval, the credible interval can be defined as the range of values within an unobserved parameter value which falls with a subjective probability. This interval is in the domain of a posterior probability distribution or a predictive distribution. The credible interval is like the confidence intervals in frequentist statistics, despite the differences in their respective philosophies. Therefore, it is interesting to mention the differences between credible intervals and confidence intervals. In Bayesian statistics, intervals' bounds are treated as fixed and the estimated parameter as a random variable, in contrast frequentist confidence intervals treat their bounds as a random variable and the estimated parameter as a fixed value. In addition to that, Bayesian credible intervals need knowledge of the condition-exact prior distribution, but frequentist confidence intervals don't require that.

To find the suitable credible interval, there are some methods to follow, such as:

1- The highest posterior density interval need to choose the narrowest interval which can be done for the unimodal distribution. In this case, the mode will be chosen with those values of highest probability density.
2- The equal-tailed interval, in this case the interval can be chosen where the probability of being lower the interval is the same as being above it. This kind of interval will contain the median.

To calculate these intervals, we need the posterior distribution, however, in many cases, the posterior distribution is known only up to a proportionality constant. Therefore, we can't use the posterior directly and we need some solution to this problem. Various approaches were used in the literature including importance sampling and Markov chain Monte Carlo (MCMC) techniques. In this thesis, we will use importance sampling as (Shaw, 2018) clarified. Suppose we have any distribution $f(x)$ and we can't sample from it. But, it is possible to generate samples from another distribution $q(x)$ which approximates $f(x)$. Therefore, we can use importance sampling method to sample from $q(x)$. Suppose that we face a problem in estimating $E\{g(X)\}$ for some function $g(x)$ with respect to a density $f(x)$.

Then that satisfy the following

$$E\{g(X)\} = \int_{x} \frac{g(x)f(x)}{q(x)} q(x)dx = E\left\{\frac{g(x)f(x)}{q(x)} | X \sim q(x)\right\}, \quad (29)$$

This expectation is calculated with respect to the density $q(x)$. Therefore, if a random sample has been drawn from $q(x)$ then $E\{g(X)\}$ can be approximated by

$$\hat{I} = \frac{1}{N}\sum_{i=0}^{N}\frac{g(x_i)f(x_i)}{q(x_i)}\,, \qquad (30)$$

Noted that,

$$E\left(\hat{I}\middle|X \sim q(x)\right) = \tfrac{1}{N}\textstyle\sum_{i=0}^{N} E\left(\frac{g(x_i)f(x_i)}{q(x_i)}\,|X \sim q(x)\right),$$

$$= \tfrac{1}{N}\textstyle\sum_{i=0}^{N} E\left(\frac{g(X)f(X)}{q(X)}\,|X \sim q(x)\right) = E\{g(X)\}\,,$$

Then $\hat{I}$ is an unbiased estimator of $E\{g(X)\}$. Similarly, we can obtain that

$$var\left(\hat{I}\middle|X \sim q(x)\right) = \frac{1}{N} var\left(\frac{g(X)f(X)}{q(X)}\middle|X \sim q(x)\right), \quad (31)$$

Therefore, the variance of $\hat{I}$ will depends on the choice of N and the approximation density $q(x)$. Further results can be found in the following standard references on Bayesian analysis, in (Berger, 1980), (Bolstad, 2007), and (Lee, 1992). To compute the Bayesian estimators, we need to find the posterior distribution which will be introduced in the following section.

3.2 Bayesian Estimate for x_p

In this chapter, we will derive the Bayesian estimators for the parameters θ, σ and for the p^{th} quantile. (Gupta & Kundu, Generalized Exponential Distribution: Bayesian Estimations, 2008) derived the Bayesian estimators for the two unknown parameters of the GE distribution. They assumed gamma distribution as prior distributions for both shape and scale parameters. The Bayesian estimators couldn't be written in the explicit form. For that reason, a simulation computation had been applied. To generate posterior samples, they proposed the Gibbs sampler procedure. Monte Carlo simulation had been applied to compare between Bayesian estimators under the assumption of non-informative priors and the maximum likelihood estimators. (Kim & Han, Bayesian Estimation of Generalized Exponential Distribution Under Progressive First Failure Censored Sample, 2015) derived the maximum likelihood and Bayesian estimators of the GE distribution based on progressive first failure censored samples. They applied Markov Chain Monte Carlo method for generating samples. For estimating the parameters and predicting future observations, they used importance sampling. For their application purposes, they used a simulated data analysis. (Mohie El-Din & Shafay, 2013) considered the one and two sample Bayesian prediction intervals based on the progressively Type-II censored data. They applied their results on some distributions such as exponential, Pareto, Weibull and Burr Type-XII models. They did some numerical computations and found, in case of one-sample, that the lower bounds are nearly insensitive to the assumed values to hyper parameters, but the upper bounds are quite sensitive. In case of two-sample both lower and upper bounds are found nearly insensitive to the assumed values to hyper parameters. The empirical Bayes approach estimating can be used as a prior parameter in case of unknown vector of hyper parameters. (El-Sagheer, 2016) consider the Bayesian approach, point and interval predictions based on general progressively cosponsored data for the generalized Pareto distribution.

Based on definition 4.1 and proportional (27), we shall find the posterior distributions for our parameters θ and σ. After that, we are interested to find the posterior distribution for the p^{th}quantile.

In this study, we are interested in the Generalized Exponential distribution as mentioned in chapter 1. Equations (3 & 4) shows the cumulative and density distribution functions for the GE distribution respectively. Now, to compute the posterior distributions we need to use the likelihood function, which is defined in equation (8). For our purposes we shall rewrite equation (8) as follows.

$$L(\underline{x};\theta,\sigma)=\left(\frac{\theta}{\sigma}\right)^{m} e^{-\sum_{i=1}^{m}x_i/\sigma}\prod_{i=1}^{m}\left(1-e^{-x_i/\sigma}\right)^{\theta-1}\left(\prod_{i=1}^{m}\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)^{R_i}\right), (32)$$

The prior distributions for parameters θ and σ have been chosen to be Gamma and Inverted Gamma distributions, respectively. Basically, both of that distributions are known as a continues distribution on the positive real line, with two parameters α and β. The exponential distribution, Erlang distribution, and chi-square distribution are al special cases of Gamma distribution. Also, we can say that if $X\sim Gamma(\alpha,\beta)$ then $\frac{1}{X}\sim Inv-Gamma(\alpha,\beta)$. The inverse gamma distribution can be used as a conjugate prior for the scale parameter. In general, the probability density function for gamma and inverse gamma distributions can be presented respectively as follows:

$$f(x;\alpha,\beta)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{\alpha-1}e^{-x/\beta} \quad ,x>0$$

$$f(x;\alpha,\beta)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}x^{-(\alpha+1)}e^{-1/\beta x} \quad ,x>0$$

with α is a shape parameter and β is a scale parameter. Where $\Gamma(.)$ denotes as a gamma function.

Now, suppose $\theta\sim G(a_0,b_0)$ and $\sigma\sim IG(a_1,b_1)$. Therefore, the prior distributions for θ and

σ are

$$\prod_1(\theta) = \frac{1}{\Gamma(a_0){b_0}^{a_0}}\theta^{a_0-1}e^{-\theta/b_0} \quad , \qquad (33)$$

$$\prod_2(\sigma) = \frac{1}{\Gamma(a_1){b_1}^{a_1}}\sigma^{-(a_1+1)}e^{-1/b_1\sigma} \quad , \qquad (34)$$

The joint posterior density function of θ and σ, can be computed by substituting equations (32, 33, & 34) into proportional (27) to get

$$\prod(\theta,\sigma|x) \propto L(x|\theta,\sigma)\prod_1(\theta)\prod_2(\sigma) \ ,$$

$$\prod(\theta,\sigma|x) \propto \left(\frac{\theta}{\sigma}\right)^m e^{-\sum_{i=1}^{m}x_i/\sigma}e^{(\theta-1)\sum_{i=1}^{m}ln\left(1-e^{-x_i/\sigma}\right)}e^{\sum_{i=1}^{m}R_i ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)}.$$

$$\frac{1}{\Gamma(a_0){b_0}^{a_0}}\theta^{a_0-1}e^{-\theta/b_0}.\frac{1}{\Gamma(a_1){b_1}^{a_1}}\sigma^{-(a_1+1)}e^{-1/b_1\sigma} \ , \quad (35)$$

$$\prod(\theta,\sigma|x) \propto \theta^{a_0+m-1}e^{-\theta\left(1/b_0-\sum_{i=1}^{m}ln\left(1-e^{-x_i/\sigma}\right)\right)}\sigma^{-(a_1+m+1)}e^{-1/\sigma\left(\sum_{i=1}^{m}x_i+1/b_1\right)}$$

$$e^{-\sum_{i=1}^{m}ln\left(1-e^{-x_i/\sigma}\right)+\sum_{i=1}^{m}R_i ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)} \ , \quad (36)$$

Equation (36) can be written as follows:

$$\prod(\theta,\sigma|x) \propto G_{\theta|\sigma}(a_0^*,b_0^*)IG_\sigma(a_1^*,b_1^*)G_3(\theta,\sigma) \ , \qquad (37)$$

where $G_{\theta|\sigma}(a_0^*,b_0^*)$ is a gamma density function with parameters $a_0^* = a_0 + m$ and $b_0^* = \frac{1}{1/b_0-\sum_{i=1}^{m}ln\left(1-e^{-x_i/\sigma}\right)}$, which are known as shape and scale parameters, respectively. The same for $IG_\sigma(a_1^*,b_1^*)$ is an inverse gamma density with parameters $a_1^* = a_1 + m$ and $b_1^* = \frac{1}{\sum_{i=1}^{m}x_i+1/b_1}$, which are known as shape and scale parameters, respectively. On the other hand,

$$G_3(\theta,\sigma) = \frac{1}{\left(\frac{1}{1/b_0 - \sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right)}\right)^{a_0+m}} e^{R_i ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right) - \sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right)}.$$

In case of noninformative prior distributions given by

$$\prod_2(\theta) = 1 , \qquad (38)$$

$$\prod_2(\sigma) = \frac{1}{\sigma} , \qquad (39)$$

We proceed as follows, the joint posterior density function can be done by substituting equations (32), (38), and (39) into proportional (27) to get the following

$$\prod(\theta,\sigma|x) \propto \theta^m e^{\theta \sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right)} \sigma^{-(m+1)} e^{-1/\sigma \sum_{i=1}^{m} x_i}$$

$$e^{-\sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right) + \sum_{i=1}^{m} R_i ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)} , \quad (40)$$

And then equation (40) can be written as follows:

$$\prod(\theta,\sigma|x) \propto G_{\theta|\sigma}\left(m+1, \frac{-1}{\sum_{i=1}^{m} ln(1-e^{-x_i/\sigma})}\right) IG_{\sigma}\left(m, \frac{1}{\sum_{i=1}^{m} x_i}\right) G_4(\theta,\sigma) , \qquad (41)$$

where $G_4(\theta,\sigma) = \frac{1}{\Gamma(m+1)\left(\frac{-1}{\sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right)}\right)^{(m+1)}} e^{-\sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right) + \sum_{i=1}^{m} R_i ln\left(1-\left(1-e^{-x_i/\sigma}\right)^{\theta}\right)}.$

3.3 Simulation Study

We have been applied importance sampling to simulate our results. The same process that we used in classical way, similarly, we generated a progressively type II censored sample from generalized exponential distribution. The schemes that displayed in table 3 has been used again in Bayesian simulation. We choose the same initial values as, $\theta_0 = 2$ and $\sigma_0 = 1.2$. We mentioned in section 3.5 in this study that our parameters θ and

σ priors distributions are Gamma and Inverse Gamma distributions, respectively; i.e $\theta \sim G(a_0, b_0)$ and $\sigma \sim IG(a_1, b_1)$. We choose different values of hyper parameters a_0, b_0, a_1, and b_1 for each scheme to avoid high bias results. For all hyper parameters, we choose nonnegative values. The simulation has been repeated for $N = 2000$, and $B = 1000$ as several samples.

The following steps are the steps that we use in simulation:

1- Generate $\sigma_1 \sim IG(\, a_1 + m, {}^{1}/_{b_1} + \sum_{i=1}^{m} x_i)$

2- Generate $\theta_1 \sim G(a_0 + m, {}^{1}/_{b_0} - \sum_{i=1}^{m} ln\left(1 - e^{-x_i/\sigma}\right))$

3- Calculate $\hat{x}_{B_p}$ from $\hat{\theta}$ and $\hat{\sigma}$ that we got from steps 1 and 2 for all values of *p*.

4- Repeat step 1, 2, and 3 N times to obtain our parameters $(\theta_1, \sigma_1), \dots, (\theta_N, \sigma_N)$.

5- The Bayes estimate is considering by

$$\hat{x}_{B_p} \approx \frac{{}^{1}/_{N} \sum_{i=1}^{N} \hat{x}_p(\theta_i,\sigma_i)\; G_3(\theta_i,\sigma_i)}{{}^{1}/_{N} \sum_{i=1}^{N} G_3(\theta_i,\sigma_i)} .$$

Bias and MSE of our Bayesian estimators $\hat{x}_{B_p}$ have been calculated too. The way of calculating bias and MSE have been explained in section 2.4.

Now, we need to calculate the equal tail Bayesian interval. And we obtain the highest posterior density region similarly as (Kundu & Pradhan, 2009), who inspired by (Chen & Shao, 1999) and (Raqab & Madi, 2005) ideas, which known as the HPD credible interval for x_{B_p}. To do so, let $\pi\left(x_{B_p}|data\right)$ and $\Pi\left(x_{B_p}|data\right)$ denote as posterior density function and posterior distribution function of x_{B_p} , respectively. And the p^{th} quantile of x_{B_p} is

$$x_{B_p}^{(p)} = inf\left\{x_{B_p}: \Pi\left(x_{B_p}|data\right) \geq p\right\},$$

where $0 < p < 1$. Therefore, for a given $x_{B_p}^*$,

$$\Pi\left(x_{B_p}^*|data\right) = E\left(I_{x_{B_p} \leq x_{B_p}^*}|data\right),$$

where $I_{x_{B_p} \leq x_{B_p}^*}$ is known as an indicator function. So, to simulate the expectation of Bayes estimator, we can obtain the following

$$\Pi\left(x_{B_p}^*|data\right) = \frac{{}^1/_N \sum_{i=1}^{N} I_{x_{B_p} \leq x_{B_p}^*} \; G_3(\theta_i,\sigma_i)}{{}^1/_N \sum_{i=1}^{N} G_3(\theta_i,\sigma_i)},$$

The next step, is order the value of $\left\{x_{B_p i}\right\}$, which has been calculated from step 3 in the simulation steps above. Then, we calculate the following

$$w_i = \frac{G_3(\theta_i,\sigma_i)}{\sum_{i=1}^{N} G_3(\theta_i,\sigma_i)}, \forall\ i = 1, \dots, N$$

After that we have

$$\Pi\left(x_{B_p}^*|data\right) = \begin{cases} 0 & if \quad x_{B_p}^* < x_{B_{p(1)}}, \\ \sum_{j=1}^{i} w_j & if \quad x_{B_{p(i)}} < x_{B_p}^* < x_{B_{p(i+1)}}, \\ 1 & if \quad x_{B_p}^* \geq x_{B_{p(n)}}. \end{cases}$$

The equal tail Bayes interval can be calculated by

$$\hat{x}_{B_p}\left(\frac{\alpha}{2}\right) = x_{B_{p(i)}} \quad if \quad \sum_{j=1}^{i-1} w_j < \frac{\alpha}{2} \leq \sum_{j=1}^{i} w_j\ ,$$

and

$$\hat{x}_{B_p}\left(1-\frac{\alpha}{2}\right) = x_{B_{p(i)}} \quad if \quad \sum_{j=1}^{i-1} w_j < 1-\frac{\alpha}{2} \leq \sum_{j=1}^{i} w_j\ .$$

But to calculate the HPD credible interval for x_{B_p} we need to approximate $x_{B_p}^{(p)}$ by

$$\hat{x}_{B_p}{}^{(p)} = \begin{cases} x_{B_p(1)} & if \qquad p = 0, \\ x_{B_p(i)} & if \quad \sum_{j=1}^{i-1} w_j < p \leq \sum_{j=1}^{i} w_j \,. \end{cases}$$

At the end, to calculate a $100(1-p)\%$ HPD credible interval for x_{B_p}, as the following

$$H_j = \left(\hat{x}_{B_p}{}^{\left(\frac{j}{N}\right)}, \hat{x}_{B_p}{}^{\left(\frac{j+(1-p)N}{N}\right)} \right), \forall\, j = 1, \dots, p\,N$$

Noted that we choose the smallest width H_{j^*} from all H_j's.

For the case of non-informative prior distributions, the following simulation is used in the simulations

Then the simulation steps will be similar as for Gamma and inverse gamma priors, but with little differences:

1- Generate $\sigma_1 \sim IG(\, m, \frac{1}{\sum_{i=1}^{m} x_i})$

2- Generate $\theta_1 \sim G(m+1, \frac{-1}{\sum_{i=1}^{m} ln\left(1-e^{-x_i/\sigma}\right)})$

3- Calculate $\hat{x}_{B_p}$ from $\hat{\theta}$ and $\hat{\sigma}$ that we got from steps 1 and 2 for all values of *p*.

4- Repeat step 1, 2, and 3 N times to obtain our parameters $(\theta_1, \sigma_1), \dots, (\theta_N, \sigma_N)$.

5- The Bayes estimate is considering by

$$\hat{x}_{B_p} \approx \frac{{}^{1}/{}_{N} \sum_{i=1}^{N} \hat{x}_p(\theta_i, \sigma_i)\; G_4(\theta_i, \sigma_i)}{{}^{1}/{}_{N} \sum_{i=1}^{N} G_4(\theta_i, \sigma_i)} \,.$$

Now, we need to calculate the equal tail Bayesian interval. And the same as above for gamma and inverse gamma prior distributions, we have done the same steps. But the simulation of the Bayes estimator's expectation will be

$$\Pi\left(x^*_{B_p}|data\right) = \frac{{}^{1}/_{N}\sum_{i=1}^{N} I_{x_{B_p} \le x^*_{B_p}} \; G_4(\theta_i,\sigma_i)}{{}^{1}/_{N}\sum_{i=1}^{N} G_4(\theta_i,\sigma_i)}.$$

And ordering the value of $\left\{x_{B_p i}\right\}$, which has been calculated from step 3 in the simulation steps above will be as follows

$$w_i = \frac{G_4(\theta_i,\sigma_i)}{\sum_{i=1}^{N} G_4(\theta_i,\sigma_i)}, \forall\ i = 1,\dots,N$$

Also, we have

$$\Pi\left(x^*_{B_p}|data\right) = \begin{cases} 0 & if \quad x^*_{B_p} < x_{B_{p(1)}}, \\ \sum_{j=1}^{i} w_j & if \quad x_{B_{p(i)}} < x^*_{B_p} < x_{B_{p(i+1)}}, \\ 1 & if \quad x^*_{B_p} \ge x_{B_{p(n)}}. \end{cases}$$

The equal tail Bayes interval can be calculated by

$$\hat{x}_{B_p}\left(\frac{\alpha}{2}\right) = x_{B_{p(i)}} \quad if \quad \sum_{j=1}^{i-1} w_j < \frac{\alpha}{2} \le \sum_{j=1}^{i} w_j\ ,$$

and

$$\hat{x}_{B_p}\left(1-\frac{\alpha}{2}\right) = x_{B_{p(i)}} \quad if \quad \sum_{j=1}^{i-1} w_j < 1-\frac{\alpha}{2} \le \sum_{j=1}^{i} w_j.$$

But to calculate the HPD credible interval for x_{B_p} we need to approximate $x^{(p)}_{B_p}$ by

$$\hat{x}_{B_p}{}^{(p)} = \begin{cases} x_{B_{p(1)}} & if \quad p = 0, \\ x_{B_{p(i)}} & if \quad \sum_{j=1}^{i-1} w_j < p \le \sum_{j=1}^{i} w_j\ . \end{cases}$$

At the end, to calculate a $100(1-p)\%$ HPD credible interval for x_{B_p}, as the following

$$H_j = \left(\hat{x}_{B_p}\left(\frac{j}{N}\right), \hat{x}_{B_p}\left(\frac{j+(1-p)N}{N}\right) \right), \forall\, j = 1, \dots, p\,N$$

Noted that we choose the smallest width H_{j^*} from all H_j's. Tables below display our results for bias, MSE, average lengths and coverage probabilities for Bayesian approach. Noticed that the following tables are presenting the results of gamma and inverse gamma prior distributions

Table 10. Bias and MSE results for informative priors.

Scheme		**Hyper parameters**	**P=0.1**	**P=0.25**	**P=0.5**	**P=0.75**	**P=0.9**
1	Bias	$a_0 = 1.5, b_0 = 3$	0.012	0.009	0.008	0.016	0.03
	MSE	$a_1 = 1.5, b_1 = 0.7$	0.008	0.015	0.036	0.109	0.282
2	Bias	$a_0 = 1.5, b_0 = 3$	0.016	0.012	0.01	0.011	0.016
	MSE	$a_1 = 1.5, b_1 = 0.7$	0.008	0.014	0.029	0.077	0.192
3	Bias	$a_0 = 2.4, b_0 = 1$	0.001	0.004	0.016	0.043	0.082
	MSE	$a_1 = 1.5, b_1 = 0.7$	0.005	0.01	0.023	0.067	0.168
4	Bias	$a_0 = 3, b_0 = 0.5$	-0.005	-0.003	0.009	0.036	0.076
	MSE	$a_1 = 1.5, b_1 = 0.5$	0.002	0.003	0.008	0.024	0.066
5	Bias	$a_0 = 1.3, b_0 = 0.7$	-0.018	-0.016	0.004	0.047	0.072
	MSE	$a_1 = 1.2, b_1 = 0.6$	0.004	0.007	0.016	0.049	0.137
6	Bias	$a_0 = 2, b_0 = 0.5$	-0.008	-0.008	0.003	0.03	0.07
	MSE	$a_1 = 1.5, b_1 = 0.5$	0.001	0.003	0.006	0.019	0.054
7	Bias	$a_0 = 1.5, b_0 = 1.2$	-0.003	-0.004	0	0.012	0.03
	MSE	$a_1 = 1.6, b_1 = 0.7$	0.003	0.005	0.01	0.03	0.08

Note that these hyper parameters presented in table 10 has been chosen to get the expectation of gamma prior and inverse gamma prior to be the value of $\theta_0 = 2$ and $\sigma_0 = 1.2$ or nearly to $\theta_0 = 2$ and $\sigma_0 = 1.2$ to get the best results. Also, the same hyper parameters of each scheme have been used to get the results of both Bayesian intervals under each censoring scheme.

From table 10, we noticed that both values of bias and MSE are increasing over all the p^{th} quantiles values. Also, it is notable that in general there are no significant difference between bias and MSE all over our schemes, but we can say that values of MSE are decreasing while m and n are getting bigger.

Now, the following tables 11 and 12 present the results of the average length and error rates for equal tail intervals and highest posterior density region. The simulation has been done for both $\alpha = 0.1$ and $\alpha = 0.05$.

Table 11. Coverage probability and expected lengths results when $\alpha = 0.1$ for informative priors.

		$\alpha = 0.1$									
Scheme	**Interval Type**	**P=0.1**		**P=0.25**		**P=0.5**		**P=0.75**		**P=0.9**	
		A L	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**
1	E T	0.295	0.104	0.396	0.124	0.609	0.132	1.053	0.136	1.697	0.128
	HPD	0.27	0.149	0.368	0.154	0.567	0.165	0.956	0.169	1.518	0.166
2	E T	0.295	0.103	0.377	0.118	0.54	0.128	0.899	0.112	1.447	0.107
	HPD	0.276	0.147	0.357	0.141	0.512	0.14	0.835	0.146	1.321	0.146
3	E T	0.236	0.103	0.317	0.125	0.475	0.137	0.792	0.136	1.262	0.129
	HPD	0.215	0.152	0.294	0.157	0.441	0.161	0.724	0.161	1.135	0.172
4	E T	0.127	0.119	0.183	0.111	0.295	0.102	0.524	0.102	0.86	0.099
	HPD	0.113	0.151	0.166	0.14	0.27	0.132	0.48	0.137	0.787	0.121
5	E T	0.196	0.154	0.262	0.151	0.409	0.117	0.729	0.101	1.202	0.097
	HPD	0.173	0.176	0.235	0.172	0.376	0.148	0.674	0.132	1.109	0.118
6	E T	0.125	0.1	0.173	0.105	0.261	0.121	0.451	0.122	0.744	0.115
	HPD	0.113	0.13	0.158	0.123	0.242	0.137	0.42	0.148	0.688	0.137
7	E T	0.151	0.139	0.207	0.138	0.318	0.138	0.553	0.122	0.903	0.112
	HPD	0.133	0.178	0.187	0.176	0.294	0.168	0.512	0.156	0.834	0.143

Table 12. Coverage probability and expected lengths results when $\alpha = 0.05$ for informative priors.

		$\alpha = 0.05$									
Scheme	**Interval Type**	**P=0.1**		**P=0.25**		**P=0.5**		**P=0.75**		**P=0.9**	
		A L	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**
1	E T	0.345	0.056	0.476	0.066	0.742	0.076	1.278	0.076	2.059	0.072
	HPD	0.319	0.08	0.445	0.087	0.692	0.088	1.176	0.093	1.869	0.093
2	E T	0.352	0.043	0.456	0.064	0.654	0.069	1.088	0.069	1.75	0.058
	HPD	0.328	0.07	0.431	0.08	0.621	0.086	1.016	0.074	1.608	0.072
3	E T	0.276	0.055	0.382	0.063	0.573	0.078	0.953	0.089	1.518	0.088
	HPD	0.255	0.078	0.357	0.082	0.539	0.098	0.886	0.094	1.391	0.094
4	E T	0.14	0.096	0.207	0.08	0.342	0.06	0.615	0.047	1.018	0.049
	HPD	0.129	0.099	0.191	0.095	0.318	0.074	0.573	0.067	0.943	0.063
5	E T	0.217	0.136	0.295	0.124	0.473	0.078	0.857	0.058	1.421	0.052
	HPD	0.198	0.148	0.271	0.136	0.44	0.095	0.801	0.079	1.327	0.07
6	E T	0.139	0.065	0.198	0.071	0.305	0.068	0.532	0.062	0.88	0.059
	HPD	0.128	0.082	0.183	0.08	0.285	0.083	0.498	0.081	0.821	0.074
7	E T	0.169	0.098	0.236	0.098	0.372	0.08	0.654	0.069	1.073	0.059
	HPD	0.153	0.132	0.217	0.118	0.348	0.09	0.613	0.093	1	0.08

Mostly, the coverage probability of those Bayesian intervals doesn't attain the nominal error rate. And this is very clear for the HPD intervals. Where the nominal error rate in this situation has been explained in chapter 2, section 2.4. The results for the average lengths for both kinds of intervals are increased while the p^{th} quantile values are getting bigger. On the other hand, the error rates didn't show similar remark.

The average lengths of the highest posterior density region are smaller than the average lengths of the equal tail intervals all over our different schemes and p^{th} quantile values. In contrast the error rates of HPD intervals for most situations are bigger than the equal tail intervals'. For $\alpha = 0.1$ the average lengths for both kind of Bayesian intervals are smaller than the average lengths when $\alpha = 0.05$.

It is worth to mention that error probabilities result for scheme 7 seems to have a problem. Specially for $\alpha = 0.05$, the error probability for the HPD interval is around 0.1 when P=0.1 and P=0.25 under scheme 6.

We have chosen four schemes to plot their results. These schemes are 1, 2, 3, and 6. This step is done to clarify our results. As what we have done in chapter 2, we plotted p^{th} quantile values on the x-access and the results of the bias, MSE, intervals lengths, and coverage probabilities. Figures 7-11 represent our purposes.

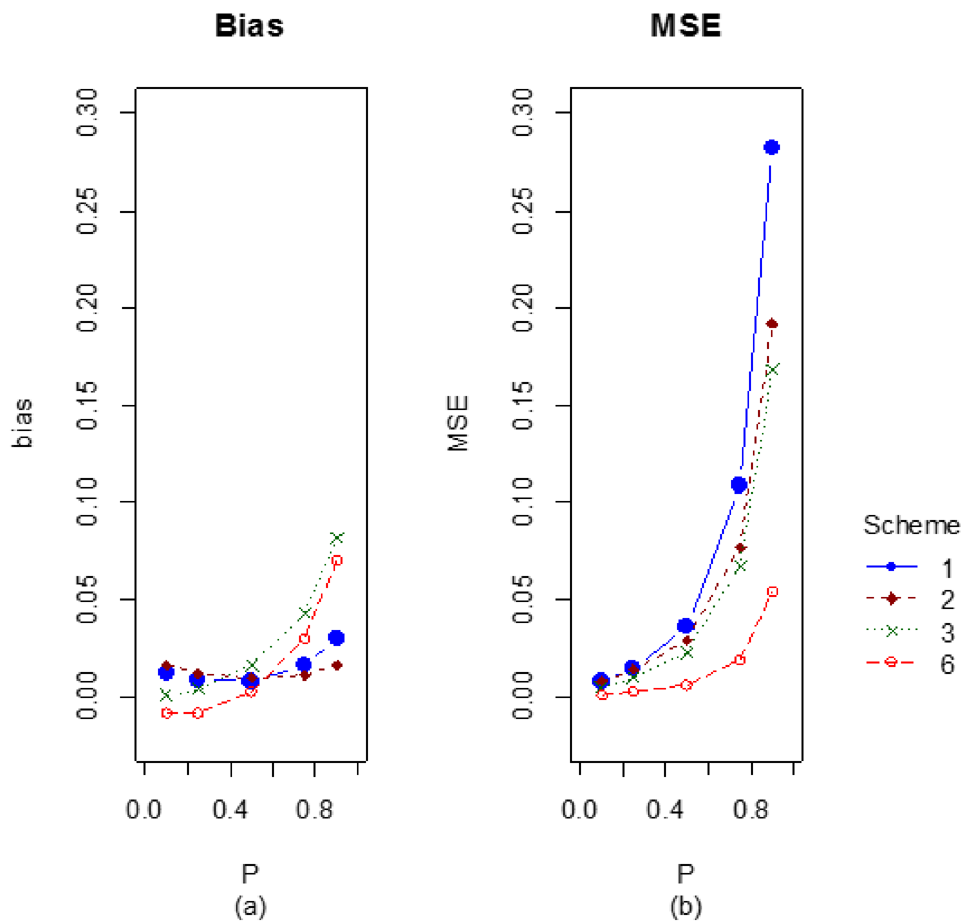

Figure 7. Bayesian bias and MSE plots for informative priors.

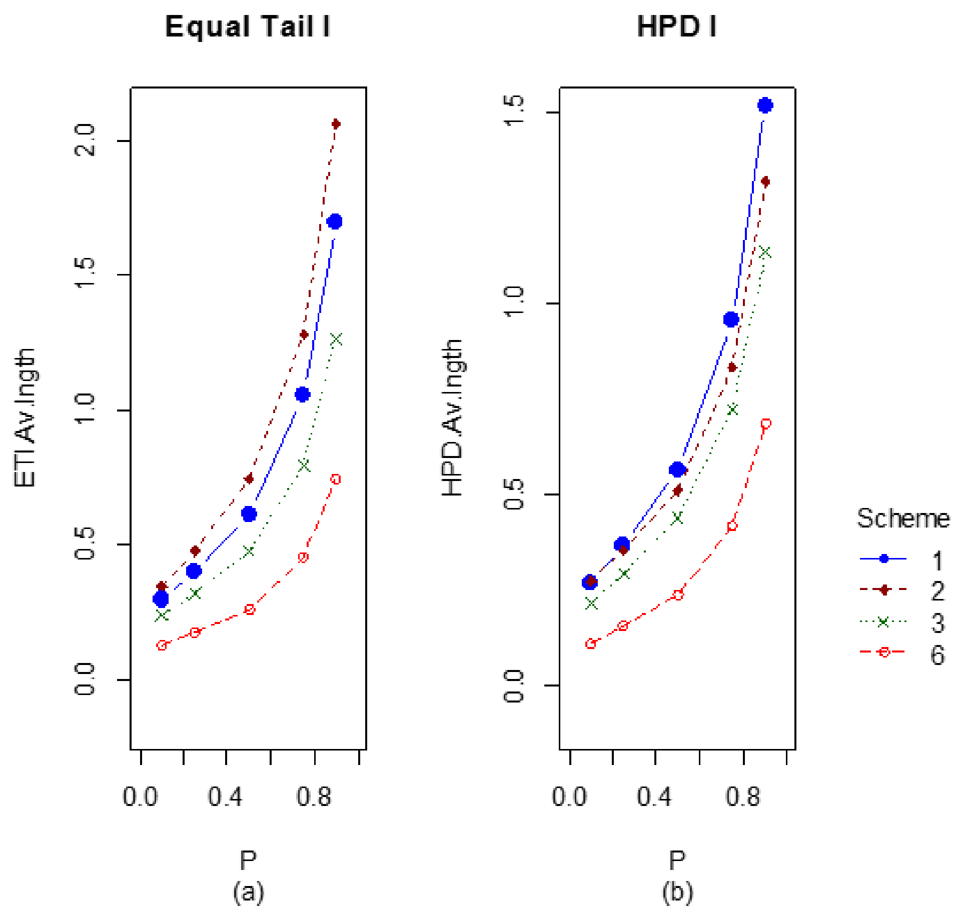

Figure 8. Bayesian expected lengths plots when $\alpha = 0.1$ for informative priors

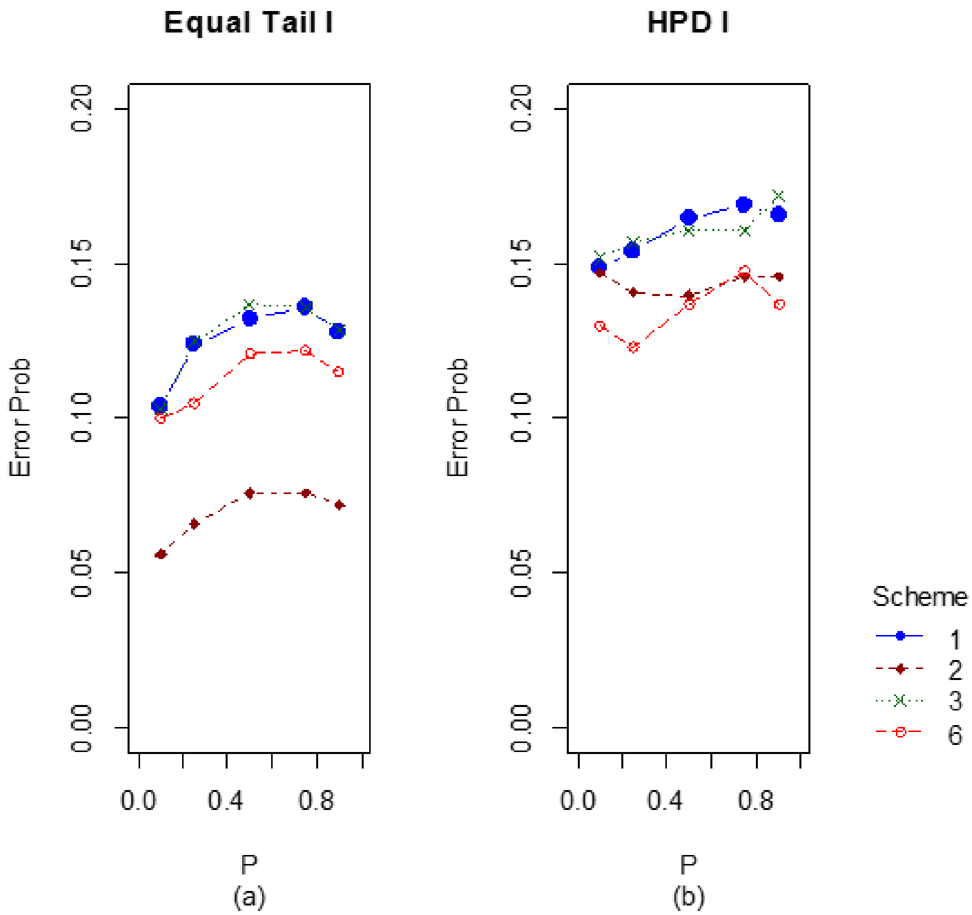

Figure 9. Bayesian coverage probability plots when $\alpha = 0.1$ for informative priors

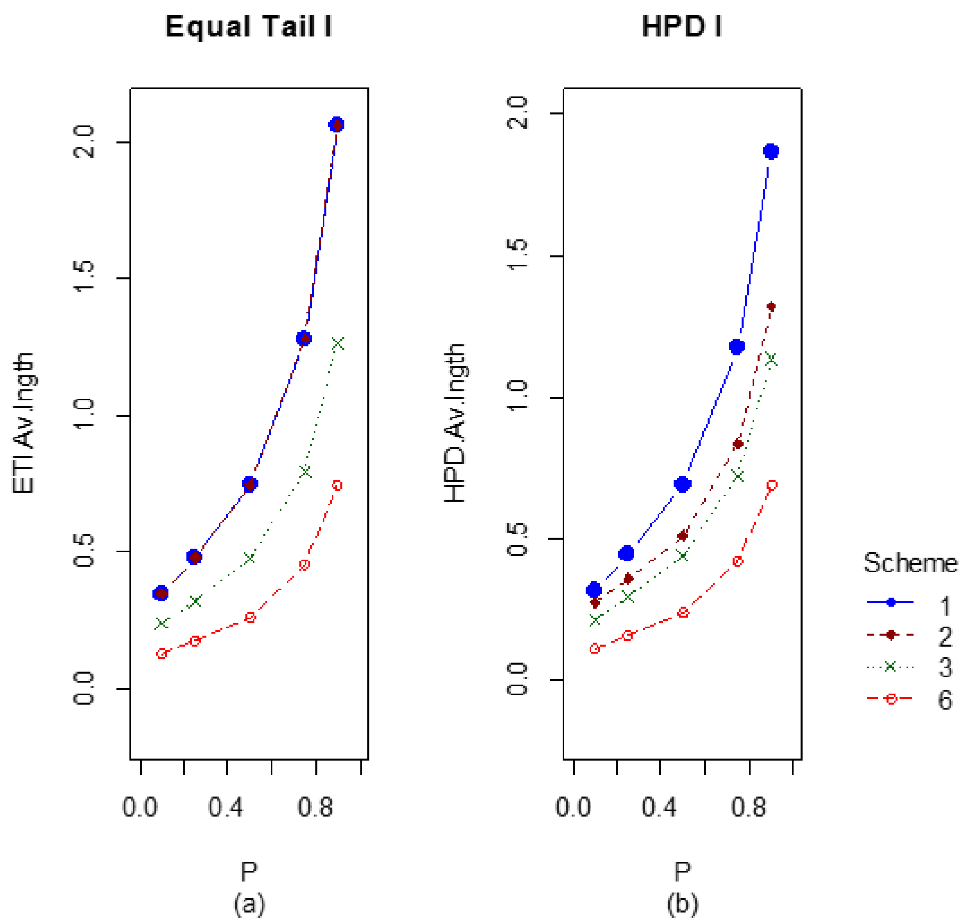

Figure 10. Bayesian expected lengths plots when $\alpha = 0.05$ for informative priors

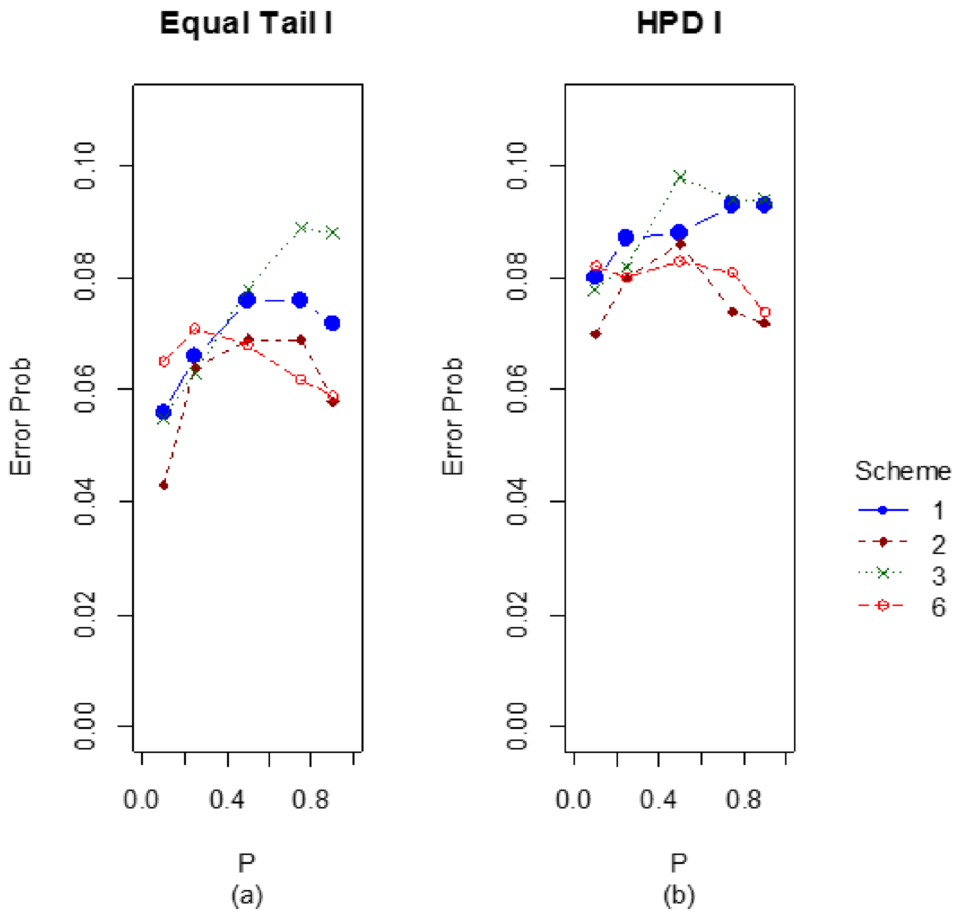

Figure 11. Bayesian coverage probability plots when $\alpha = 0.05$ for informative priors

Same process that we have done for the classical way, which implies choosing other values of θ and σ. We also repeated our simulation for Bayesian statistics, choosing the same values of θ and σ, just to make sure that our processes are perfectly done. Noted that the same technique that we depend in choosing the hyper parameters for $\theta = 2$ and $\sigma = 1.2$, we do the same for other values of θ and σ. The tables below show our results for different values of θ and σ.

Table 13. Bias and MSE results for different values of θ and σ for informative priors

P=0.5					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Hyper parameters**	**Bias**	**MSE**
1	$\theta = 1.2$	$\sigma = 0.5$	$a_0 = 1, b_0 = 1.2$ $a_1 = 1.5, b_1 = 0.25$	0.057	0.009
2			$a_0 = 1, b_0 = 1.2$ $a_1 = 1.5, b_1 = 0.25$	0.036	0.006
3			$a_0 = 1, b_0 = 1.2$ $a_1 = 1.5, b_1 = 0.25$	0.032	0.004
1	$\theta = 1.5$	$\sigma = 0.7$	$a_0 = 1.5, b_0 = 3$ $a_1 = 1.5, b_1 = 0.7$	0.025	0.012
2			$a_0 = 1.5, b_0 = 3$ $a_1 = 1.5, b_1 = 0.7$	0.022	0.009
3			$a_0 = 2.4, b_0 = 1$ $a_1 = 1.5, b_1 = 0.7$	0.015	0.007
1	$\theta = 1.7$	$\sigma = 0.9$	$a_0 = 1.9, b_0 = 0.9$ $a_1 = 1.5, b_1 = 0.45$	0.031	0.006
2			$a_0 = 1.9, b_0 = 0.9$ $a_1 = 1.5, b_1 = 0.45$	0.019	0.005
3			$a_0 = 1.9, b_0 = 0.9$ $a_1 = 1.5, b_1 = 0.45$	0.019	0.003

Table 14. Coverage probability and expected lengths results of different values of θ and σ and $\alpha = 0.1$ for informative priors

$\boldsymbol{\alpha = 0.1}$ **&** $\boldsymbol{p = 0.5}$					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Interval Type**	**A L**	**E R**
1	$\theta = 1.2$	$\sigma = 0.5$	E T	0.241	0.152
			HPD	0.217	0.249
2			E T	0.214	0.147
			HPD	0.194	0.215
3			E T	0.18	0.15
			HPD	0.162	0.225
1	$\theta = 1.5$	$\sigma = 0.7$	E T	0.354	0.126
			HPD	0.328	0.143
2			E T	0.313	0.11
			HPD	0.294	0.108
3			E T	0.275	0.122
			HPD	0.256	0.138
1	$\theta = 1.7$	$\sigma = 0.9$	E T	0.232	0.145
			HPD	0.211	0.178
2			E T	0.204	0.127
			HPD	0.186	0.168
3			E T	0.175	0.144
			HPD	0.159	0.204

Table 15. Coverage probability and expected lengths results of different values of θ and σ and $\alpha = 0.05$ for informative priors

$\boldsymbol{\alpha = 0.05}$ & $\boldsymbol{p = 0.5}$					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Interval Type**	**A L**	**E R**
1	$\theta = 1.2$	$\sigma = 0.5$	E T	0.262	0.09
			HPD	0.243	0.145
2			E T	0.24	0.088
			HPD	0.222	0.133
3			E T	0.197	0.094
			HPD	0.182	0.144
1	$\theta = 1.5$	$\sigma = 0.7$	E T	0.417	0.074
			HPD	0.389	0.087
2			E T	0.372	0.057
			HPD	0.35	0.065
3			E T	0.325	0.075
			HPD	0.304	0.091
1	$\theta = 1.7$	$\sigma = 0.9$	E T	0.255	0.1
			HPD	0.237	0.13
2			E T	0.229	0.085
			HPD	0.212	0.118
3			E T	0.194	0.097
			HPD	0.179	0.13

Comparing between results in Tables 10 and 13, in case of $p = 0.5$, we found that results of bias for $\theta = 2$ and $\sigma = 1.2$ are smaller than the results for other different values of θ and σ. On the other hand, MSE results show the opposite. When considering tables 11, 12, 14, and 15. It seems that the error rates when $\alpha = 0.1$ or $\alpha = 0.05$ for $p = 0.5$, mostly don't attain the nominal error rate for each interval. Then, we can say that Bayesian intervals for any value of θ and σ are not comparable. Note that, the HPD intervals show more problems to attain the nominal error rate.

Now, the following tables are presenting results of the noninformative prior distributions.

Table 16. Bias and MSE results of noninformative prior distributions.

Scheme		**P=0.1**	**P=0.25**	**P=0.5**	**P=0.75**	**P=0.9**
1	Bias	0.02	0.019	0.021	0.031	0.049
	MSE	0.01	0.017	0.04	0.122	0.319
2	Bias	0.02	0.018	0.019	0.026	0.039
	MSE	0.009	0.014	0.028	0.08	0.212
3	Bias	0.009	0.014	0.032	0.065	0.111
	MSE	0.006	0.012	0.027	0.075	0.188
4	Bias	0.004	0.007	0.019	0.044	0.079
	MSE	0.005	0.009	0.021	0.062	0.159
5	Bias	0.005	0.004	0.007	0.017	0.031
	MSE	0.004	0.007	0.017	0.048	0.124
6	Bias	-0.003	0	0.013	0.039	0.076
	MSE	0.004	0.007	0.016	0.045	0.115
7	Bias	-0.007	-0.006	0.004	0.026	0.058
	MSE	0.003	0.005	0.013	0.038	0.096

In general, bias and MSE values are increased while the p^{th} quantiles are increased. In contrast, these results are decreasing when m and n are getting bigger. But the bias value of scheme 5 for P=0.9 is 0.111, which is quite high. On the other hand, MSE values of all schemes are around 0.1 for P=0.9, which is also quite high.

Table 17. Coverage probability and expected lengths results when $\alpha = 0.1$ for noninformative prior distributions.

		$\alpha = 0.1$									
Scheme	**Interval Type**	**P=0.1**		**P=0.25**		**P=0.5**		**P=0.75**		**P=0.9**	
		A L	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**
1	Bayes	0.291	0.126	0.39	0.135	0.611	0.142	1.074	0.148	1.744	0.142
	HPD	0.271	0.162	0.368	0.16	0.574	0.17	0.985	0.17	1.567	0.17
2	Bayes	0.292	0.1	0.374	0.104	0.544	0.122	0.918	0.122	1.488	0.12
	HPD	0.277	0.136	0.359	0.128	0.52	0.135	0.859	0.145	1.363	0.148
3	Bayes	0.236	0.134	0.318	0.142	0.478	0.152	0.805	0.158	1.286	0.16
	HPD	0.219	0.17	0.299	0.17	0.45	0.178	0.744	0.182	1.165	0.183
4	Bayes	0.199	0.133	0.274	0.156	0.426	0.16	0.729	0.156	1.163	0.154
	HPD	0.181	0.168	0.254	0.184	0.397	0.192	0.669	0.189	1.051	0.188
5	Bayes	0.182	0.156	0.242	0.166	0.38	0.145	0.675	0.127	1.107	0.116
	HPD	0.16	0.199	0.218	0.206	0.35	0.182	0.622	0.157	1.013	0.156
6	Bayes	0.192	0.12	0.256	0.134	0.377	0.154	0.624	0.166	0.99	0.168
	HPD	0.176	0.156	0.239	0.158	0.354	0.182	0.579	0.188	0.903	0.185
7	Bayes	0.151	0.163	0.207	0.172	0.321	0.161	0.548	0.165	0.874	0.168
	HPD	0.134	0.204	0.189	0.208	0.297	0.208	0.504	0.196	0.796	0.198

Table 18. Coverage probability and expected lengths results when $\alpha = 0.05$ for noninformative prior distributions.

$\alpha = 0.05$											
Scheme	**Interval Type**	**P=0.1**		**P=0.25**		**P=0.5**		**P=0.75**		**P=0.9**	
		A L	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**	**A L**	**E R**
1	Bayes	0.343	0.069	0.468	0.08	0.741	0.077	1.305	0.08	2.122	0.083
	HPD	0.323	0.092	0.446	0.094	0.702	0.094	1.209	0.093	1.931	0.094
2	Bayes	0.346	0.06	0.449	0.062	0.656	0.067	1.109	0.065	1.794	0.064
	HPD	0.329	0.068	0.433	0.07	0.631	0.074	1.042	0.081	1.656	0.086
3	Bayes	0.278	0.071	0.378	0.078	0.576	0.09	0.973	0.102	1.553	0.102
	HPD	0.26	0.098	0.359	0.098	0.547	0.102	0.908	0.106	1.425	0.102
4	Bayes	0.233	0.086	0.326	0.098	0.51	0.094	0.872	0.098	1.393	0.102
	HPD	0.215	0.108	0.305	0.111	0.481	0.11	0.813	0.118	1.281	0.113
5	Bayes	0.206	0.125	0.278	0.122	0.444	0.094	0.798	0.078	1.314	0.071
	HPD	0.186	0.145	0.255	0.138	0.415	0.117	0.746	0.102	1.22	0.088
6	Bayes	0.224	0.076	0.303	0.078	0.452	0.086	0.749	0.101	1.187	0.107
	HPD	0.208	0.092	0.286	0.092	0.429	0.108	0.701	0.118	1.096	0.119
7	Bayes	0.174	0.118	0.244	0.114	0.382	0.097	0.656	0.098	1.048	0.099
	HPD	0.158	0.136	0.225	0.132	0.358	0.12	0.611	0.123	0.968	0.124

The general observation on both tables 17 and 18 is that most of the coverage probability for all schemes didn’t attain the nominal error rate, therefore these intervals in this case are not comparable.

We also choose the same schemes as in informative priors’ case and plot their results. Figures 12-16 represent our purposes.

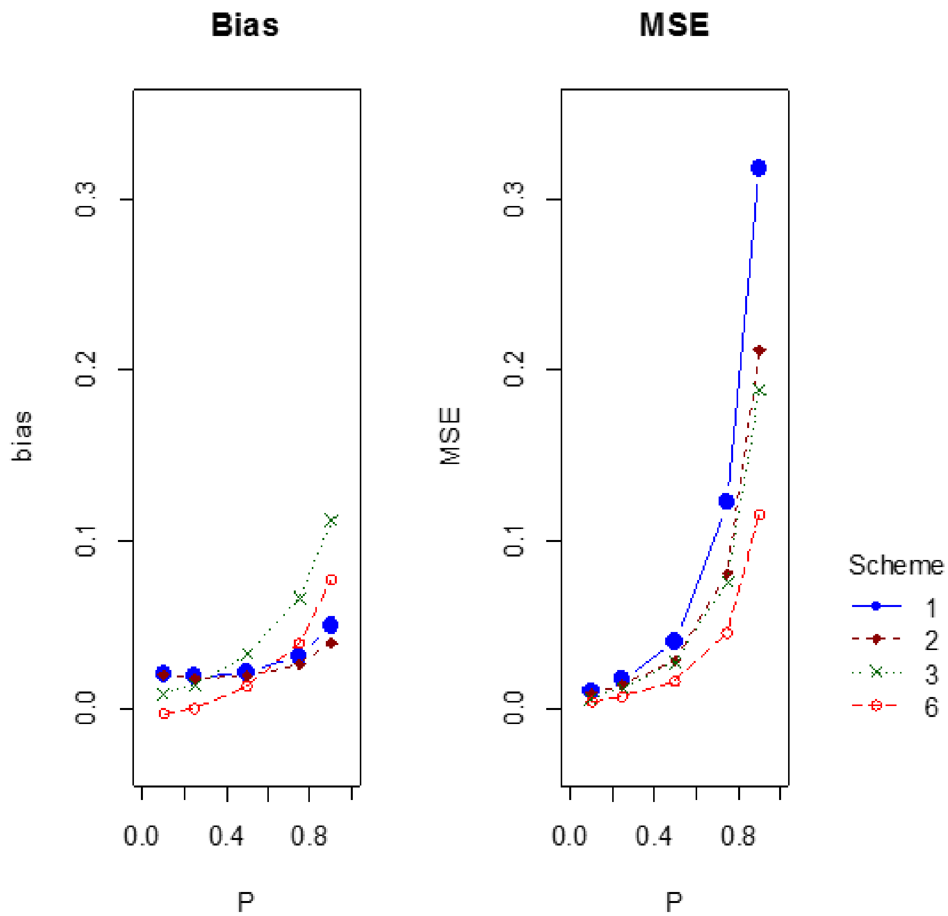

Figure 12. Bayesian bias and MSE plots for noninformative priors.

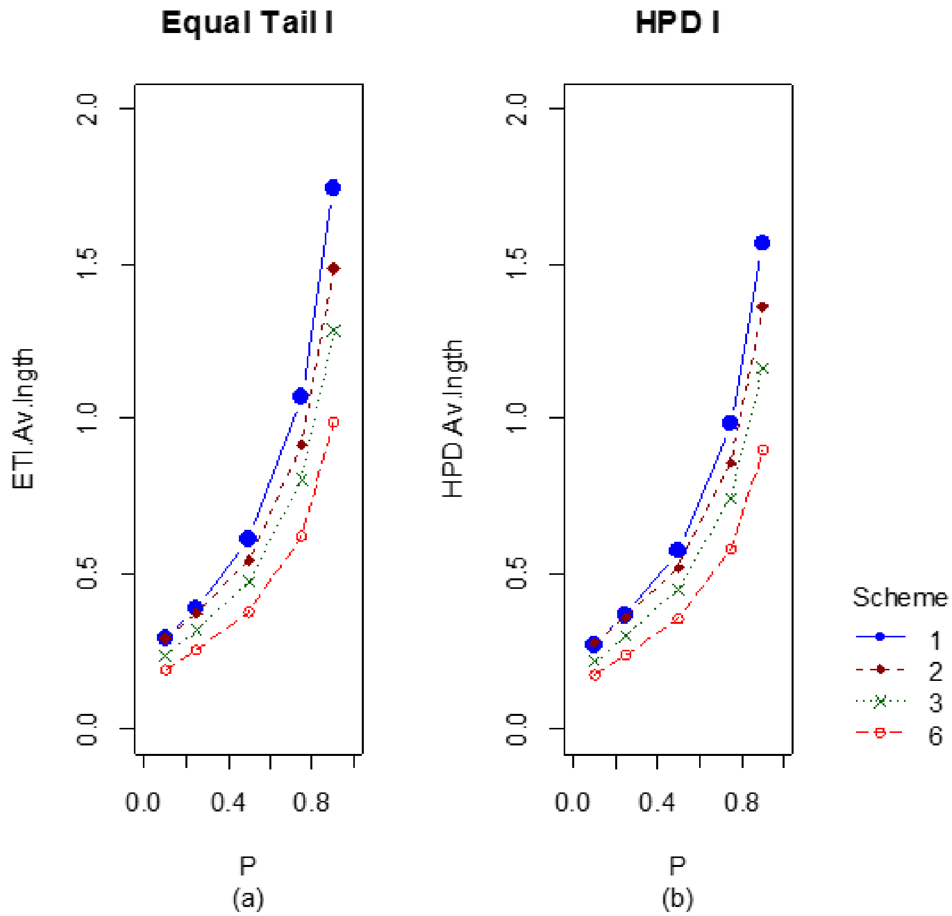

Figure 13. Bayesian expected lengths plots when $\alpha = 0.1$ for noninformative priors

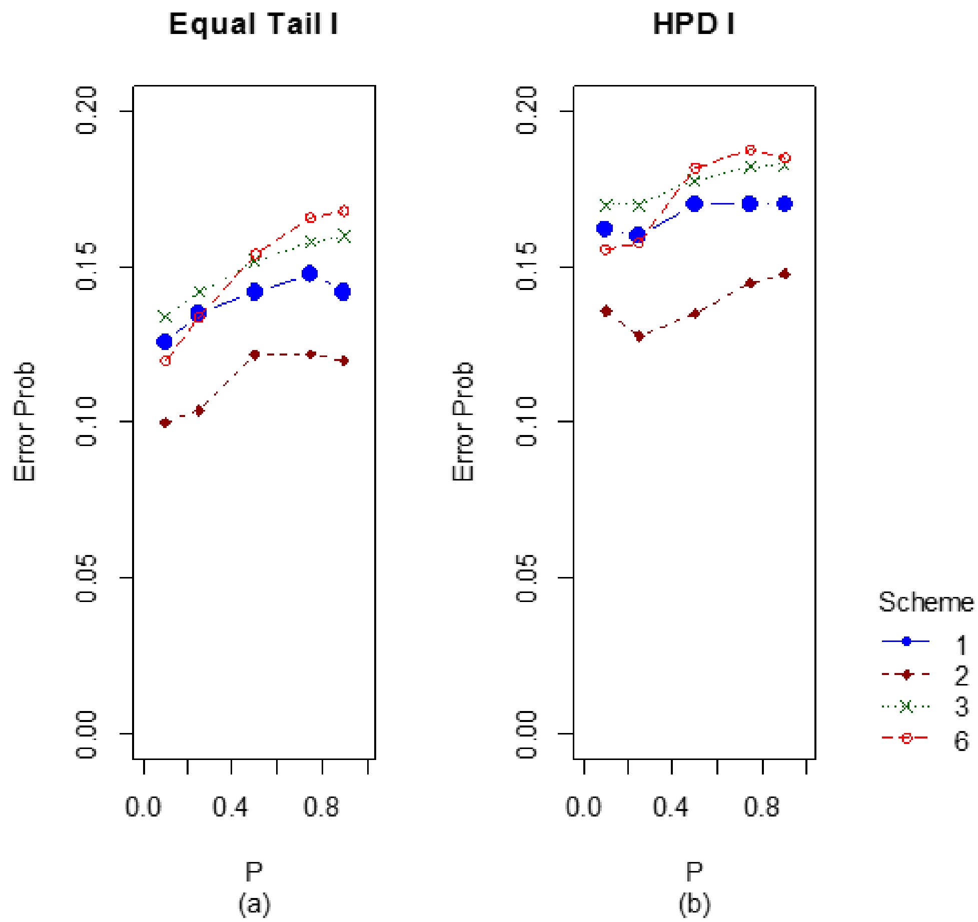

Figure 14. Bayesian coverage probability plots when $\alpha = 0.1$ for noninformative priors

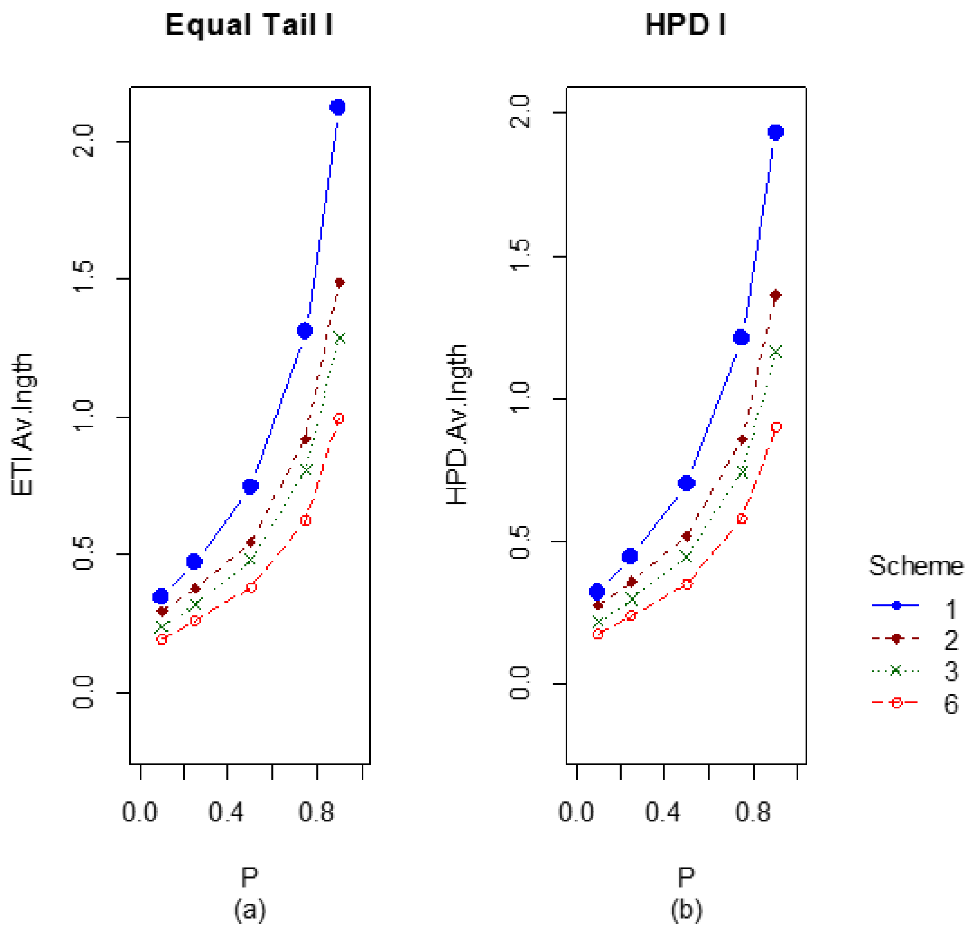

Figure 15. Bayesian expected lengths plots when $\alpha = 0.05$ for noninformative priors

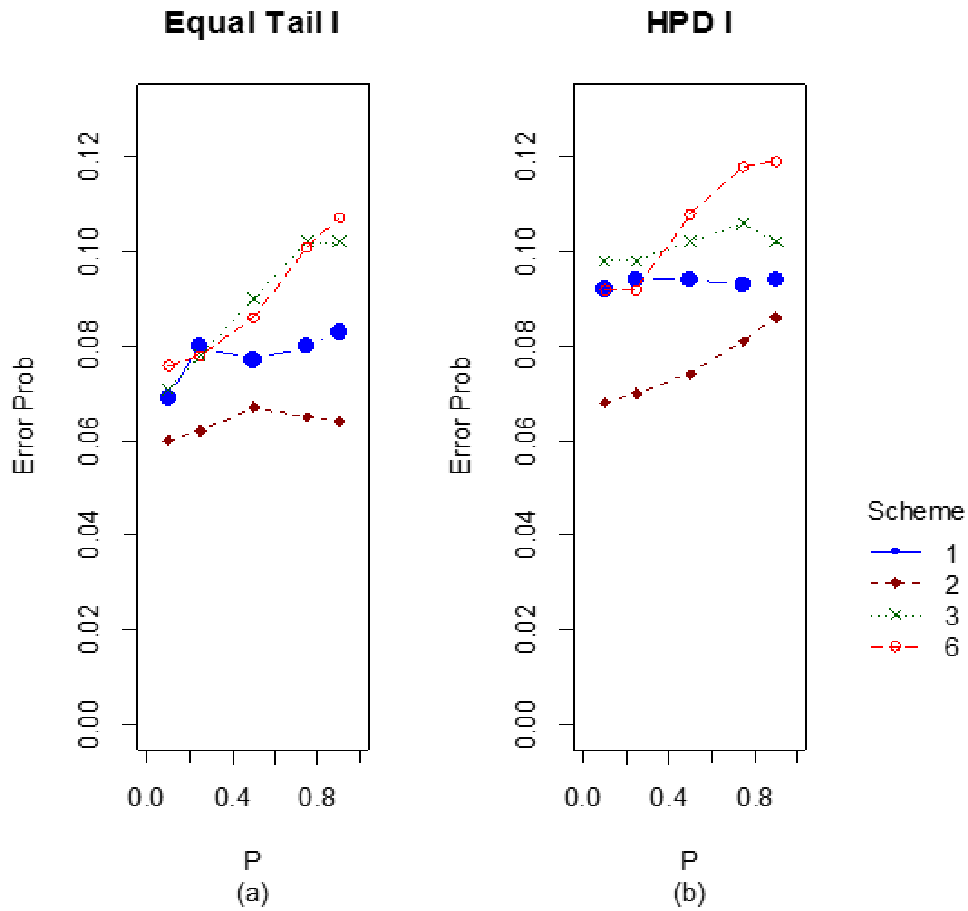

Figure 16. Bayesian coverage probability plots when $\alpha = 0.05$ for noninformative priors.

To be consistent with the informative prior distribution case, we apply the same process of choosing different values of θ and σ to apply it for the noninformative prior distribution. Our results in this case are shown in the tables below.

Table 19. Bias and MSE results for different values of θ and σ for noninformative priors.

P=0.5				
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Bias**	**MSE**
1	$\theta = 1.2$	$\sigma = 0.5$	0.004	0.005
2			0.008	0.004
3			0.001	0.003
1	$\theta = 1.5$	$\sigma = 0.7$	0.010	0.011
2			0.011	0.009
3			0.009	0.007
1	$\theta = 1.7$	$\sigma = 0.9$	0.013	0.019
2			0.014	0.016
3			0.007	0.012

Table 20. Coverage probability and expected lengths results of different values of θ and σ and $\alpha = 0.1$ for noninformative priors.

$\boldsymbol{\alpha = 0.1}$ & $\boldsymbol{p = 0.5}$					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Interval Type**	**A L**	**E R**
1	$\theta = 1.2$	$\sigma = 0.5$	E T	0.213	0.156
			HPD	0.197	0.189
2			E T	0.188	0.157
			HPD	0.176	0.176
3			E T	0.167	0.154
			HPD	0.154	0.181
1	$\theta = 1.5$	$\sigma = 0.7$	E T	0.339	0.122
			HPD	0.316	0.147
2			E T	0.301	0.120
			HPD	0.285	0.139
3			E T	0.263	0.128
			HPD	0.247	0.155
1	$\theta = 1.7$	$\sigma = 0.9$	E T	0.451	0.133
			HPD	0.421	0.149
2			E T	0.395	0.126
			HPD	0.376	0.147
3			E T	0.346	0.124
			HPD	0.326	0.152

Table 21. Coverage probability and expected lengths results of different values of θ and σ, and $\alpha = 0.05$ for noninformative priors.

$\boldsymbol{\alpha = 0.05}$ & $\boldsymbol{p = 0.5}$					
Scheme	$\boldsymbol{\theta}$	$\boldsymbol{\sigma}$	**Interval Type**	**A L**	**E R**
1	$\theta = 1.2$	$\sigma = 0.5$	E T	0.244	0.112
			HPD	0.229	0.130
2			E T	0.219	0.117
			HPD	0.206	0.125
3			E T	0.191	0.104
			HPD	0.179	0.125
1	$\theta = 1.5$	$\sigma = 0.7$	E T	0.400	0.076
			HPD	0.376	0.088
2			E T	0.356	0.073
			HPD	0.338	0.085
3			E T	0.311	0.069
			HPD	0.293	0.083
1	$\theta = 1.7$	$\sigma = 0.9$	E T	0.540	0.069
			HPD	0.508	0.080
2			E T	0.474	0.076
			HPD	0.452	0.083
3			E T	0.415	0.077
			HPD	0.391	0.087

It is clear from tables 16 and 19, that the bias and MSE results are getting smaller for smaller values of θ and σ , especially for $\theta = 1.2$ and $\sigma = 0.5$. By concentrating on the results related to $p = 0.5$, tables 17 and 18 show that Bayesian intervals are mostly anti-conservative. And this is true for different values of θ and σ , which their results are clearly presented in tables 20 and 21. Therefore, we can say that Bayesian intervals in case of noninformative priors are not comparable.

CHAPTER 4: DATA ANALYSIS

In this chapter, we shall present two data sets applications to clarify our inferential procedures, which have been explained in the previous units. For each data set we will apply both the classical and Bayesian statistics. In this chapter, we are interested in calculating the maximum likelihood for our parameters $\theta, \sigma, x_p,$ and x_{B_p}. Also, we will find the confidence intervals for quantiles and will include the MLE, Bootstrap, and Bayes intervals. The first data set is generated from scheme 5 and the second one is a real-life data set. Of course, we use same simulation techniques that we used in chapter 2 and chapter 3 for running both examples 1 and 2.

Example 1: Generated Data from Scheme 5

A progressively type II censored data has been generated from Scheme 5 ($n = 70, m = 50,$ and $R = (0^{20}, 5,5,4,3, 0^{6}, 1,1, 0^{10}, 1, 0^{7})$) based on generalized exponential distribution with parameters $\theta = 2$ and $\sigma = 1.2$. The data we got are

0.0851	0.1127	0.1368	0.2465	0.2644	0.2905	0.3342	0.3378	0.3773	0.4160
0.4866	0.5046	0.5559	0.5730	0.6359	0.6451	0.6716	0.6856	0.7168	0.7591
0.7955	0.8151	0.8157	0.8380	0.9440	0.9541	1.1145	1.1826	1.1964	1.3082
1.3109	1.4631	1.6386	1.7388	1.9520	2.0443	2.0617	2.1800	2.4350	2.4380
2.5894	2.6489	2.9877	3.3289	3.7082	3.8538	3.9058	3.9783	4.0314	4.1352

a- We used the same simulation which has been explained in chapter 2. We set $B = 500$. Our estimation results for our estimators are $\hat{\theta} = 1.596, \hat{\sigma} = 1.312, \hat{x}_{p_{0.1}} = 0.354,$ $\hat{x}_{p_{0.25}} = 0.714, \hat{x}_{p_{0.5}} = 1.369, \hat{x}_{p_{0.75}} = 2.365,$ and $\hat{x}_{p_{0.9}} = 3.609.$ It appears that estimation results are very close to the initial values of our parameters.

Now, we shall present the three types of confidence intervals (Asymptotic, percentile, and bootstrap-t confidence intervals) for our estimators $\hat{x}_p$, for both values of $\alpha = 0.1$ and $\alpha = 0.05$. The table below shows our results.

Table 22. Example 1 (a) confidence intervals for classical statistics methods.

C.I	$\alpha = 0.1$				
	$\hat{x}_{p_{0.1}}$	$\hat{x}_{p_{0.25}}$	$\hat{x}_{p_{0.5}}$	$\hat{x}_{p_{0.75}}$	$\hat{x}_{p_{0.9}}$
A I	(0.232, 0.475)	(0.549, 0.879)	(1.114, 1.624)	(1.906, 2.824)	(2.842, 4.377)
P R C	(0.247, 0.505)	(0.568, 0.915)	(0.915, 1.67)	(1.911, 2.908)	(2.842, 4.436)
Boot-t	(0.218, 0.479)	(0.535, 0.886)	(1.108, 1.676)	(1.939, 2.97)	(2.937, 4.641)
C.I	$\alpha = 0.05$				
	$\hat{x}_{p_{0.1}}$	$\hat{x}_{p_{0.25}}$	$\hat{x}_{p_{0.5}}$	$\hat{x}_{p_{0.75}}$	$\hat{x}_{p_{0.9}}$
A I	(0.209, 0.499)	(0.517, 0.91)	(1.065, 1.673)	(1.818, 2.912)	(2.695, 4.524)
P R C	(0.237, 0.552)	(0.528, 0.964)	(1.082, 1.706)	(1.847, 2.962)	(2.765, 4.584)
Boot-t	(0.192, 0.504)	(0.484, 0.935)	(1.084, 1.712)	(1.887, 3.083)	(2.852, 4.813)

For $\alpha = 0.05$, the confidence intervals seem to be wider than it is for $\alpha = 0.1$. As we increase α the interval width decreases. For both values of α , we notice that the length of these different types of confidence intervals are not very different from each other for each estimator $\hat{x}_p$, and this result doesn’t conflict with the results presented in table 5 and table

6 in chapter 2.

b- In this part, we apply Bayesian simulation which has been explained in chapter 3 to get Bayesian estimation and calculate Bayesian intervals (Bayes and HPD intervals). To do so, we also choose same parameters initial values, $\theta = 2$ and $\sigma = 1.2$, and we use the same generated data which has been generated from scheme 5 based on Generalized exponential distribution. Therefore, results of Bayes estimators are $\hat{\theta}_B = 1.17, \hat{\sigma}_B = 1.804, \hat{x}_{B_{p_{0.1}}} = 0.356$, $\hat{x}_{B_{p_{0.25}}} = 0.721, \hat{x}_{B_{p_{0.5}}} = 1.393, \hat{x}_{B_{p_{0.75}}} = 2.424$, and $\hat{x}_{B_{p_{0.9}}} = 3.72$. Also, these Bayes estimators are very close to our initial guesses. While the Bayesian intervals for each estimator $\hat{x}_{B_p}$ are indicated in the following table.

Table 23. Example 1 (b) confidence intervals for Bayesian statistics methods.

C.I			$\boldsymbol{\alpha = 0.1}$		
	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{p_{0.9}}}$
Bayes	(0.23, 0.499)	(0.546, 0.925)	(1.139, 1.681)	(1.981, 2.973)	(2.896, 4.647)
HPD	(0.236, 0.499)	(0.541, 0.896)	(1.165, 1.689)	(1.933, 2.826)	(2.87, 4.405)
C.I			$\boldsymbol{\alpha = 0.05}$		
	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{p_{0.9}}}$
Bayes	(0.208, 0.533)	(0.514, 1.011)	(1.09, 1.74)	(1.933, 3.091)	(2.896, 4.895)
HPD	(0.23, 0.533)	(0.498, 0.926)	(1.082, 1.689)	(1.912, 3.01)	(2.87, 4.728)

Table 23 shows that the length of both kinds of Bayesian confidence intervals are higher when $\alpha = 0.05$ than the lengths of these confidence intervals when $\alpha = 0.1$. It is notable that this result is consistent with our findings in chapter 3.

Comparing between classical way results and Bayesian results in this example, we find that $\hat{\theta} = 1.596$ and $\hat{\sigma} = 1.312$ estimators are much closer than Bayesian estimators $\hat{\theta}_B = 1.17, \hat{\sigma}_B = 1.804$, to the initial values $\theta = 2$ & $\sigma = 1.2$. On the other hand, by calculating the values of x_p's based on equation (12) in chapter 2, we get $x_{p_{0.1}} = 0.456, x_{p_{0.25}} = 0.832, x_{p_{0.5}} = 1.474, x_{p_{0.75}} = 2.412$ and $x_{p_{0.9}} = 3.564$. In general, values of $\hat{x}_p$'s are like $\hat{x}_{B_p}$'s. However, it should be said that $\hat{x}_{B_p}$'s values are much closer to x_p's values than $\hat{x}_p$'s values, except for $p = 0.9$, opposite is true. About confidence intervals, generally we can say that lengths of HPD confidence intervals for both values of α , are smaller than Bootstrap-t confidence intervals. In contrast, the lengths of percentile confidence intervals are smaller than Bayes confidence intervals.

Example 2: Real Data from (Lawless, 2003)

In this example, we take a real data which has been taken from (Lawless, 2003). The data represents the lifetime of automobile brake pads for 98 cars, where the number of miles or kilometers are driven, is known to be the pads lifetime. For our purposes we only present the lifetime t_i (in km) data which is left truncated:

18.6	20.8	24.8	27.8	31.8	32.9	33.6	34.3	37.2	38.7	38.8	39.3	42.4	42.4	42.4	43.4
43.8	44.1	44.2	44.8	45.2	46.3	46.7	46.8	47.4	49.2	49.2	49.8	50.5	50.8	51.5	52.0
53.9	54.0	54.0	54.9	55.0	55.9	56.2	56.2	58.4	59.3	59.4	60.3	61.4	61.9	63.7	64.0

65.0 65.1 65.5 67.6 68.8 68.8 68.9 68.9 69.0 69.0 69.6 72.2 72.8 73.8 74.7 74.8 75.2 77.2 77.6 78.1 78.7 79.4 79.5 81.6 82.6 83.0 83.0 83.6 83.8 86.7 87.6 88.0 89.1 89.5 92.5 92.6 95.7 100.6 101.2 101.9 103.6 105.6 105.6 107.8 110.0 123.5 124.5 124.6 143.6 165.5

Before explaining our work, we shall mention that dealing with our data has been inspired by (Pradhan & Kundu, 2009) and (Asgharzadeh, 2009) . Where (Pradhan & Kundu, 2009) obtained the maximum likelihood estimators of the generalized exponential distribution based on progressive censoring. They used EM algorithm in their application. For application, they used only one real data as an example. First, they tested the complete data set if fitted the generalized exponential distribution by applying the Kolmogorov-Smirnov distance. After that, they took three different samples from the data set with $m = 12$ for the application. And (Asgharzadeh, 2009) derived the scale parameter of the generalized exponential distribution and approximated the likelihood function by providing a simple method of deriving an explicit estimator. They used a Monte Carlo simulation to find that estimator. They applied two examples for applications. In the first example, to obtain the MLE of their parameter λ, they considered $n = m = 23,$ & $R = (0^{23})$ since the whole data set contains 23 observations. In example two, they generated two different progressively type II censored samples from GE distribution to obtain the approximated MLE and the confidence intervals for parameter σ.

First, we shall present a summary statistic which display in the following table:

Table 24: Descriptive statistics for the real data.

Mean	Standard deviation	Minimum	Median	Maximum
67.73	26.73	18.60	65.05	165.50

And the histogram plot for the data is shown in the figure below:

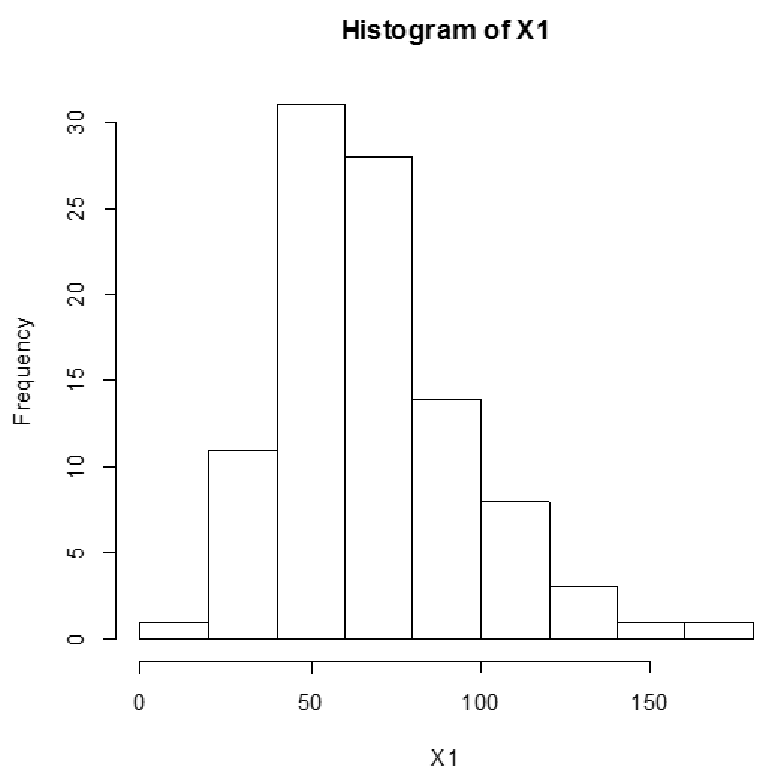

Figure 17. Histogram plot for the real data.

We must test if our data set fits the GE distribution. We use Kolmogorov-Smirnov to do that. We apply the test by using R software, and the Kolmogorov-Smirnov distance is $d = 0.056$ and $p - value = 0.915$. Where d is the test statistics represents the maximum absolute distance between the expected and the observed distribution. The blot below explains that distance.

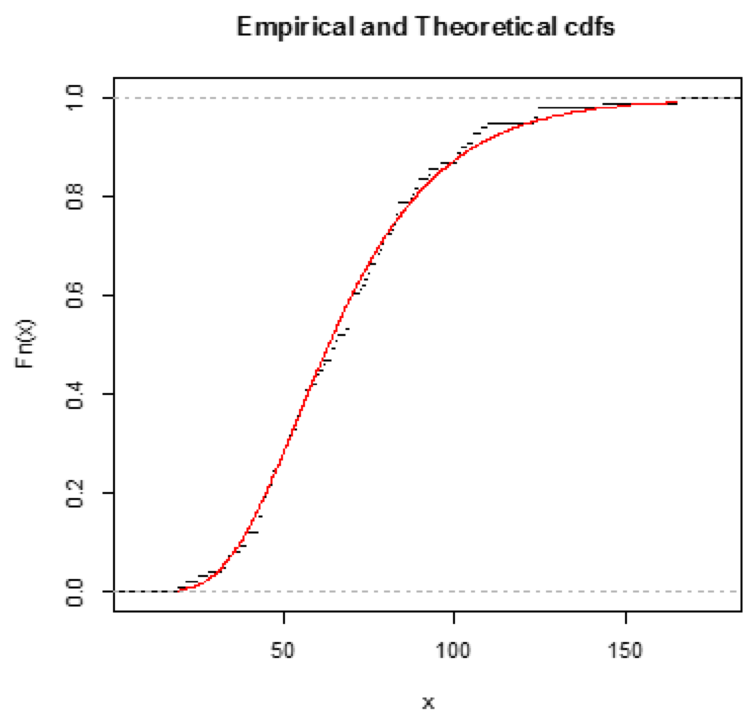

Figure 18. Plot between the expected and the observed distribution.

Our hypothesis test is H_0 : Data come from GE distribution vs H_1 : Data doesn't come from GE distribution. Now, we need to find the critical values at %95. By using one-sample-Kolmogorov Smirnov-table, we find that $d_{0.05} = 1.36$. Then, $d_{0.05} = \frac{1.36}{\sqrt{n}} = 0.137$, where $n = 98$. Since $d = 0.056 < d_{0.05} = 0.137$, then we conclude that H_0 can't be rejected for both values of α. Therefore, our data follow the GE distribution. To run a complete data set simulation, setting $n = m = 98$, and $\mathrm{R} = (0^{98})$. Our estimation results are $\hat{\theta} = 10.364, \hat{\sigma} = 22.983, \hat{x}_{p_{0.1}} = 37.08,\ \hat{x}_{p_{0.25}} = 47.756, \hat{x}_{p_{0.5}} = 62.931, \hat{x}_{p_{0.75}} = 82.696$, and $\hat{x}_{p_{0.9}} = 105.58$.

We have generated progressively type II censored samples by using two different schemes. We set $m = 50$ and $m = 70$, and the censoring schemes below are applied for both ways (classical and Bayesian statistics):

Censoring Scheme 1: $\mathrm{R1} = (0^{49}, 48)$

The progressively censored type II sample 1:

18.6 20.8 24.8 27.8 31.8 32.9 33.6 34.3 37.2 38.7 38.8 39.3 42.4 42.4 42.4 43.4
43.8 44.1 44.2 44.8 45.2 46.3 46.7 46.8 47.4 49.2 49.2 49.8 50.5 50.8 51.5 52.0
53.9 54.0 54.0 54.9 55.0 55.9 56.2 56.2 58.4 59.3 59.4 60.3 61.4 61.9 63.7 64.0
65.0 65.1

Censoring Scheme 2: $\mathrm{R2} = (0^{69}, 28)$

The progressively censored type II sample 2:

18.6 20.8 24.8 27.8 31.8 32.9 33.6 34.3 37.2 38.7 38.8 39.3 42.4 42.4 42.4 43.4

43.8 44.1 44.2 44.8 45.2 46.3 46.7 46.8 47.4 49.2 49.2 49.8 50.5 50.8 51.5 52.0 53.9 54.0 54.0 54.9 55.0 55.9 56.2 56.2 58.4 59.3 59.4 60.3 61.4 61.9 63.7 64.0 65.0 65.1 65.5 67.6 68.8 68.8 68.9 68.9 69.0 69.0 69.6 72.2 72.8 73.8 74.7 74.8 75.2 77.2 77.6 78.1 78.7 79.4

a- The MLEs results of $\left(\theta, \sigma, x_{p_{0.1}}, x_{p_{0.25}}, x_{p_{0.5}}, x_{p_{0.75}}, x_{p_{0.9}}\right)$ for all censoring schemes 1 and 2 are presented in the following table:

Table 25. MLEs results of $\left(\theta, \sigma, x_{p_{0.1}}, x_{p_{0.25}}, x_{p_{0.5}}, x_{p_{0.75}}, x_{p_{0.9}}\right)$ parameters.

Scheme	$\hat{\boldsymbol{\theta}}$	$\hat{\boldsymbol{\sigma}}$	$\hat{\boldsymbol{x}}_{\boldsymbol{p}_{0.1}}$	$\hat{\boldsymbol{x}}_{\boldsymbol{p}_{0.25}}$	$\hat{\boldsymbol{x}}_{\boldsymbol{p}_{0.5}}$	$\hat{\boldsymbol{x}}_{\boldsymbol{p}_{0.75}}$	$\hat{\boldsymbol{x}}_{\boldsymbol{p}_{0.9}}$
1	8.716	25.12	36.683	48.155	64.588	86.1	111.071
2	9.381	24.149	36.824	47.937	63.8	84.518	108.541

MLE results are increased while p^{th} quantiles are increased for $\hat{x}_p$ estimators. Comparing between the MLE results presented in table 25 for censored samples with the MLEs of the complete data, it seems that the MLEs of scheme 2 are much closer to the complete data MLEs. Also, the MLE results of scheme 2 are lower than the MLE results of scheme 1. On the other hand, it doesn't appear that MLEs for censored data are lower than the MLEs of the complete data, but on the contrary except for

$\hat{\theta}$ and $\hat{x}_{p_{0.1}}$. Now, tables 25 and 26 represent the confidence intervals of our estimators $\hat{x}_p$ in this case:

Table 26. Scheme 1 Confidence Intervals of the classical statistics methods.

C.I	$\boldsymbol{\alpha = 0.1}$				
	$\hat{x}_{p_{0.1}}$	$\hat{x}_{p_{0.25}}$	$\hat{x}_{p_{0.5}}$	$\hat{x}_{p_{0.75}}$	$\hat{x}_{p_{0.9}}$
A I	(32.796, 40.571)	(44.206, 52.105)	(59.016, 70.16)	(76.905, 95.295)	(96.975, 125.167)
P R C	(33.023, 40.48)	(44.245, 52.144)	(60.449, 68.757)	(81.173, 90.744)	(105.167, 116.5)
Boot-t	(32.909, 40.447)	(44.186, 52.098)	(60.271, 68.815)	(81.123, 91.04)	(105.265, 116.912)
C.I	$\boldsymbol{\alpha = 0.05}$				
	$\hat{x}_{p_{0.1}}$	$\hat{x}_{p_{0.25}}$	$\hat{x}_{p_{0.5}}$	$\hat{x}_{p_{0.75}}$	$\hat{x}_{p_{0.9}}$
A I	(32.051, 41.315)	(43.449, 52.861)	(57.949, 71.228)	(75.143, 97.056)	(94.274, 127.867)
P R C	(32.501, 41.427)	(43.693, 53.011)	(59.516, 69.547)	(80.209, 91.548)	(104.239, 117.179)
Boot-t	(31.949, 41.033)	(43.38, 52.846)	(59.481, 69.838)	(80.013, 92.132)	(104.371, 117.931)

Table 27. Scheme 2 Confidence Intervals of the classical statistics methods.

C.I	$\boldsymbol{\alpha = 0.1}$				
	$\hat{x}_{p_{0.1}}$	$\hat{x}_{p_{0.25}}$	$\hat{x}_{p_{0.5}}$	$\hat{x}_{p_{0.75}}$	$\hat{x}_{p_{0.9}}$
A I	(33.077, 40.572)	(44.194, 51.679)	(59.037, 68.562)	(77.171, 91.866)	(97.509, 119.573)
P R C	(32.41, 39.353)	(43.111, 50.212)	(58.429, 65.592)	(78.316, 86.066)	(101.429, 109.771)
Boot-t	(34.193, 41.55)	(45.604, 52.878)	(61.915, 69.639)	(82.757, 91.78)	(106.976, 117.166)
C.I	$\boldsymbol{\alpha = 0.05}$				
	$\hat{x}_{p_{0.1}}$	$\hat{x}_{p_{0.25}}$	$\hat{x}_{p_{0.5}}$	$\hat{x}_{p_{0.75}}$	$\hat{x}_{p_{0.9}}$
A I	(32.359, 41.29)	(43.477, 52.397)	(58.125, 69.474)	(75.763, 93.273)	(95.396, 121.686)
P R C	(31.726, 40.4)	(42.212, 51.319)	(57.354, 67.068)	(77.202, 87.292)	(100.284, 110.825)
Boot-t	(33.137, 42.363)	(44.507, 53.909)	(60.372, 70.699)	(81.363, 92.787)	(105.721, 118.196)

For all schemes in this example, the confidence intervals lengths when $\alpha = 0.1$ are smaller than the lengths of same confidence intervals when $\alpha = 0.05$.

b- In this part, the same censoring schemes and its corresponding samples are applied to obtain Bayesian estimates and their confidence intervals. we have been considered the noninformative prior distribution in this case. The Bayesian estimates results of $\left(\theta_B, \sigma_B, x_{B_{p0.1}}, x_{B_{p0.25}}, x_{B_{p0.5}}, x_{B_{p0.75}}, x_{B_{p0.9}}\right)$ for schemes 1 and 2 are presented in the following table:

Table 28. Bayesian estimates of $\left(\theta_B, \sigma_B, x_{B_{p0.1}}, x_{B_{p0.25}}, x_{B_{p0.5}}, x_{B_{p0.75}}, x_{B_{p0.9}}\right)$ parameters.

Scheme	$\hat{\theta}_B$	$\hat{\sigma}_B$	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{0.9}}$
1	4.064	34.968	28.973	43.006	64.335	93.342	127.672
2	3.866	37.948	30.402	45.509	68.557	99.977	137.205

It is clear that $\hat{x}_{B_p}$ estimators are increased while p^{th} quantiles are increased for both schemes. Now, tables below present Bayesian confidence intervals for all three schemes.

Table 29. Scheme 1 Bayesian Confidence Intervals.

C.I	$\alpha = 0.1$				
	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{p_{0.9}}}$
Bayes	(26.614, 33.72)	(40.352, 46.904)	(59.682, 72.323)	(85.32, 106.777)	(115.518, 147.75)
HPD	(27.739, 32.734)	(40.322, 46.423)	(58.889, 69.092)	(84.895, 101.759)	(115.518, 140.974)
C.I	$\alpha = 0.05$				
	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{p_{0.9}}}$
Bayes	(25.734, 32.842)	(38.72, 47.615)	(58.539, 73.064)	(85.32, 109.157)	(115.82, 152.223)
HPD	(26.495, 32.835)	(39.579, 47.615)	(59.682, 73.215)	(84.895, 107.986)	(115.732, 150.275)

Table 30. Scheme 2 Bayesian Confidence Intervals.

C.I			**$\alpha = 0.1$**		
	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{p_{0.9}}}$
Bayes	(27.41, 32.621)	(41.316, 47.162)	(63.741, 71.348)	(93.531, 105.885)	(128.543, 146.974)
HPD	(27.636, 32.328)	(41.801, 47.517)	(63.59, 70.956)	(93.299, 102.829)	(128.485, 141.178)
C.I			**$\alpha = 0.05$**		
	$\hat{x}_{B_{p_{0.1}}}$	$\hat{x}_{B_{p_{0.25}}}$	$\hat{x}_{B_{p_{0.5}}}$	$\hat{x}_{B_{p_{0.75}}}$	$\hat{x}_{B_{p_{0.9}}}$
Bayes	(27.113, 35.5)	(41.253, 53.749)	(62.447, 78.515)	(93.61, 113.681)	(129.815, 155.301)
HPD	(26.253, 32.352)	(40.449, 47.951)	(61.11, 71.348)	(92.7, 106.945)	(128.485, 147.427)

As in part (a), the lengths of confidence intervals are smaller when $\alpha = 0.1$ than lengths of confidence intervals when $\alpha = 0.05$, and that satisfied for all schemes. Comparing between part (a) and part (b), we find that the corresponding parameters estimates results for MLE and Bayesian are close to each other, but $\hat{x}_{B_p}$'s Bayesian estimators are much higher than $\hat{x}_p$'s estimators for $p = 0.75$ and $p = 0.9$.

CHAPTER 5: COMPARISON, CONCLUSION, AND SUGGESTIONS FOR FURTHER STUDIES

In conclusion, this study is about studying the inference procedures for the Generalized exponential quantiles based on the progressively censored type II data. The maximum likelihood estimators have been derived for both shape and scale parameters of the GE distribution. After that, we derived the maximum likelihood estimator for the p^{th} quantile. For calculating the approximate confidence intervals for θ and σ for large m, an observed Fisher information matrix is needed to do our purpose calculations. As we mentioned before that we are interested in the maximum likelihood of p^{th} quantiles. Therefore, we find the asymptotic variance of the MLE of the p^{th} quantiles. After that, we used the delta method for multivariate to compute the asymptotic distribution. Then, we compute three kinds of confidence intervals, the asymptotic confidence interval, bootstrap-t confidence interval, and percentile confidence interval. We chose different censoring schemes with different choices of n and m. And we applied a simulation study to calculate bias and MSE for parameter (x_p). Also, we calculated the average lengths and coverage probability for the three types of confidence intervals. Another statistical method has been used in this study, which is known as Bayesian statistics. In this case, we derived the posterior distribution, choosing prior distributions for parameters θ and σ to be Gamma and Inverted Gamma distributions, respectively. Another case of Bayesian application is using noninformative prior distributions for parameters θ and σ.We find out the conditional distribution for both $f(\theta|x)$ and $f(\sigma|x)$ for the both cases. Then, we compute two kinds of Bayesian intervals, equal tail interval and the highest posterior density interval. For simulation study we applied importance sampling to calculate the bias and MSE for Bayesian estimators. And again, we calculate the average lengths and coverage probability for each type of intervals, using same censoring schemes used in classical way.

Comparison

In this section, we compared between results presented in chapter 2 and results presented in chapter 3. In other words, this comparison will be between the classical statistics results in chapter 2 and Bayesian statistics results in chapter 3. We compared between table 4, table 10 and table 16, these tables present the bias and MSE for classical statistics and both cases of Bayesian statistics respectively. It is obvious that the higher pth quantiles, the lower the bias values, and this is expressed table 4 which represents the results of the classical statistics. In contrast, table 9 and table 12 show that the higher p^{th} quantiles, the higher the bias values, and these tables represents Bayesian work, but this is not true for schemes 1 and 2 in table 9. Table 9 shows that the smallest bias for schemes 1 and 2 is when $p = 0.5$. It is known that if the bias is zero then we can say that we have an unbiased estimator. Therefore, we can see that bias is mostly insignificant for both estimators $\hat{x}_p$'s and $\hat{x}_{B_p}$'s. Now, we shall describe the MSE results. In our study, the closest values of MSE to zero are when $p = 0.1$, and these values move away from zero as p values increase. Note that this is applied to both classical and Bayesian cases and applies to all censoring schemes. But, table 16 shows that MSE results when $p = 0.9$ are quite high. Generally, MSE results for both classical and Bayesian are similar, but larger samples, such as schemes 4,5, … etc, MSE results for the first case of Bayesian statistics are better than the classical statistics.

In general, the bias and MSE results for classical and Bayesian method, is that the larger values of n and m in each censoring scheme, the lower the values of bias and MSE.

It is important to compare between classical method confidence intervals which present in tables 5 and 6, and Bayesian methods intervals which present in tables 11 and 12. Note that Bayesian intervals in the case of noninformative prior distributions are not comparable. Before comparing, we shall mention that coverage probabilities must attain the nominal error rate, which are $\alpha = 0.1$ and $\alpha = 0.05$. Whereas, an interval is said to be reached the nominal error rate, if the observed coverage probability is close to the nominal one. For example, when $\alpha = 0.1$ or $\alpha = 0.05$, the interval is said to be valid if the observed coverage probability result is between 0.08 and 0.12, or between 0.04 and 0.06, respectively. It is obvious that some coverage probabilities results presented in table 5 didn't attain the nominal error rate and this is mostly obvious for percentile and bootstrap-t confidence intervals for schemes 1, 2, 3 and 7, especially when $p = 0.75$ and $p = 0.9$. And this is true only for the percentile confidence interval for scheme 2, when $p = 0.1$. Also, the coverage probability of the asymptotic confidence interval for scheme 1 when $p = 0.75$ and $p = 0.9$ show some problems. On the other hand, the coverage probabilities of the both types of Bayesian intervals show some problems for all schemes and all p^{th} quantiles, and clearly this is true especially for the HPD interval. According to the coverage probability results, the comparison between the intervals average widths is only for those intervals which attain the nominal error rate. Generally, Bayesian intervals average widths are narrower than classical confidence intervals average widths. For scheme 1, the average widths of the HPD intervals are not comparable because they didn't attain the nominal error rate. But equal tail interval shows better average lengths than the classical confidence intervals when $p = 0.1$. On the other hand, when $p = 0.25$, asymptotic and percentile confidence intervals show better average lengths than equal tail interval with a very small difference. When $p = 0.9$, the only intervals can be compared together are bootstrap-t and equal tail intervals, and equal tail interval has better average length. For scheme 2, the equal tail intervals have the best average lengths all over the p^{th} quantiles compared with the classical confidence intervals, except when $p = 0.5$, percentile confidence interval is the best. And this is true for schemes 4 and 6 for all over the p^{th} quantiles. For scheme 3, the only situations that we can compare are when $p = 0.1$, $p = 0.25$ and $p = 0.9$, and for all these situations equal tail intervals show the best average lengths. The same remark for scheme 7, but this is true only when $p = 0.75$ and $p = 0.9$.

The coverage probability of the intervals for $\alpha = 0.05$, which are presented in tables 6 and 12 for both classical and Bayesian methods, respectively, have more problems than the coverage probability of intervals when $\alpha = 0.1$. In table 6, coverage probabilities don't reach the nominal coverage probability especially for percentile confidence interval. And this is true for all p^{th} quantiles and all schemes, except for $p = 0.5$ and scheme 4. For Bayesian intervals, the coverage probabilities mostly don't attain the nominal error rate, especially for the HPD intervals. The only case that HPD intervals attain the nominal error rate is for scheme 4 when $p = 0.75$ and $p = 0.9$. When comparing between the two methods, classical and Bayesian intervals, equal tail interval has the best average width in scheme 1 when $p = 0.1$, but when $p = 0.25$ the percentile confidence interval has the best average width with a very small difference. Scheme 3 shows the same comparison, equal tail interval has the best average width, but for both $p = 0.1$ and $p = 0.25$. For scheme 2, the equal tail interval has the best average width when $p = 0.1$, but when $p = 0.25$ and $p = 0.5$, asymptotic confidence interval and percentile confidence interval are the best, respectively. In scheme 4, there is a clear difference between Bayesian intervals average widths and classical confidence intervals average widths, where Bayesian intervals average widths are smaller.

It is remarkable that all comparable classical confidence intervals and Bayesian intervals average widths are decreased when n and m increase in each censoring scheme. In contrast, the average lengths are increased when p^{th} quantiles increase.

We have applied the classical and Bayesian methods again, but we chose other three different values of θ and σ with only one p^{th} quantile value; i.e $p = 0.5$. In this case, bias and MSE results have been decreased when θ and σ were decreased. For the coverage probability, classical confidence intervals mostly attain the nominal coverage probability better than when $\theta = 2$ and $\sigma = 1.2$ for both cases $\alpha = 0.1$ or $\alpha = 0.05$, except for the coverage probabilities of the classical confidence intervals when $\alpha = 0.05$. But this not true for both cases of Bayesian intervals since Bayesian intervals' coverage probability don't attain the nominal error rates when $p = 0.5$. Also, the average widths of the comparable intervals for smaller values of θ & σ are smaller than their counterpart when $\theta = 2$ and $\sigma = 1.2$.

In conclusion, we need to find out which inference procedures is better in a case of progressively censored data. It is clear from bias and MSE results that the bias for estimates of x_p's is closer to zero than the bias for estimates of x_{B_p}'s. Which means that x_p's estimates are indicated to be more unbiased than x_{B_p}'s estimates. Of course, for smaller parameters values, we get smaller bias and MSE.

For those intervals which attain the nominal error rate (0.1 or 0.05), Bayesian intervals have the best average widths. For classical methods confidence intervals, percentile confidence interval generally has the best average lengths compared with asymptotic and bootstrap-t confidence intervals. And this is true for $\alpha = 0.1$ or $\alpha = 0.05$. Therefore, a general conclusion is that equal tail intervals are the best regardless of the parameter's values.

Suggestion for Further Studies

It is notable that some of the coverage probability of the classical confidence intervals and Bayesian intervals don't attain the nominal coverage error, especially for Bayesian intervals of the noninformative prior distributions. To solve this problem, we suggest using another method of simulation, namely Markov Chain Monte Carlo (MCMC), especially Metropolis-Hasting within Gibbs sampler. This can be done as a further study that will compare between importance sampling and the Markov Chain Monte Carlo (MCMC). Also, we can study other types of censoring and compare it with our type of censoring. Another different loss function may be able to study it as further studies. Another suggestion for further studies is studying the effectiveness of some bias reduction techniques like the bootstrap, Jackknife or asymptotic corrections for bias. On the other hand, the same procedures can be applied for other continuous distribution and use another real data set for application.

REFERENCES

Aldrich, J. (1997). R. A. Fisher and The Making of Maximum Likelihood 1912-1922. *Statistical Science*, 162-176.

Asgharzadeh, A. (2009). Approximate MLE for The Scaled Generalized Exponential Distribution Under Progressive Type-II Censoring. *Journal of The Korean Statistical Society*, 223-229.

Baklizi, A. (2008). Confidence Intervals for The Quantiles of The Weibull Distribution Based on Record Values. *Journal of Applied Statistical Science*, 119-130.

Baklizi, A. (2009). Confidence Intervals for Quantiles Using Record Values from The Generalized Exponential Distribution. *Journal of Statistical Theory and Applications*, 373-381.

Balakrishnan, N. (2007). Progressive Censoring Methodology: An Appraisal (with Discussions). *Test*, 211-296.

Balakrishnan, N., & Cramer, E. (2014). *The Art of Progressive Censoring.* New York: Birkhäuser.

Balakrishnan, N., & Li, T. (2006). Confidence Intervals for Quantiles and Tolerance Intervals Based on Ordered Ranked Set Samples. *The Institute of Statistical Mathematics*, 78-82.

Balakrishnan, N., & Sandhu, R. A. (1995). A Simple Simulational Algorithm for Generating Progressive Type II Censored Samples. *The American Statistician*, 229-230.

Bang, H., & Zhao, H. (2012). Average Cost-Effectiveness Ratio with Censored Data. *Journal of Biopharmaceutical Statistics*, 401-415.

Berger, J. O. (1980). *Statistical Decision Theory*. New York: Springer Science+Business Media.

Bolstad, W. M. (2007). *Introduction to Bayesian Statistics*. New Jersey: Wiley.

Casella, G., & Berger, R. L. (2002). *Statistical Inference*. Pacific Grove: DUXBURY THOMSON LEARNING.

Chen, M.-H., & Shao, Q.-M. (1999). Monte Carlo Estimation of Bayesian Credible and HPD Intervals. *Journal of Computational and Graphical Statistics*, 69-92.

Costa, Veber, Frenandes, & Wilson. (2017). Bayesian Estimation of Extreme Flood Quantiles Using A Rainfall Runoff Model and A Stochastic Daily Rainfall Generator. *Journal of Hydrology*, 137-154.

Davidson, R., & MacKinnon, J. (2001). Bootstrap tests: How many bootstraps? *Queen's Economics Department*.

De San Roman, A. G., & Rebollo-Sanz, Y. F. (2018). An Estimation of Worker and Firm Effects with Censored Data. *Bulletin of Economic Research*, 459-482.

Edgeworth, F. (1886). *Progressive Means*. JRSS.

Edgeworth, F. (1893). *Review of Fisher's Mathematical Investigations*. EJ.

Efron, B. (1979a). Bootstrap Methods: Another Look at The Jackknife. *Ann*, 1-26.

El-Sagheer, R. M. (2016). Bayesian Prediction Based on General Progressive Censored Data from Generalized Pareto Distribution. *Journal of Statistics Applications & Probabilities*, 43-51.

Eubank, R. L. (1984). Quantiles. *Naval Research*.

Fienberg, S. (2006). When Did Bayesian Inference Become Bayesian? *Bayesian Analysis*, 1-40.

Galton, F. (1889). *Natural Inheritance.* New York: MacMillan.

Gelman, A., Carlin, J., Stern, H., Dunson, D., Vehtari, A., & Rubin, D. (2013). *Bayesian Data Analysis.* Chapman and Hall/CRC.

Gelman, A., Simpson, D., & Betancourt, M. (2017). The Prior Can Often Only Be Understood in The Context of The Likelihood. *Entropy*, 1-13.

Gugushvili, S. (2014, October 20). Maximum Likelihood Estimation: Multidimensional Case. Leiden, Netherlands.

Gupta, R. D., & Kundu, D. (1999). Generalized Exponential Distributions. *Austral. & New Zealand J. Statist.*, 173-188.

Gupta, R. D., & Kundu, D. (2002). Generalized Exponential Distribution: Statistical Inferences. *J Statist Theory Appl*, 101-118.

Gupta, R. D., & Kundu, D. (2007). Generalized Exponential Distribution: Existing Results and Some Recent Developments. *J Statist Plan Inference*, 3537-3547.

Gupta, R. D., & Kundu, D. (2008). Generalized Exponential Distribution: Bayesian Estimations. *Computational Statistics & Data Analysis*, 1873-1883.

Hasan, M. R., & Baizid, A. R. (2016). Bayesian Estimation Under Different Loss Function Using Gamma Prior for The Case of Exponential Distribution. *Journal of Scientific Research*, 67-78.

Held, L., & Sabanés Bové, D. (2014). *Applied Statistical Inference: Likelihood and Bayes.* New York: Springer.

Jones, M. C. (1992). Estimating Densities, Quantiles, Quantile Density and Density Quantiles. *Ann. Inst. Statist. Math*, 721-727.

Khan, M. (1987). A Generalized Exponential Distribution . *Biom.J*, 121-127.

Kim, C., & Han, K. (2015). Bayesian Estimation of Generalized Exponential Distribution Under Progressive First Failure Censored Sample. *Applied Mathematical Sciences*, 2037-2047.

Kim, C., & Song, S. (2010). Bayesian Estimation of The Parameters of The Generalized Exponential Distribution from Doubly Censored Samples. *Stat Papers*, 583-597.

Kim, N. (2013). Approximate MLE for The Scale Parameter of The Generalized Exponential Distribution Under Random Censoring. *Journal of The Korean Statistical Society*, 119-131.

Krishna, H., & Kumar, K. (2011). Reliability Estimation in Lindley Distribution with Progressively Type II Right Censored Sample. *Mathematics and Computers in Simulation*, 281-294.

Kundu, D., & Pradhan, B. (2009). Estimating The Parameters of The Generalized Exponential Distribution in Presence of Hybrid Censoring. *Naval Res Logist*, 687-698.

Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data.* New Jersey: John Wiley & Sons.

Lee, P. M. (1992). *Bayesian Statistics: An Introduction.* Wiley.

Lewis, N. (2018). Combining Independent Bayesian Posteriors into A Confidence Distribution, with Application to Estimate Sensitivity. *Journal of Statistical Planning and Inference*, 80-92.

Li, X. (2011). Interval Estimation for The Quantile of A Two-Parameter Exponential Distribution. *International Journal of Innovative Management, Information & Production*, 78-82.

Mohi El-Din, M., Amein, M., Shafay, A., & Mohamed, S. (2016). Estimation of Generalized Exponential Distribution Based on an Adaptive Progressively Type II Censored Sample. *Journal of Statistical Computation and Simulation*, 1292-1304.

Mohie El-Din, M. M., & Shafay, A. R. (2013). One- and Two-Sample Bayesian Prediction Intervals Based on Progressively Type-II Censored Data. *Statist Pap*, 287-307.

Mosteller, F. (1946). On some useful "inefficient" statistics. *Annals of Mathematical Statistics*, 377-408.

Ng, H. K., Kundu, D., & Chan, P. (2009). Statistical Analysis of Exponential Lifetimes Under an Adaptive Type-II Progressive Censoring Scheme. *Naval Res Logist*, 687-698.

Ogawa, J. (1951). Contributions to the theory of systematic statistics. I. *Osaka Math. J*, 175-213.

Panichkitkosolkul, W., & Saothayanun, L. (2012). Bootstrap Confidence Intervals for The Process Capability Index Under Half-Logistic Distribution. *Journal of Science and Technology*, 272-281.

Parzen, E. (1979a). *A Density-quantile function perspective on robust estimation.* New York: Academic Press.

Pearson, G. (1920). Factors Controlling the Distribution of Forest Types, Part I. *ECOLOGY SOCIETY of AMERICA*, 139-159.

Pfeiffer, P. (1990). Some Properties of the Quantile Function . *Probability for Applications*, 266-271.

Pradhan, B., & Kundu, D. (2009). On Progressively Censored Generalized Exponential Distribution. *TEST*.

Raqab, M. Z., & Madi, M. T. (2005). Bayesian Inference for The Generalized Exponential Distribution. *Journal of Statistical Computation and Simulation*, 841-852.

Sarhan, A. M., & Abuammoh, A. (2008). Statistical Inference Using Progressively Type-II Censored Data with Random Scheme. *International Mathematical Forum*, 1713-1725.

Shaw, S. (2018). *Topics in Bayesian statistics.* Retrieved from MA40189: Topics in Bayesian statistics: http://people.bath.ac.uk/masss/ma40189.html

Sheppard, W. (1899). A Method for Extending the Accuracy of Certain Mathematical Tables. *London Mathematical Society*, 432-448.

Simon, J. (1969). *Basic Research Methods in Social Science.* New York: Random House.

Strawderman, R. L., Parzen, M., & Wells, M. T. (1997). Accurate Confidence Limits for Quantiles Under Random Censoring. *International Biometric Society*, 1399-1415.

Tukey, J. (1977). *Exploratory Data Analysis.* Massachusetts: Addison-Wesley.

Ye, Z. S., Chan, P. S., Xie, M., & Ng, H. K. (2014). Statistical Inference for The Extreme Value Distribution Under Adaptive Type-II Progressive Censoring Schemes. *Journal of Statistical Coputation and Simulation*, 1099-1114.