T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

STATISTICAL INFERENCE FOR SOME DISTRIBUTIONS BASED ON RECORD VALUES

FATİH KIZILASLAN

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
DEPARTMENT OF MATHEMATICS

GEBZE
2015

T.R.
GEBZE TECHNICAL UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

STATISTICAL INFERENCE FOR SOME DISTRIBUTIONS BASED ON RECORD VALUES

FATİH KIZILASLAN

A THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY
DEPARTMENT OF MATHEMATICS

THESIS SUPERVISOR
ASSOC. PROF. DR. COŞKUN YAKAR
II. THESIS SUPERVISOR
ASSOC. PROF. DR. MUSTAFA NADAR

GEBZE
2015

T.C.
GEBZE TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

REKOR DEĞERLER KULLANILARAK BAZI DAĞILIMLAR İÇİN İSTATİSTİKSEL ÇIKARIM

FATİH KIZILASLAN
DOKTORA TEZİ
MATEMATİK ANABİLİM DALI

DANIŞMANI
DOÇ. DR. COŞKUN YAKAR
II. DANIŞMANI
DOÇ. DR. MUSTAFA NADAR

GEBZE
2015

GEBZE TEKNİK ÜNİVERSİTESİ	DOKTORA JÜRİ ONAY FORMU

GTÜ Fen Bilimleri Enstitüsü Yönetim Kurulu'nun 23/06/2014 tarih ve 2014/37 sayılı kararıyla oluşturulan jüri tarafından 21/01/2015 tarihinde tez savunma sınavı yapılan Fatih KIZILASLAN'ın tez çalışması Matematik Anabilim Dalında DOKTORA tezi olarak kabul edilmiştir.

JÜRİ

ÜYE
(TEZ DANIŞMANI) : Doç. Dr. Coşkun YAKAR

ÜYE
(II. TEZ DANIŞMANI) : Doç. Dr. Mustafa NADAR

ÜYE : Prof. Dr. Tahir Aliyev AZEROĞLU

ÜYE : Prof. Dr. Fatih TAŞÇI

ÜYE : Prof. Dr. Mansur İSMAİLOV

ONAY

Gebze Teknik Üniversitesi Fen Bilimleri Enstitüsü Yönetim Kurulu'nun

......../........./......... tarih ve ..../......... sayılı kararı.

İMZA/MÜHÜR

SUMMARY

In this dissertation, the methods of statistical inference of record values are considered for the Burr Type XII, the generalized exponential and the Kumaraswamy distributions. This includes estimates of the distribution parameters, the stress-strength reliability and prediction of the future record values. Both frequentist and Bayesian techniques, namely maximum likelihood, uniformly minimum variance unbiased, Bayesian and empirical Bayesian estimates are used for the unknown parameters and stress-strength reliability of the distributions. All these estimates are obtained based on record values or based on record values with their corresponding inter-record times. Furthermore, the asymptotic confidence interval using Fisher information or observed information matrix, Bayesian credible, highest probability density credible intervals and the exact confidence interval, when it is available, are constructed. In order to draw a statistical inference a simulation study is carried out for each of these distributions. The performance of all these estimates are compared by using the Monte Carlo simulation. A numerical findings of the estimates are presented for the generated data in every case and a real life data when it is available.

Keywords: Record Values, Stress-Strength Reliability, Bayesian Estimation, Prediction, Lindley's Approximation, Markov Chain Monte Carlo (MCMC) Method.

ÖZET

Bu tezde, Burr XII, genelleştirilmiş üstel ve Kumaraswamy dağılımları için rekor değerlerin istatistiksel çıkarım methodları ele alınmıştır. Bu methodlar dağılımların parametreleri, güvenilirlik ve gelecek rekor değerlerin kestrimi tahminlerini içermektedir. Klasik ve Bayes tekniklerinden en çok olabilirlik, düzgün en küçük varyanslı yansız, Bayes ve empirik Bayes tahmin edicileri kullanılmıştır. Tüm bu tahmin ediciler rekor değerler veya rekor değerler ve onların rekor zamanları kullanılarak elde edilmiştir. Ayrıca, Fisher bilgisi veya gözlemlenmiş bilgi matrisi kullanılarak asimptotik güven aralığı, Bayes güven aralıkları ve mümkün olduğunda kesin güven aralığı oluşturulmuştur. Bu dağılımların herbiri için istatistiksel çıkarım elde etmek amacıyla simülasyon çalışması yapılmıştır. Bu tahmin edicilerin performansları Monte Carlo simülasyon yöntemi ile karşılaştırılmıştır. Her durum için üretilmiş veriler ve mümkün olduğunda gerçek hayat verileri kullanılarak bahsedilen tahmin ediciler için nümerik sonuçlar sunulmuştur.

Anahtar Kelimeler: Rekor Değerler, Stres-Dayanıklılık Güvenilirliği, Bayes Tahmin Edicisi, Kestirim, Lindley Yaklaşımı, Markov Zinciri Monte Carlo (MCMC) Methodu.

ACKNOWLEDGEMENTS

I am grateful to my doctoral supervisor Assoc. Prof. Dr. Mustafa Nadar for sharing his experiences, knowledge and valuable comments during my graduate study.

Besides my advisor, I would like to thank the rest of my thesis committee: Assoc. Prof. Coşkun YAKAR, Prof. Dr. Tahir AZEROĞLU, Prof. Dr. Mansur İSMAİLOV and Prof. Dr. Fatih TAŞÇI.

I am thankful to my friends who have always helped and supported me in different stages of my graduate study.

I am thankful to the Scientific and Technological Research Council of Turkey (TÜBİTAK) for their financial supporting during my graduate study.

Last but not least, I am deeply grateful to my family their patience, help and constantly support during my life.

TABLE of CONTENTS

LIST of ABBREVIATIONS and ACRONYMS

Abbreviations and Acronyms		**Explanations**
$Burr(a, b)$	:	Burr Type XII distribution with parameters a and b
$Gamma(a, b)$	:	Gamma distribution with parameters a and b
$GE(a, b)$	:	Generalized exponential distribution with parameters a and b
$Kum(a, b)$	:	Kumaraswamy distribution with parameters a and b
cdf	:	Cumulative Density Function
cp	:	Coverage Probability
cps	:	Coverage Probabilities
C.I.	:	Confidence Interval
ER	:	Estimated Risk
HPD	:	Highest Probability Density
i.i.d.	:	Independent and Identically Distributed
K-S	:	Kolmogorov-Smirnov
LINEX	:	Linear Exponential
MCMC	:	Markov Chain Monte Carlo
ML	:	Maximum Likelihood
MLE	:	Maximum Likelihood Estimation
MLP	:	Maximum Likelihood Predictor
MSE	:	Mean Square Error
MSPE	:	Mean Square Predictor Error
pdf	:	Probability Density Function
PI	:	Prediction Interval
PMLE	:	Predictive Maximum Likelihood Estimation
SE	:	Square Error
UMVU	:	Uniformly Minimum Variance Unbiased
UMVUE	:	Uniformly Minimum Variance Unbiased Estimation

LIST of TABLES

1. INTRODUCTION

1.1. Overview and Motivation

Record values and the associated statistics are of interest in many real life applications. For example, predicting the flood level of a river that is greater than the previous ones is of importance to climatologists and a hydrologist, while predicting the magnitude of an earthquake which has a greater magnitude than the previous ones, in a given region, is of importance to seismologists. A meteorologist may want to know how much flooding will occur the next time the current rainfall record is broken. The statistician must estimate the next record value of rainfall from a data set consisting of past record values. In industry and reliability studies, many products may fail under stress. For example, a wooden beam breaks when sufficient perpendicular force is applied to it, an electronic component ceases to function in an environment of too high temperature, and a battery dies under the stress of time. But the precise breaking stress or failure point varies even among identical items. Hence, in such experiments, measurements may be made sequentially and only values smaller (or larger) than all previous ones are recorded. This type data is called record data or records. Thus, the number of measurements made is considerably smaller than the complete sample size. Therefore, the measurement saving can be important when the measurements of these experiments are costly or the entire sample is very big or destroyed. For more examples, see [Gulati and Padgett, 1994].

The theory of record values was first introduced by [Chandler, 1952] and it has been extensively studied in the literature since then. A number of statisticians have worked on interesting problems about the records. The distributions of lower records and inter-record times for independent and identically distributed sequences of random variables were studied by [Chandler, 1952]. The theory of the limiting distributions concerning the random variables, which was constituted by the index of the record values, were studied by [Rényi, 1962]. Record values, inter-record times and their limiting properties were studied by many authors. These studies were summarized by [Glick, 1978]. A likelihood function for estimating unknown parameters based on record samples was given by [Arnold et al., 1998]. A predictive likelihood function for future record values was given by [Basak and Balakrishnan, 2003].More details and references can be found in [Ahsanullah, 1995], [Arnold et al., 1998], [Nevzorov, 2001].

1.2. Definitions

- Definition of Record Values and Record Times

Suppose that $X_1, X_2, ...$ be a sequence of independent and identically distributed (i.i.d.) random variables from a continuous distribution. An observation X_j is called an upper record value and j is called upper record time if X_j exceeds that of than all preceding observations. That is X_j is an upper record value if $X_j > X_i$ for all $i < j$. The record times are the indices at which record values occur. The record time sequence for upper record values $\{U(n),\ n \geq 1\}$ is defined in the following manner: $U(1) = 1$ with probability 1 and nth record time, for $n > 1$

$$U(n) = \min\left\{j : j > U(n-1), X_j > X_{U(n-1)}\right\}. \tag{1.1}$$

Then, the nth upper record value is $X_{U(n)}$. Similarly, X_j is called a lower record value if its value is smaller than all preceding observations. The record time sequence for lower record values is denoted by $L(n)$. It is clear that X_1 is a lower (upper) record value and $U(1) = L(1) = 1$ by definitions.

Let $\Delta_r = U(r+1) - U(r)$ and $\Delta_{(r)} = L(r+1) - L(r)$, $r = 1, 2, ...$ Δ_r and $\Delta_{(r)}$ are called upper and lower inter-record times, respectively. Inter-record times correspond roughly to the number of non-record observations between record values.

- Distributions of Record Values and Record Times

Let $X_1, X_2, ...$ be a sequence of i.i.d. random variables with continuous cumulative distribution function (cdf) F and probability density function (pdf) f. The pdf of nth upper record value $X_{U(n)}$,say $f_n(x)$, is

$$f_n(x) = \frac{(R(x))^{n-1}}{(n-1)!} f(x), \quad -\infty < x < \infty \tag{1.2}$$

where $R(x) = -\ln \overline{F}(x)$, $0 < \overline{F}(x) < 1$ and $\overline{F}(x) = 1 - F(x)$. Then the joint pdf $f(x_1, x_2, ..., x_n)$ of n upper record values $X_{U(1)}, X_{U(2)}, ..., X_{U(n)}$ is given by

$$f(x_1, x_2, ..., x_n) = r(x_1)r(x_2)...r(x_{n-1})f(x_n)$$

$$= f(x_n)\prod_{i=1}^{n}\frac{f(x_i)}{1 - F(x_i)}, \tag{1.3}$$

for $-\infty < x_1 < x_2 < ... < x_{n-1} < x_n < \infty$ where $r(x) = dR(x)/dx = f(x)/(1 - F(x))$, $0 < F(x) < 1$ and $r(x)$ is known as hazard rate function.

The pdf of nth lower record value $X_{L(n)}$, say $f_{(n)}(x)$, is

$$f_{(n)}(x) = \frac{(H(x))^{n-1}}{(n-1)!}f(x), \quad -\infty < x < \infty \tag{1.4}$$

and the joint pdf of n lower record values $X_{L(1)}, X_{L(2)}, ..., X_{L(n)}$ is given by

$$f_{(1),...,(n)}(x_1, x_2, ..., x_n) = h(x_1)h(x_2)...h(x_{n-1})f(x_n)$$

$$= f(x_n)\prod_{i=1}^{n}\frac{f(x_i)}{F(x_i)}, \tag{1.5}$$

for $-\infty < x_n < x_{n-1} < ... < x_1 < \infty$ where $H(x) = -\ln F(x)$, $0 < F(x) < 1$ and $h(x) = -dH(x)/dx = f(x)/F(x)$.

An important question that the number of record values among the sequence of observations $X_1, X_2, ..., X_n$. Let M_n be the number of record values among the sequence $X_1, X_2, ..., X_n$. [Rényi, 1962] showed that the mean and variance of M_n are

$$E(M_n) = \sum_{i=1}^{n}\frac{1}{i} \text{ and } Var(M_n) = \sum_{i=1}^{n}\frac{1}{i}\left(1 - \frac{1}{i}\right). \tag{1.6}$$

Moreover, the mean and variance approximately equal to $\log n + \gamma$ and $\log n + \gamma - (\pi^2/6)$, respectively, where γ is Euler's constant 0.5772... and $\log$ is used for the natural logarithm (see [Arnold et al., 1998]). Therefore, record values are clearly not common. A sequence of n i.i.d. continuous random variables only about $\log n$ records are expected.

For all $n \geq 1$, Δ_n and $\Delta_{(n)}$ are identically distributed. [Nevzorov, 2001] shows that the inter-record times are conditionally independent given the record values, and the nth inter-record time has probability mass function

$$P(\Delta_n = k \left| X_{U(1)}, X_{U(2)}, ... \right.) = (1 - F(X_{u(n-1)})) \left(F(X_{U(n-1)}) \right)^{k-1}, \qquad (1.7)$$

for $k = 1, 2, ...$ and $n = 1, 2, 3, ...$ Thus, the nth inter-record time follows a geometric distribution. The pdfs of Δ_n and $\Delta_{(n)}$ are independent of the parent distribution $F(x)$ and are given as

$$P(\Delta_n = k) = P(\Delta_{(n)} = k) = \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^i \frac{1}{(2+i)^n}. \qquad (1.8)$$

Moreover,

$$E(\Delta_n \left| X_{U(n)} = x_n \right.) = \frac{1}{1 - F(x_n)},\ Var(\Delta_n \left| X_{U(n)} = x_n \right.) = \frac{F(x_n)}{(1 - F(x_n))^2}, \quad (1.9)$$

$$E(\Delta_{(n)} \left| X_{L(n)} = x_{(n)} \right.) = \frac{1}{F(x_{(n)})},\ Var(\Delta_{(n)} \left| X_{L(n)} = x_{(n)} \right.) = \frac{1 - F(x_{(n)})}{\left(F(x_{(n)}) \right)^2}. (1.10)$$

The various probabilities of the record times can be easily obtained by using the probability of the inter-record times. It is known that the record times are independent from [Rényi, 1962]. Then, the joint probability mass function of the first n record times is

$$P(U(2) = j_2, U(3) = j_3, ..., U(n) = j_n) = \frac{1}{(j_2 - 1)(j_3 - 1)...(j_n - 1)j_n}, \qquad (1.11)$$

for $1 = j_1 < j_2 < ... < j_n$.

When the consider the sample which consists of the record values and their corresponding inter-record times, we have two sampling schemes for generating record data known as inverse sampling and random sampling schemes. Let K_i is the number of trials required to obtain a new record value, namely $K_i = \Delta_i$ (or $\Delta_{(i)}$). Under the inverse sampling scheme, units are taken sequentially and sampling is terminated when the mth maximum (or minimum) is observed. In this case, the total number of sampled unit is a random number, and K_m is defined to be one for convenience, while under the random sampling scheme, a random sample $X_1, \ldots, X_n$ is examined sequentially and successive maximum (minimum) values are recorded. In this setting the number of records $N^{(n)}$ obtained is a random and $\sum_{i=1}^{N^{(n)}} K_i = n$.

The distribution of K_i given the past upper records and inter-record times is

$$P(K_i = k \,|X_{U(i)} = x_i) = (1 - F(x_i))\,(F(x_i))^{k-1}\,, \tag{1.12}$$

for $k = 1, 2, ...$ It follows that the joint pdf or likelihood function associated with the sequence $\left\{X_{U(1)}, K_1, \ldots, X_{U(m)}, K_m\right\}$ under the inverse sampling scheme is given by [Samaniego and Whitaker, 1986] as

$$L(\mathbf{x}, \mathbf{k}) = \prod_{i=1}^{m} f(x_i)\,\{F(x_i)\}^{k_i - 1}\, I_{(x_{i-1}, \infty)}(x_i), \tag{1.13}$$

where $x_0 \equiv -\infty$, $k_m \equiv 1$ and $I_A(x)$ is the indicator function of the set A. Similarly, the distribution of K_i given the past lower records and inter-record times is

$$P(K_i = k \,|X_{L(i)} = x_i) = F(x_i)\,(1 - F(x_i))^{k-1}\,, \tag{1.14}$$

for $k = 1, 2, ...$ It follows that the joint pdf or likelihood function associated with the sequence $\left\{X_{L(1)}, K_1, \ldots, X_{L(m)}, K_m\right\}$ under the inverse sampling scheme is

$$L(\mathbf{x}, \mathbf{k}) = \prod_{i=1}^{m} f(x_i)\,\{1 - F(x_i)\}^{k_i - 1}\, I_{(-\infty, x_{i-1})}(x_i), \tag{1.15}$$

where $x_0 \equiv \infty$ and $k_m \equiv 1$.

- Definition of The Stress-Strength Reliability

Let X and Y be independent random variables, the quantity of $R = P(X < Y)$ is commonly referred as stress-strength parameter or reliability. In the simplest terms this can be described as an assessment of reliability of a component in terms of random variables X representing stress experienced by the component and Y representing the strength of the component available to overcome the stress. If the stress exceeds the strength, i.e. $X > Y$, then the component will fail.

Assume that a random vector (X, Y) has pdf $f(x, y\,|\theta)$ with an unknown scalar or vector-valued parameter $\theta \in \Theta$. The aim is to estimate R on the basis of observations $(X_1, Y_1), ..., (X_n, Y_n)$. Note that if X and Y are independent with the pdf of the form $f(x, y\,|\theta) = f(x\,|\theta)f(y\,|\theta)$ the number of observations for X and Y need not be the same.

The reliability R can be calculated as

$$R = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x, y\,|\theta)I(x < y)dxdy. \tag{1.16}$$

If X and Y are independent with the pdfs $f(x\,|\theta)$ and $f(y\,|\theta)$ and the cdfs $F_X(x\,|\theta)$ and $F_Y(y\,|\theta)$, respectively, equation (1.16) can be rewritten as

$$R = \int_{-\infty}^{\infty} F_X(z\,|\theta)f_Y(z\,|\theta)dz = \int_{-\infty}^{\infty} \left(1 - F_Y(z\,|\theta)\right) f_X(z\,|\theta)dz. \tag{1.17}$$

The main idea was introduced by [Birnbaum, 1956] and developed by [Birnbaum and McCarty, 1958]. The problem of estimating of R on random samples has been extensively studied under various distributional assumptions on X and Y. A comprehensive account of this topic is presented by [Kotz et al., 2003]. It is provided an excellent review of the development of the stress-strength under classical and Bayesian point of views up to the year 2003. For most recent results on the topic see [Kundu and Gupta, 2005], [Mokhlis, 2005], [Baklizi, 2008], [Rezaei et al., 2010], [Nadar et al., 2014] and the references therein.

- Inferential Methods For Bayesian Analysis

In Bayesian methods, to evaluate various characteristics of posterior and predictive distributions, especially their densities, means and variances are very important. When the problem under consideration does not involve a conjugate prior likelihood pair, these tasks can not be obtained in closed form. In this case, an analytical or a numerical approximation methods are needed. Because the Lindley approximation, the Tierney-Kadane approximation and Markov Chain Monte Carlo (MCMC) methods are used frequently in this thesis, the summary of these methods using [Press, 2002], [Gelman et al., 2003], [Tierney and Kadane, 1986], [Lindley, 1980], [Soliman et al., 2011] are given below.

- The Lindley Approximation

Let $u(\theta)$ be a smooth, positive function on the parameter space. The posterior mean of $u(\theta)$ for given data $\mathbf{x} = (x_1, ..., x_n)$ can be written as

$$E(u(\theta)\,|\mathbf{x}) = \frac{\int u(\theta)e^{l(\theta)+\rho(\theta)}d\theta}{\int e^{l(\theta)+\rho(\theta)}d\theta}, \tag{1.18}$$

where $l(\theta)$ is the logarithm of the likelihood function, $\rho(\theta)$ is the logarithm of the prior density of θ and $\theta = (\theta_1, ..., \theta_m)$ is a parameter. The Lindley approximation is developed by [Lindley, 1980] and is given in the following theorem.

Theorem 1.1: For n sufficiently large and $l(\theta)$ defined in equation (1.18) concentrates around a unique maximum likelihood estimator $\widehat{\theta} = (\widehat{\theta}_1, ..., \widehat{\theta}_m)$ for θ, the ratio of integrals in equation (1.18) is given by approximately as

$$E(u(\theta)\,|\mathbf{x}) = \Bigg[u + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}(u_{ij} + 2u_i\rho_j)\sigma_{ij} + \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{k=1}^{m} L_{ijk}\sigma_{ij}\sigma_{kl}u_l\Bigg]_{\widehat{\theta}} \tag{1.19}$$

where $u_i = \partial u(\theta)/\partial\theta_i$, $u_{ij} = \partial^2 u(\theta)/\partial\theta_i\partial\theta_j$, $L_{ijk} = \partial^3 l(\theta)/\partial\theta_i\partial\theta_j\partial\theta_k$, $\rho_j = \partial\rho(\theta)/\partial\theta_j$, and $\sigma_{ij} = (i,j)$th element in the inverse of the matrix $\{-L_{ij}\}$ all evaluated at the MLE of the parameters.

Proof 1.1: For the proof of theorem see [Lindley, 1980]. ■

Remark 1.1: The first term in equation (1.19) is $O(1)$; the other terms are $O(1/n)$ and are called correction terms. The overall approximation in the Theorem 1.1 is $O(1/n)$, so the first term neglected is $O(1/n^2)$.

Remark 1.2: When $m = 3$, equation (1.19) reduces to

$$E(u(\theta)\,|\mathbf{x}) = u + (u_1c_1 + u_2c_2 + u_3c_3 + c_4 + c_5) + \frac{A}{2}\,(u_1\sigma_{11} + u_2\sigma_{12} + u_3\sigma_{13})$$

$$+\frac{B}{2}(u_1\sigma_{21} + u_2\sigma_{22} + u_3\sigma_{23}) + \frac{C}{2}(u_1\sigma_{31} + u_2\sigma_{32} + u_3\sigma_{33}), \quad (1.20)$$

evaluated at $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3)$ where

$$c_i = \rho_1\sigma_{i1} + \rho_2\sigma_{i2} + \rho_3\sigma_{i3},\ i = 1, 2, 3, \quad (1.21)$$

$$c_4 = u_{12}\sigma_{12} + u_{13}\sigma_{13} + u_{23}\sigma_{23}, \quad (1.22)$$

$$c_5 = \frac{1}{2}(u_{11}\sigma_{11} + u_{22}\sigma_{22} + u_{33}\sigma_{33}), \quad (1.23)$$

$$A = \sigma_{11}L_{111} + 2\sigma_{12}L_{121} + 2\sigma_{13}L_{131} + 2\sigma_{23}L_{231} + \sigma_{22}L_{221} + \sigma_{33}L_{331}, \quad (1.24)$$

$$B = \sigma_{11}L_{112} + 2\sigma_{12}L_{122} + 2\sigma_{13}L_{132} + 2\sigma_{23}L_{232} + \sigma_{22}L_{222} + \sigma_{33}L_{332}, \quad (1.25)$$

$$C = \sigma_{11}L_{113} + 2\sigma_{12}L_{123} + 2\sigma_{13}L_{133} + 2\sigma_{23}L_{233} + \sigma_{22}L_{223} + \sigma_{33}L_{333}. \quad (1.26)$$

Remark 1.3: When $m = 2$, the following notations in [Jaheen, 2005] are used

$$E(u(\theta)\,|\mathbf{x}) = u + \frac{1}{2}\left[B + Q_{30}B_{12} + Q_{21}C_{12} + Q_{12}C_{21} + Q_{03}B_{21}\right], \quad (1.27)$$

where $B = \sum_{i=1}^{2}\sum_{j=1}^{2} u_{ij}\tau_{ij}$, $Q_{ij} = \partial^{i+j}Q/\partial^i\theta_1\partial^j\theta_2$, for $i, j = 0, 1, 2, 3$ and $i + j = 3$, $u_i = \partial u(\theta)/\partial\theta_i$, $u_{ij} = \partial^2 u(\theta)/\partial\theta_i\partial\theta_j$ for $i, j = 1, 2$ and $B_{ij} = (u_i\tau_{ii} + u_j\tau_{ij})\tau_{ii}$ and $C_{ij} = 3u_i\tau_{ii}\tau_{ij} + u_j(\tau_{ii}\tau_{ij} + 2\tau_{ij}^2)$ for $i \neq j$, where τ_{ij} is the (i, j)th element in the inverse of the matrix $Q^* = (-Q^*_{ij})$, $i, j = 1, 2$ such that $Q^*_{ij} = \partial^2 Q/\partial\theta_i\partial\theta_j$, Q is the logarithm of the posterior density function of θ except for the normalizing constant. The equation (1.27) is to be evaluated at $(\widehat{\theta}_1, \widehat{\theta}_2)$, the mode of the posterior density density function of θ.

- The Tierney-Kadane Approximation

Another analytical approximation result for evaluation of Bayesian integrals is Tierney-Kadane approximation. It is developed by [Tierney and Kadane, 1986] and is given in Theorem 1.2.

Theorem 1.2: For n sufficiently large, if the posterior distribution of $u(\theta)$ (given the data) is concentrated on the positive (or negative) half-line, and if $(l(\theta)+\rho(\theta))$, defined in equation (1.18), concentrates around a unique maximum, under suitable regularity conditions the ratio of integrals in equation (1.18) is given approximately by

$$E(u(\theta)\,|\mathbf{x}) = \left[\frac{\det \Sigma^*}{\det \Sigma}\right]^{1/2} \exp\left(n\left[\Lambda^*(\widetilde{\theta}^*) - \Lambda(\widetilde{\theta})\right]\right), \tag{1.28}$$

where $n\Lambda(\theta) = l(\theta) + \rho(\theta)$, $n\Lambda^*(\theta) = \ln u(\theta) + l(\theta) + \rho(\theta)$, $\widetilde{\theta}^*$ maximizes $\Lambda^*(\theta)$, $\widetilde{\theta}$ is the posterior mode and therefore maximizes $l(\theta)$, and Σ^* and Σ are the negatives of the inverse Hessians of $\Lambda^*(\widetilde{\theta}^*)$ and $\Lambda(\widetilde{\theta})$, respectively.

Remark 1.4: The terms omitted in the approximation in equation (1.28) are $O(1/n^2)$, as in the results in equation (1.19).

- MCMC Method

The MCMC method is a general simulation method for sampling from posterior distributions. The MCMC methods sample successively from a target distribution. A Markov chain is generated by sampling the current point based on the previous one. The MCMC method works successfully in Bayesian computing. The analytical forms of the posterior distributions can only be recognized in the simplest models. Most posterior densities are computationally intensive to work with directly. With the MCMC method, it is possible to generate samples from the posterior distribution and to use these samples to approximate expectations of quantities of interest. In addition, the simulation algorithm can be easily extensible to models with a large number of parameters or high complexity. The MCMC techniques, including the Metropolis–Hastings algorithm and the Gibbs sampler have become very popular in recent years as methods for generating a sample from a complicated model. Details of the MCMC method can be found in [Gelman et al., 2003].

- The Gibbs Sampling Algorithm

The Gibbs sampler is a special case of an MCMC algorithm. It generates a sequence of samples from the full conditional probability distributions of two or more random variables. Gibbs sampling requires decomposing the joint posterior distributions into full conditional distributions for each parameter in the model and then sampling from them. The sampler can be efficient when the full conditional distributions are easy to sample from. Suppose that $\theta_1, ..., \theta_k$ denote some grouping (blocking) of θ and $\pi_1^*(\theta_1 | \theta_2, ..., \theta_k, \underline{x}), ..., \pi_k^*(\theta_k | \theta_1, ..., \theta_{k-1}, \underline{x})$ denote the associated conditional densities, often called the full conditional densities. After having derived the full conditional posterior distributions for the parameters $\theta_1, ..., \theta_k$, the Gibbs sampler works as follows:

- Step 1: Begin with some initial values $\theta^{(0)} = (\theta_1^{(0)}, ..., \theta_k^{(0)})$.
- Step 2: Set $j = 1$.
- Step 3: Generate $\theta_1^{(j)}$ from conditional distribution $\pi_1^*(\theta_1 | \theta_2, ..., \theta_k, \underline{x})$.
- Step 4: Generate $\theta_2^{(j)}$ from conditional distribution $\pi_2^*(\theta_2 | \theta_1, ..., \theta_k, \underline{x})$.
- Step 5: Generate $\theta_k^{(j)}$ from conditional distribution $\pi_k^*(\theta_k | \theta_1, ..., \theta_{k-1}, \underline{x})$.
- Step 6: Set $j = j + 1$ and repeats steps 3-5, $j = 1, 2, ..., N$.

- The Metropolis-Hastings Algorithm

The Metropolis-Hastings algorithm is a very general MCMC method. It can be used to obtain random samples from any arbitrarily complicated target distribution of any dimension that is known up to a normalizing constant. In fact, the Gibbs sampler is just a special case of the Metropolis-Hastings algorithm. For specificity, suppose that the full conditional density $\pi_1^*(\theta_1 | \theta_2, ..., \theta_k, \underline{x})$ is intractable. Let $q(\theta_1^* | \theta_2, ..., \theta_k, \underline{x})$ denote a proposal density that generates a candidate θ_1^*. The Metropolis-Hastings algorithm for intractable $\pi_1^*(\theta_1 | \theta_2, ..., \theta_k, \underline{x})$ can be summarized as follows:

- Step 1: Specify an initial value $\theta^{(0)} = (\theta_1^{(0)}, ..., \theta_k^{(0)})$.
- Step 2: Propose a value for θ_1 by drawing $\theta^* \sim q(\theta_1^* | \theta_2^{(j-1)}, ..., \theta_k^{(j-1)}, \underline{x})$.
- Step 3: Calculate the acceptance probability

$$\rho(\theta_1^{(j-1)}, \theta^*) = \min\left[1, \frac{\pi_1^*(\theta_1^* | \theta_2^{(j-1)}, .., \theta_k^{(j-1)}, \underline{x}) q(\theta_1^{(j-1)} \big| \theta_2^{(j-1)}, .., \theta_k^{(j-1)}, \underline{x})}{\pi_1^*(\theta_1^{(j-1)} \big| \theta_2^{(j-1)}, .., \theta_k^{(j-1)}, \underline{x}) q(\theta_1^* | \theta_2^{(j-1)}, .., \theta_k^{(j-1)}, \underline{x})}\right] \quad (1.29)$$

- Step 4: Generate $U \sim Uniform(0,1)$.
- Step 5: If $U \leq \rho(\theta_1^{(j-1)}, \theta^*)$, accept the proposal and set $\theta_1^{(j)} = \theta^*$. Otherwise, reject the proposal and set $\theta_1^{(j)} = \theta_1^{(j-1)}$.
- Step 6: Set $j = j+1$ and repeats steps 1-4, $j = 1, 2, ..., N$.

If the proposal distribution is symmetric, then we have $\rho(\theta^{(t-1)} \big| \theta^*) = \rho(\theta^* | \theta^{(t-1)})$, so that the acceptance probability is given by

$$\rho(\theta_1^{(j-1)}, \theta^*) = \min\left[1, \frac{\pi_1^*(\theta_1^* | \theta_2^{(j-1)}, ..., \theta_k^{(j-1)}, \underline{x})}{\pi_1^*(\theta_1^{(j-1)} \Big| \theta_2^{(j-1)}, ..., \theta_k^{(j-1)}, \underline{x})}\right]. \tag{1.30}$$

A similar approach is used to sample $\theta_2, ..., \theta_k$.

1.3. The Aim of The Thesis

In recent years, the record values and the stress-strength reliability models are getting more popular among the statisticians. Many authors have investigated the statistical inferences of the record values and the stress-strength reliability models for the different distributions. However, the estimation of the unknown distribution parameters based on record values with their corresponding inter-record times and the estimation of the stress-strength reliability based on record values have not paid much attention in the literature. Hence, we basically concentrate on these subjects in this thesis when the underlying distributions are the Burr Type XII, the generalized exponential and the Kumaraswmay.

2. STATISTICAL ANALYSIS FOR THE BURR TYPE XII DISTRIBUTION

2.1. Introduction

The Burr system of distributions includes twelve types of cumulative distribution functions which yield a variety of density shapes and were listed in,[Burr, 1942]. It has applied in business, chemical engineering, quality control, medical and reliability studies. The Burr XII distribution is one of the different distributions introduced by [Burr, 1942] for modeling data.

If a random variable X follows a Burr Type XII distribution, denoted by $Burr(\alpha, \beta)$, then its pdf and cdf are given by

$$F(x;\alpha,\beta) = 1-(1+x^{\alpha})^{-\beta},\ x>0,\ (\alpha>0,\ \beta>0) \tag{2.1}$$

$$f(x;\alpha,\beta) = \alpha\beta x^{\alpha-1}(1+x^{\alpha})^{-(\beta+1)},\ x>0 \tag{2.2}$$

$\alpha, \beta > 0$ are the shape parameters. The mean and variance of a Burr Type XII distribution are given by

$$E(X) = \beta B\left(\beta-\frac{1}{\alpha}, 1+\frac{1}{\alpha}\right), \tag{2.3}$$

and

$$Var(X) = \beta B\left(\beta-\frac{2}{\alpha}, 1+\frac{2}{\alpha}\right) - \left\{\beta B\left(\beta-\frac{1}{\alpha}, 1+\frac{1}{\alpha}\right)\right\}^2. \tag{2.4}$$

The Burr Type XII distribution is unimodal and its mode $x_{mode} = (\alpha-1/(\alpha\beta+1))^{1/\alpha}$ if $\alpha>1$. If $\alpha>1$, its pdf increases on $(0, x_{mode}]$ and decreases on $[x_{mode}, \infty)$. If $\alpha \geq 1$, its pdf is strictly decreasing.

The Burr Type XII distribution has been studied by many authors. For example, the Bayes estimates of the shape parameter and reliability function were derived by [Papadopoulos, 1978] when the other shape parameter was known. The Bayes estimates of the parameters, the reliability and failure rate functions based on a Type-2 censored sample were obtained by [Al-Hussaini and Jaheen, 1992]. The Bayesian prediction bounds for certain order statistics were considered by [Al-Hussaini and Jaheen, 1995]. The maximum likelihood (ML) estimates of the parameters based on randomly right censored data were obtained by [Ghitany and Al-Awadhi, 2002]. The ML and Bayes estimates of the parameters based on generalized order statistics were derived by [Jaheen, 2005]. The ML and Bayes estimates for some life time parameters (reliability and hazard functions) as well as the shape parameters based on progressively Type-II censored samples were obtained by [Soliman, 2005].

The rest of this chapter is organized as follows. In Section 2.1, the statistical inferences for the Burr Type XII distribution based on record values are mentioned. In Section 2.2, the statistical inferences for the Burr Type XII distribution based on record values with their corresponding inter-record times are considered. In Section 2.3, the statistical inferences for the stress-strength reliability of the Burr Type XII distribution based on record values are considered.

2.2. Estimation of The Parameters Based on Record Values

The Bayesian estimates for the two shape parameter of the Burr Type XII distribution based on upper record values were obtained by [Ahmadi and Doostparast, 2006] using the symmetric loss function. Bayesian prediction bounds for future upper record values was also derived. When the first shape parameter was known, the Bayes and empirical Bayes estimates for the unknown shape parameter of the Burr Type XII distribution based on upper record values were considered by [Wang and Shi, 2010]. The Bayesian and empirical Bayesian prediction bounds for future upper record values were also obtained. The frequentist and Bayesian point estimates for the two shape parameters based on upper record values were derived by [Nadar and Papadopoulos, 2011] using the asymmetric loss function. The prediction for the future record values was also obtained by using non-Bayesian and Bayesian approach.

2.3. Estimation of the Parameters Based on Records and Inter-Record Times

When the underlying distribution is exponential, estimation of the mean parameter by using record values and their corresponding inter-record times was obtained by [Samaniego and Whitaker, 1986] under random sampling and inverse sampling schemes. The optimal random sampling plan and associated cost analysis for the exponential distribution were studied by [Doostparast and Balakrishnan, 2010]. Non-Bayesian and Bayesian estimates were derived by [Doostparast, 2009] for the two parameters of the exponential distribution based on record values and their corresponding inter-record times under the inverse sampling scheme. The optimal confidence intervals, uniformly most powerful tests for one-sided alternatives were derived by [Doostparast and Balakrishnan, 2011] when the underlying distribution is the two-parameter exponential distribution. Also, they obtained as generalized likelihood ratio test, uniformly unbiased and invariant tests for a two-sided alternative. The optimal statistical procedures including point and interval estimation as well as most powerful tests based on record data from a two-parameter Pareto model were obtained by [Doostparast and Balakrishnan, 2013]. When the underlying distribution is lognormal, non-Bayesian and Bayesian point estimates as well as asymptotic confidence intervals for the unknown parameters were obtained by [Doostparast et al., 2013].

Prediction of future records becomes a problem of great interest. For example, while studying the record rainfalls or snowfalls, having observed the record values until the present time, we will be naturally interested in predicting the amount of rainfall or snowfall that is to be expected when the present record is broken for the first time in future. Let $R_1, ..., R_m$ be the first m lower record values observed from a specific distribution. Then, we may be interested in predicting (either point or interval prediction) the value of the next record (R_{m+1}), or, more generally, the value of the s-th record (R_s) for some $s > m$ (see [Arnold et al., 1998]). Prediction of future records has been studied by many authors such that [Ahmadi and Doostparast, 2006], [Soliman et al., 2006], [Raqab et al., 2007].

In this section, the parameter estimates of Burr Type XII distribution based on lower record values and their corresponding inter-record times are obtained under the classical and Bayesian frameworks. The Lindley approximation and MCMC technique are used to obtain the Bayes estimates under different loss functions. The non-Bayesian and Bayesian point predictors and the Bayesian prediction interval of future lower record values are obtained based on the observed lower record values with their corresponding inter-record times. Also, the Bayesian point predictors and the Bayesian prediction interval of future lower record values are constructed based on just the lower record values. Finally, the two approach are compared by using Monte Carlo simulations to see the effect of the inter-record times in prediction.

2.3.1. ML Estimation

Let $X_1, X_2, \ldots$ be i.i.d. random variables, coming from a population with the cdf and the pdf $F(.)$ and $f(.)$, respectively. Then, the likelihood function associated with the sequence $\{R_1, K_1, \ldots, R_m, K_m\}$ is given by [Samaniego and Whitaker, 1986] as

$$L(\mathbf{r}, \mathbf{k}) = \prod_{i=1}^{m} f(r_i) \{1 - F(r_i)\}^{k_i - 1} I_{(-\infty, r_{i-1})}(r_i), \tag{2.5}$$

where $r_0 \equiv \infty$, $k_m \equiv 1$ and $I_A(x)$ is the indicator function of the set A. From equations (2.1), (2.2) and (2.5), we have

$$L(\alpha, \beta; \mathbf{r}, \mathbf{k}) = \alpha^m \beta^m \exp\left\{ (\alpha - 1) \sum_{i=1}^{m} \ln r_i - \sum_{i=1}^{m} (\beta k_i + 1) \ln(1 + r_i^{\alpha}) \right\}, \tag{2.6}$$

where $r_1 > \ldots > r_m$ and so the log-likelihood function is

$$l(\alpha, \beta; \mathbf{r}, \mathbf{k}) = m(\ln \alpha + \ln \beta) + (\alpha - 1) \sum_{i=1}^{m} \ln r_i - \sum_{i=1}^{m} (\beta k_i + 1) \ln(1 + r_i^{\alpha}). \tag{2.7}$$

The ML estimates of α and β are given by

$$\widehat{\beta} = \frac{m}{T_\alpha}, \tag{2.8}$$

where $T_\alpha = \sum_{i=1}^{m} k_i \ln(1 + r_i^{\alpha})$ and $\widehat{\alpha}$ is the solution of the following non-linear equation

$$\frac{m}{\alpha}+\sum_{i=1}^{m}\ln r_i-\sum_{i=1}^{m}\left(k_i\frac{m}{T_\alpha}+1\right)\frac{r_i^\alpha\ln r_i}{1+r_i^\alpha}=0. \tag{2.9}$$

Therefore, $\widehat{\alpha}$ can be obtained as the solution of the non-linear equation of the form $h(\alpha)=\alpha$ where

$$h(\alpha)=m\left[\frac{m}{T_\alpha}\sum_{i=1}^{m}\frac{k_i r_i^\alpha\ln r_i}{1+r_i^\alpha}-\sum_{i=1}^{m}\frac{\ln r_i}{1+r_i^\alpha}\right]^{-1}. \tag{2.10}$$

Since, $\widehat{\alpha}$ is a fixed point solution of the non-linear equation (2.10), its value can be obtained using an iterative scheme as: $\alpha_{(j+1)}=h(\alpha_{(j)})$, where $\alpha_{(j)}$ is the jth iterate of $\widehat{\alpha}$. The iteration procedure should stopped when $\left|\alpha_{(j)}-\alpha_{(j+1)}\right|$ is sufficiently small. After $\widehat{\alpha}$ is obtained, $\widehat{\beta}$ can be obtained from equation (2.8).

Next, we establish the existence and uniqueness of the maximum likelihood estimation (MLE) of the parameters α and β of the Burr Type XII distribution based on lower record data. Similar result has been obtained by [Ghitany and Al-Awadhi, 2002] for the Burr Type XII distribution using randomly right censored data. We present the following lemma that will be used in the proof of Theorem 2.1.

Lemma 2.1: Let

$$w_m(x_1,...,x_m)=\left\{\sum_{i=1}^{m}k_i\ln(1+x_i)\right\}^2-\left\{\sum_{i=1}^{m}k_i\xi(x_i)\right\}^2+\sum_{i=1}^{m}k_i\ln(1+x_i)\sum_{i=1}^{m}k_i\frac{\xi^2(x_i)}{x_i}, \tag{2.11}$$

where $\xi(x)=x\ln x/(1+x)$, $x\geq 0$. *Then,* $w_m(x_1,...,x_m)\geq 0$ *for all* $x_i\geq 0$ *and* $k_i\geq 1$ *,*$i=1,...,m$.

Proof 2.1: For a proof, one may refer to [Ghitany and Al-Awadhi, 2002]. ■

Theorem 2.1: The ML estimates of the parameters α *and* β *are unique and given by*

$\widehat{\beta} = m/T_{\widehat{\alpha}}$ *where* $\widehat{\alpha}$ *is the solution of the non-linear equation:*

$$G(\alpha) = \frac{m}{\alpha} - \frac{m}{T_\alpha}\sum_{i=1}^{m}\frac{k_i r_i^\alpha \ln r_i}{1+r_i^\alpha} + \sum_{i=1}^{m}\frac{\ln r_i}{1+r_i^\alpha} = 0. \tag{2.12}$$

Proof 2.1: We have

$$G(0) \equiv \lim_{\alpha\to 0} G(\alpha) = \lim_{\alpha\to 0}\frac{m}{\alpha} - \frac{m}{2\ln 2}\frac{\sum_{i=1}^{m} k_i \ln r_i}{\sum_{i=1}^{m} k_i} + \frac{1}{2}\sum_{i=1}^{m}\ln r_i = \infty. \tag{2.13}$$

The limit of $G(\alpha)$ *as* $\alpha \to \infty$ *is considered in four cases.*

i) If at least one record value is greater than unity, then

$$G(\infty) \equiv \lim_{\alpha\to\infty} G(\alpha) = \lim_{\alpha\to\infty}\left(\frac{m}{\alpha} - \frac{m}{T_\alpha}\sum_{i=1}^{m}\frac{k_i r_i^\alpha \ln r_i}{1+r_i^\alpha} + \sum_{i=1}^{m}\frac{\ln r_i}{1+r_i^\alpha}\right)$$

$$= -m\lim_{\alpha\to\infty}\frac{\sum_{i=1(r_i<1)}^{m}(k_i r_i^\alpha \ln r_i/(1+r_i^\alpha)) + \sum_{i=1(r_i>1)}^{m}(k_i r_i^\alpha \ln r_i/(1+r_i^\alpha))}{\sum_{i=1(r_i<1)}^{m} k_i \ln(1+r_i^\alpha) + \sum_{i=1(r_i>1)}^{m} k_i \ln(1+r_i^\alpha)}$$

$$+ \lim_{\alpha\to\infty}\left(\sum_{i=1(r_i<1)}^{m}\frac{\ln r_i}{1+r_i^\alpha} + \sum_{i=1(r_i>1)}^{m}\frac{\ln r_i}{1+r_i^\alpha}\right) \tag{2.14}$$

$$= \sum_{i=1(r_i<1)}^{m}\ln r_i < 0.$$

ii) If at least one record value is less than unity, then

$$G(\infty) = \sum_{i=1(r_i<1)}^{m}\ln r_i < 0. \tag{2.15}$$

iii) If all record values are less than unity, that is $r_i < 1, i = 1, ..., m$*, then* $(r_i/r_1) < 1$, $\lim_{\alpha\to\infty}(r_i/r_1)^\alpha = 0$ *and* $\lim_{\alpha\to\infty}(\ln(1+r_i^\alpha))/r_1^\alpha = 0, i = 2, ..., m$*. By using these limits and dividing the numerator and denominator of the second term of* $G(\alpha)$

by r_1^α, then we obtain

$$G(\infty) = -m \lim_{\alpha\to\infty} \frac{\sum_{i=1}^m (k_i \ln r_i/(1+r_i^\alpha))(r_i/r_1)^\alpha}{\sum_{i=1}^m k_i(\ln(1+r_i^\alpha)/r_1^\alpha)} + \lim_{\alpha\to\infty}\sum_{i=1}^m \frac{\ln r_i}{1+r_i^\alpha}$$

$$= -m\frac{k_1 \ln r_1}{k_1} + \sum_{i=1}^m \ln r_i = \sum_{i=1}^m (\ln r_i - \ln r_1) < 0. \tag{2.16}$$

iv) If all record values are greater than unity, that is $r_i > 1, i = 1, ..., m$, then

$$G(\infty) = -m \lim_{\alpha\to\infty} \frac{\sum_{i=1}^m k_i \ln r_i (r_i^\alpha/(1+r_i^\alpha))}{\sum_{i=1}^m k_i \ln(1+r_i^\alpha)} + \lim_{\alpha\to\infty}\sum_{i=1}^m \frac{\ln r_i}{1+r_i^\alpha}$$

$$< \sum_{i=1}^m \ln r_i \lim_{\alpha\to\infty}\left(\frac{1}{(1+r_m^\alpha)} - \frac{1}{\ln(1+r_m^\alpha)}\right) < 0. \tag{2.17}$$

Hence, we obtain that $\lim_{\alpha\to 0} G(\alpha) = \infty$ and $\lim_{\alpha\to\infty} G(\alpha) < 0$. By the intermediate value theorem $G(\alpha)$ has at least one root in $(0,\infty)$. If it can be shown that $\partial G(\alpha)/\partial\alpha < 0$ then the proof will be completed. It is easily obtained that

$$\frac{\partial G(\alpha)}{\partial\alpha} = -\frac{1}{\alpha^2}\left[m + \sum_{i=1}^m r_i^\alpha\left(\frac{\ln r_i^\alpha}{1+r_i^\alpha}\right)^2\right.$$

$$\left. + \frac{m}{T_\alpha^2}\left\{\sum_{i=1}^m k_i r_i^\alpha\left(\frac{\ln r_i^\alpha}{1+r_i^\alpha}\right)^2 \sum_{i=1}^m k_i \ln(1+r_i^\alpha) - \left(\sum_{i=1}^m \frac{k_i r_i^\alpha \ln r_i^\alpha}{1+r_i^\alpha}\right)^2\right\}\right] \tag{2.18}$$

$$= -\frac{1}{\alpha^2}\left[\sum_{i=1}^m \frac{\xi^2(r_i^\alpha)}{r_i^\alpha} + \frac{m}{T_\alpha^2} w_m(r_1^\alpha, ..., r_m^\alpha)\right].$$

It is clear that $\partial G(\alpha)/\partial\alpha < 0$ by using Lemma 1.

Finally, we will show that the ML estimates of (α,β) maximizes the log-likelihood function $l(\alpha,\beta;\mathbf{r},\mathbf{k})$. Let $H(\alpha,\beta)$ be the Hessian matrix of $l(\alpha,\beta;\mathbf{r},\mathbf{k})$ at (α,β). It is clear that $H_{11}(\widehat{\alpha},\widehat{\beta}) < 0$ and the determinant of the Hessian matrix

$$D(\widehat{\alpha},\widehat{\beta}) = H_{11}(\widehat{\alpha},\widehat{\beta})H_{22}(\widehat{\alpha},\widehat{\beta}) - \Big(H_{12}(\widehat{\alpha},\widehat{\beta})\Big)^2$$

$$= \frac{1}{\widehat{\alpha}^2}\left[\frac{m}{\widehat{\beta}^2}\sum_{i=1}^{m}\frac{\xi^2(r_i^{\widehat{\alpha}})}{r_i^{\widehat{\alpha}}} + w_m(r_1^{\widehat{\alpha}},...,r_m^{\widehat{\alpha}})\right], \tag{2.19}$$

and $D(\widehat{\alpha},\widehat{\beta}) > 0$ by Lemma 2.1. Hence, $(\widehat{\alpha},\widehat{\beta})$ is the local maximum of $l(\alpha,\beta;\mathbf{r},\mathbf{k})$. Since there is no singular point of $l(\alpha,\beta;\mathbf{r},\mathbf{k})$ and it has a single critical point then, it is enough to show that the absolute maximum of the function is indeed the local maximum. Assume that there exist a $\widehat{\alpha}_0$ in the domain in which $l^*(\widehat{\alpha}_0) > l^*(\widehat{\alpha})$, where $l^*(\widehat{\alpha}) = l(\widehat{\alpha},\widehat{\beta};\mathbf{r},\mathbf{k})$. Since $\widehat{\alpha}$ is the local maximum there should be some point α_1 in the neighborhood of $\widehat{\alpha}_{ML}$ such that $l^*(\widehat{\alpha}) > l^*(\alpha_1)$. Let $k(\alpha) = l^*(\alpha) - l^*(\widehat{\alpha})$ then $k(\widehat{\alpha}_0) > 0$, $k(\alpha_1) < 0$ and $k(\widehat{\alpha}) = 0$. This implies that α_1 is a local minimum of the $l^*(\alpha)$, but $\widehat{\alpha}$ is the only critical point so it is a contradiction. Therefore, $(\widehat{\alpha},\widehat{\beta})$ is the absolute maximum of $l(\alpha,\beta;\mathbf{r},\mathbf{k})$. ■

2.3.1.1. ML Estimation When α Is Known

Without loss of generality, the parameter α is assumed to be known, say $\alpha = \alpha_0$. Then, by equation (2.6)

$$L(\alpha_0,\beta;\mathbf{r},\mathbf{k}) = \alpha_0^m \exp\left\{(\alpha_0-1)\sum_{i=1}^{m}\ln r_i - \sum_{i=1}^{m}(\beta k_i+1)\ln(1+r_i^{\alpha_0})\right\}, \tag{2.20}$$

where $r_1 > \ldots > r_m$. In this case, T_{α_0} is a sufficient statistic for β and the MLE of β is $\widehat{\beta}_{ML} = m/T_{\alpha_0}$. The moment generating function of T_{α_0} is $M(t) = 1/(1-t/\beta)^m, \beta > t$. By the uniqueness of the moment generating function, T_{α_0} is distributed as $Gamma(m, 1/\beta)$ and its mean and variance are m/β and m/β^2, respectively. Therefore, $E(\widehat{\beta}_{ML}) = (m/(m-1))\beta$ and an unbiased estimator for β is $\widehat{\beta}_U = (m-1)/T_{\alpha_0}$. Notice that, $MSE(\widehat{\beta}_U) < MSE(\widehat{\beta}_{ML})$ and $MSE(\widehat{\beta}_{ML}) \to 0$ as $m \to \infty$ then $\widehat{\beta}_{ML}$ and $\widehat{\beta}_U$ converge to β in mean square.

2.3.1.2. Asymptotic Confidence Interval

In practice, the observed information matrix is used as a consistent estimator of the Fisher information matrix. An asymptotic confidence intervals for the parameters α and β based on the record values and their corresponding inter-record times are obtained by using the observed information matrix. The observed information matrix $\mathbf{J}_m(\alpha, \beta)$ is given by

$$\mathbf{J}_m(\alpha, \beta) = -\begin{bmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \beta} \\ \frac{\partial^2 l}{\partial \beta \partial \alpha} & \frac{\partial^2 l}{\partial \beta^2} \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}, \tag{2.21}$$

where $J_{11} = (m/\alpha^2) + \sum_{i=1}^{m}(\beta k_i + 1) r_i^{\alpha} \left(\ln r_i/(1 + r_i^{\alpha})\right)^2$, $J_{12} = J_{21} = \sum_{i=1}^{m} \left(k_i r_i^{\alpha} \ln r_i\right) / \left(1 + r_i^{\alpha}\right)$ and $J_{22} = m/\beta^2$.

By the asymptotic normality of the MLE, we have $\left[\sqrt{m}(\widehat{\alpha}_{ML} - \alpha), \sqrt{m}(\widehat{\beta}_{ML} - \beta)\right] \overset{a}{\sim} N_2(\mathbf{0}, \mathbf{I}^{-1})$ for large m, where $\overset{a}{\sim}$ means approximately distributed and $\mathbf{I}^{-1}$ is the inverse of the Fisher information matrix. If the likelihood equations have a unique solution $\widehat{\theta}_n$, then $\widehat{\theta}_n$ is consistent, asymptotically normal and efficient, see [Lehmann and Casella, 1998]. When the likelihood equations have a unique solution, the observed information matrix $J_m(\widehat{\alpha}_{ML}, \widehat{\beta}_{ML})/m$ is a consistent estimator for $I_m(\alpha, \beta)/m$ (see Appendix C in [Lawless, 2003]). Therefore, the observed information matrix can be used in the asymptotic normality of the MLE. For large m (the number of record values) under inverse sampling, the approximate $100(1 - \gamma)\%$ equi-tail confidence intervals for α and β are constructed as

$$\left(\widehat{\alpha}_{ML} \pm z_{1-\gamma/2}\sqrt{\frac{J_{22}}{J_{11}J_{22} - J_{12}^2}}\right) \text{ and } \left(\widehat{\beta}_{ML} \pm z_{1-\gamma/2}\sqrt{\frac{J_{11}}{J_{11}J_{22} - J_{12}^2}}\right), \tag{2.22}$$

where z_γ is the upper γth quantile of the standard normal distribution.

2.3.2. Bayesian Estimation

In the Bayesian inference, the most commonly used loss function is the squared error (SE) loss, $L(\theta^*, \theta) = (\theta^* - \theta)^2$, where θ^* is an estimate of θ. This loss function is symmetrical and gives equal weight to overestimation as well as underestimation. It is well known that the use of symmetric loss functions may be inappropriate in many circumstances, particularly when positive and negative errors have different consequences. A useful asymmetric loss function is the linear-exponential (LINEX) loss, $L(\theta^*, \theta) = e^{v(\theta^* - \theta)} - v(\theta^* - \theta) - 1,\ v \neq 0$, introduced by [Varian, 1975]. The sign and magnitude of v represents the direction and degree of asymmetry, respectively. For v close to zero, the LINEX loss is approximately equal to the SE loss and therefore almost symmetric.

In this section, the Bayes estimates of the parameters Burr Type XII distribution are obtained by using different loss functions for both α is known and unknown cases under the inverse sampling scheme.

2.3.2.1. Bayesian Estimation When α Is Known

When the parameter α is assumed to be known, say $\alpha = \alpha_0$, the gamma conjugate prior density is used for the parameter β, that was first used by [Papadopoulos, 1978], i.e.

$$\pi(\beta) = \frac{b_1^{a_1+1}}{\Gamma(a_1+1)} \beta^{a_1} e^{-\beta b_1},\ \beta > 0. \tag{2.23}$$

The posterior density function of β is readily obtained from equations (2.20) and (2.23) as $\beta| \left(\mathbf{r}, \mathbf{k}\right) \sim Gamma\left(m + a_1 + 1, (b_1 + T_{\alpha_0})^{-1}\right)$. Then, the Bayes estimate of α under the SE loss function, $\widehat{\beta}_{BS,1}$, is the mean of the $\beta| \left(\mathbf{r}, \mathbf{k}\right)$. Therefore

$$\widehat{\beta}_{BS,1} = \frac{m + a_1 + 1}{b_1 + T_{\alpha_0}}, \tag{2.24}$$

and the Bayes estimate of β under the LINEX loss function, $\widehat{\beta}_{BL,1}$, is given by

$$\widehat{\beta}_{BL,1} = -\frac{1}{v} \ln E_{\beta|(\mathbf{r},\mathbf{k})}(e^{-v\beta}) = \frac{m + a_1 + 1}{v} \ln\left(1 + \frac{v}{b_1 + T_{\alpha_0}}\right). \tag{2.25}$$

If we use the Jeffrey's non-informative prior, that is $\pi(\beta) = 1/\beta$, then we have $\beta|\left(\mathbf{r},\mathbf{k}\right) \sim Gamma\left(m, 1/T_{\alpha_0}\right)$. Therefore, the Bayes estimates of α under the SE and the LINEX loss functions are obtained as

$$\widehat{\beta}_{BS,0} = \frac{m}{T_{\alpha_0}} \text{ and } \widehat{\alpha}_{BL,0} = \frac{m}{v} \ln\left(1 + \frac{v}{T_{\alpha_0}}\right), \tag{2.26}$$

respectively. Notice that, $\widehat{\beta}_{BS,0}$ and $\widehat{\beta}_{BL,0}$ are the limit of $\widehat{\beta}_{BS,1}$ and $\widehat{\beta}_{BL,1}$ as $a_1 \to 0$ and $b_1 \to 0$. Moreover, $\widehat{\beta}_{BL,1} \to \widehat{\beta}_{BS,1}$ as $v \to 0$ is satisfied. The Bayesian credible interval can be easily constructed by using the posterior density function of β. It is clear that $2(b_1 + T_{\alpha_0})\beta|\left(\mathbf{r},\mathbf{k}\right) \sim \chi^2_{2(m+a_1+1)}$. Therefore, a Bayesian credible interval for β is given by

$$\left(\frac{\chi^2_{2(m+a_1+1)}(\gamma/2)}{2(b_1 + T_{\alpha_0})}, \frac{\chi^2_{2(m+a_1+1)}(1-\gamma/2)}{2(b_1 + T_{\alpha_0})}\right). \tag{2.27}$$

In the following proposition, the comparison of Bayes estimates are given under the SE and the LINEX loss functions.

Proposition 2.1:

i) $\widehat{\beta}_{BL,1} \leq \widehat{\beta}_{BS,1}$ *for* $v > 0$.

ii) $\widehat{\beta}_{BL,1} \geq \widehat{\beta}_{BS,1}$ *for* $-(b_1 + T_{\alpha_0}) < v < 0$.

Proof 2.1: It is known that

$$\ln(1+x) \leq x \text{ for every } x > -1. \tag{2.28}$$

i) Suppose $v > 0$. $v/(b_1 + T_{\alpha_0}) > 0$, *when* $b_1 > 0$ *and* $T_{\alpha_0} > 0$. *We have* $\ln\left(1 + v/(b_1 + T_{\alpha_0})\right) \leq v/(b_1 + T_{\alpha_0})$ *by the inequality (2.28). Therefore,* $\widehat{\beta}_{BL,1} \leq \widehat{\beta}_{BS,1}$.

ii) Suppose $v < 0$ *and* $-(b_1 + T_{\alpha_0}) < v$, *then* $v/(b_1 + T_{\alpha_0}) > -1$. *We have* $\ln\left(1 + v/(b_1 + T_{\alpha_0})\right) \leq v/(b_1 + T_{\alpha_0})$ *by the inequality (2.28). Therefore,* $\widehat{\beta}_{BL,1} \geq \widehat{\beta}_{BS,1}$. ■

2.3.2.2. Bayesian Estimation When α and β Are Unknown

Assume that the parameters α and β have a joint bivariate density function, suggested by [Al-Hussaini and Jaheen, 1992]. The parameters α and β follow the joint bivariate density function

$$\pi(\alpha,\beta) = \pi_1(\beta\,|\alpha)\pi_2(\alpha), \tag{2.29}$$

where

$$\pi_1(\beta\,|\alpha) = \frac{\alpha^{a_1+1}}{\Gamma(a_1+1)b_1^{a_1+1}}\beta^{a_1}e^{-\alpha\beta/b_1},\ \beta > 0\ (a_1 > -1, b_1 > 0), \tag{2.30}$$

and α has gamma distribution with parameters (a_2, b_2). From equations (2.6) and (2.29), the joint posterior density function of α and β is

$$\pi(\alpha,\beta|\,\mathbf{r},\mathbf{k}) = I(\mathbf{r},\mathbf{k})\alpha^{m+a_1+a_2}\beta^{m+a_1}$$
$$\exp\left\{-\beta\left(\frac{\alpha}{b_1}+T_\alpha\right)-\alpha\left(\frac{1}{b_2}-\sum_{i=1}^{m}\ln r_i\right)-\sum_{i=1}^{m}\ln(1+r_i^{\alpha})\right\}, \tag{2.31}$$

where

$$\frac{[I(\mathbf{r},\mathbf{k})]^{-1}}{\Gamma(m+a_1+1)} = \int_0^\infty \alpha^{m+a_1+a_2}\left(\frac{\alpha}{b_1}+T_\alpha\right)^{-m-a_1-1}$$
$$\exp\left\{-\alpha\left(\frac{1}{b_2}-\sum_{i=1}^{m}\ln r_i\right)-\sum_{i=1}^{m}\ln(1+r_i^{\alpha})\right\}d\alpha. \tag{2.32}$$

The Bayes estimator of a given measurable function of α and β, say $g(\alpha,\beta)$ under the SE loss function is

$$\widehat{g}_{BS} = E_{\alpha,\beta|\mathbf{r},\mathbf{k}}(g(\alpha,\beta)) = \frac{\int_0^\infty\int_0^\infty g(\alpha,\beta)L(\alpha,\beta;\mathbf{r},\mathbf{k})\pi(\alpha,\beta)d\alpha d\beta}{\int_0^\infty\int_0^\infty L(\alpha,\beta;\mathbf{r},\mathbf{k})\pi(\alpha,\beta)d\alpha d\beta}. \tag{2.33}$$

It is not possible to compute equation (2.33) analytically. Two approaches are suggested here to approximate equation (2.33), namely Lindley's approximation and MCMC method.

- Lindley's approximation

  For the two parameter case (α, β), we have from equation (2.31)

$$Q = \ln I(\mathbf{r}, \mathbf{k}) + (m + a_1 + a_2) \ln \alpha + (m + a_1) \ln \beta$$

$$- \beta \left(\frac{\alpha}{b_1} + T_\alpha \right) - \alpha \left(\frac{1}{b_2} - \sum_{i=1}^{m} \ln r_i \right) - \sum_{i=1}^{m} \ln(1 + r_i^\alpha). \quad (2.34)$$

The joint posterior mode is obtained from the equations $\partial Q/\partial\alpha = 0$ and $\partial Q/\partial\beta = 0$ as

$$\widetilde{\beta} = \frac{m + a_1}{(\widetilde{\alpha}/b_1) + T_{\widetilde{\alpha}}}, \quad (2.35)$$

and $\widetilde{\alpha}$ is the solution of the nonlinear equation

$$\frac{m + a_1 + a_2}{\alpha} - \frac{m + a_1}{\frac{\widetilde{\alpha}}{b_1} + T_{\widetilde{\alpha}}} \left(\frac{1}{b_1} + \sum_{i=1}^{m} \frac{k_i r_i^\alpha \ln r_i}{1 + r_i^\alpha} \right) - \frac{1}{b_2} + \sum_{i=1}^{m} \frac{\ln r_i}{1 + r_i^\alpha} = 0. \quad (2.36)$$

It can be solved by using the same procedure in equation (2.10). The elements of the Q^*

$$Q_{11}^* = \frac{m + a_1 + a_2}{\alpha^2} + \sum_{i=1}^{m} (\beta k_i + 1) r_i^\alpha \left(\frac{\ln r_i}{1 + r_i^\alpha} \right)^2, \quad (2.37)$$

$$Q_{12}^* = Q_{21}^* = \frac{1}{b_1} + \sum_{i=1}^{m} \frac{k_i r_i^\alpha \ln r_i}{1 + r_i^\alpha}, \; Q_{22}^* = \frac{m + a_2}{\beta^2}, \quad (2.38)$$

and τ_{ij}, $i, j = 1, 2$ are obtained by using equations (2.37) and (2.38). Moreover,

$$Q_{12} = 0, \; Q_{21} = -\sum_{i=1}^{m} k_i r_i^\alpha \left(\frac{\ln r_i}{1 + r_i^\alpha} \right)^2, \; Q_{03} = \frac{2\,(m + a_1)}{\beta^3}, \quad (2.39)$$

$$Q_{30} = \frac{2\,(m + a_1 + a_2)}{\alpha^3} - \sum_{i=1}^{m} (\beta k_i + 1)(1 - r_i^\alpha) r_i^\alpha \left(\frac{\ln r_i}{1 + r_i^\alpha} \right)^3. \quad (2.40)$$

Therefore, the approximate Bayes estimates of α and β under the SE and the LINEX loss functions are

$$\widehat{\alpha}_{BS,Lind} = \widetilde{\alpha} + \frac{1}{2}\left[Q_{30}\tau_{11}^2 + 3Q_{21}\tau_{11}\tau_{12} + Q_{03}\tau_{21}\tau_{22}\right], \tag{2.41}$$

$$\widehat{\alpha}_{BL,Lind} = \widetilde{\alpha} - \frac{1}{v}\ln\left[1 + \frac{v}{2}\left(v\tau_{11} - Q_{30}\tau_{11}^2 - 3Q_{21}\tau_{11}\tau_{12} - Q_{03}\tau_{21}\tau_{22}\right)\right], \tag{2.42}$$

$$\widehat{\beta}_{BS,Lind} = \widetilde{\beta} + \frac{1}{2}\left[Q_{30}\tau_{12}\tau_{11} + Q_{21}(\tau_{11}\tau_{22} + 2\tau_{12}^2) + Q_{03}\tau_{22}^2\right], \tag{2.43}$$

$$\begin{aligned}\widehat{\beta}_{BL,Lind} = \widetilde{\beta} - \frac{1}{v}\ln 1 + \frac{v^2}{2}\tau_{22} - \frac{v}{2}Q_{30}\tau_{12}\tau_{11} \\ - \frac{v}{2}Q_{21}(\tau_{11}\tau_{22} + 2\tau_{12}^2) - \frac{v}{2}Q_{03}\tau_{22}^2\end{aligned} \tag{2.44}$$

Notice that all approximate Bayes estimates are evaluated at $(\widetilde{\alpha}, \widetilde{\beta})$.

- MCMC method

In the previous section, the Bayes estimates of α and β are obtained under the SE and the LINEX loss functions by using the Lindley's approximation. Since the exact probability distributions of these estimates are not known, it is difficult to evaluate HPD credible intervals of parameters. For this reason, the MCMC method are used to compute the Bayes estimates of α and β under the SE and the LINEX loss functions as well the HPD credible intervals.

The MCMC method are considered to generate samples from the posterior distributions and then the Bayes estimates of α and β under the SE and the LINEX loss functions are computed. The joint posterior density of α and β is given by equation (2.31). It is easy to see that

$$\beta|\,\alpha, \mathbf{r}, \mathbf{k} \sim Gamma\left(m + a_1 + 1, (\alpha/b_1) + T_\alpha\right), \tag{2.45}$$

and

$$\pi(\alpha|\,\beta,\mathbf{r},\mathbf{k}) \propto \alpha^{m+a_1+a_2}\exp\left\{-\beta\left(\frac{\alpha}{b_1}+T_\alpha\right)\right\}$$

$$\exp\left\{-\alpha\left(\frac{1}{b_2}-\sum_{i=1}^{m}\ln r_i\right)-\sum_{i=1}^{m}\ln(1+r_i^{\alpha})\right\}. \quad (2.46)$$

Therefore, the samples of β can be generated by using the gamma distribution. However, the posterior distribution of α cannot be reduced analytically to well known distribution and therefore it is not possible to sample directly by standard methods. If the posterior density of α is unimodal and roughly symmetric then it is often convenient to approximate it by a normal distribution centered at the mode, (see [Gelman et al., 2003]). Since the posterior density of α is log-concave density (so unimodal) and the posterior density of α is roughly symmetric with respect to mode (by experimentation), we use the Metropolis-Hasting algorithm with the normal proposal distribution to generate a random sample from the posterior density of α. The hybrid Metropolis-Hastings and Gibbs sampling algorithm, which will be used to solve our problem, is suggested by [Tierney, 1994]. This algorithm combines the Metropolis-Hastings with Gibbs sampling scheme under the Gaussian proposal distribution.

- Step 1: Take some initial guess of α and β, say $\alpha^{(0)}$ and $\beta^{(0)}$.
- Step 2: Set $t = 1$.
- Step 3: Generate $\alpha^{(t)}$ from $\pi(\alpha|\,\beta,\mathbf{r},\mathbf{k})$ using the Metropolis-Hastings algorithm with the proposal distribution $q(\alpha) \equiv N(\widetilde{\alpha}, V_{\widetilde{\alpha}})$ where $\widetilde{\alpha}$ is a mode of $\pi(\alpha|\,\beta^{(t-1)},\mathbf{r},\mathbf{k})$ and $V_\alpha = \left(-d^2(\ln\pi(\alpha|\,\beta^{(t-1)},\mathbf{r},\mathbf{k}))/d\alpha^2\right)^{-1}$:

  -Step 3.1: Let $v = \alpha^{(t-1)}$.

  -Step 3.2: Generate w from the proposal distribution q.

  -Step 3.3: Let $p(v,w) = \min\left\{1, \dfrac{\pi(w|\,\beta^{(t-1)},\mathbf{r},\mathbf{k})\;q(v)}{\pi(v|\,\beta^{(t-1)},\mathbf{r},\mathbf{k})\;q(w)}\right\}$.

  -Step 3.4: Generate u from $Uniform(0,1)$. If $u \leq p(v,w)$ then accept the proposal and set $\alpha^{(t)} = w$; otherwise, set $\alpha^{(t)} = v$.
- Step 4: Generate $\beta^{(t)}$ from $Gamma\left(m+a_1+1, (\alpha/b_1)+\sum_{i=1}^{m} k_i\ln(1+r_i^{\alpha(t)})\right)$.
- Step 5: Set $t = t+1$.

- Step 6: Repeat Steps 3-5, N times, and obtain the posterior samples $(\alpha^{(i)}, \beta^{(i)})$, $i = 1, \ldots, N$.

The samples obtained from the algorithm are used to compute the Bayes estimates and to construct the HPD credible intervals. The Bayes estimator of $g \equiv g(\alpha, \beta)$ based on the SE and the LINEX loss function are given, respectively, by

$$\widehat{g}_{BS,MH} = E(g|\,\mathbf{r}, \mathbf{k}) = \frac{1}{N-M} \sum_{i=M+1}^{N-M} g(\alpha^{(i)}, \beta^{(i)}), \tag{2.47}$$

and

$$\widehat{g}_{BL,MH} = -\frac{1}{v} \ln \left[\frac{1}{N-M} \sum_{i=M+1}^{N-M} \exp\left(-v\, g(\alpha^{(i)}, \beta^{(i)})\right) \right], \tag{2.48}$$

where M is the burn-in period.

The HPD $100(1-\gamma)\%$ credible intervals of α and β can be obtained by the method of [Chen and Shao, 1999]. In particular for α:

From MCMC, the sequence $\alpha_1, \ldots, \alpha_N$, are obtained, and ordered as $\alpha_{(1)} < \ldots < \alpha_{(N)}$. The credible intervals are constructed as $\left(\alpha_{(j)}, \alpha_{(j+[N(1-\gamma)])}\right)$ for $j = 1, ..., N - [N(1-\gamma)]$ where $[x]$ denotes the largest integer less than or equal to x. Then, the HPD credible interval of α is that interval which has the shortest length. Similarly, the HPD credible interval of β can also be constructed.

2.3.3. Prediction of Future Record Values

In this section, the problem of prediction and prediction interval for the s th ($s > m$) lower record value are considered using non-Bayesian and Bayesian approaches.

2.3.3.1. Non-Bayesian Prediction Approach

When the first m lower record values are observed, the predictive likelihood function of $Y = R_s$, $s > m$ and the parameters θ is given by [Basak and Balakrishnan, 2003] as

$$L(y,\theta;\mathbf{r}) = \prod_{i=1}^{m} \frac{f(r_i;\theta)}{F(r_i;\theta)} \frac{[H(y;\theta) - H(r_m;\theta)]^{s-m-1}}{\Gamma(s-m)} f(y;\theta). \tag{2.49}$$

where $\theta = (\alpha,\beta)$, $\mathbf{r} = (r_1,..,r_m)$ and $H(y;\theta) = -\ln F(y;\theta)$. Moreover, the likelihood function associated with the sequence $\{R_1, K_1, \dots, R_m, K_m\}$ is given by [Samaniego and Whitaker, 1986] in equation (2.5). Similarly, the predictive likelihood function for the future record R_s based on the sequence $\{R_1, K_1, \dots, R_m, K_m, R_s\}$ is derived, we have

$$L(y,\theta;\mathbf{r},\mathbf{k}) = \prod_{i=1}^{m} f(r_i;\theta) \{1 - F(r_i;\theta)\}^{k_i-1} I_{(-\infty,r_{i-1})}(r_i)$$

$$\frac{[H(y;\theta) - H(r_m;\theta)]^{s-m-1}}{\Gamma(s-m)} f(y;\theta). \tag{2.50}$$

Notice that, $K_m \equiv 1$ is defined for convenience, when the inverse sampling is employed (see [Samaniego and Whitaker, 1986]). The predictive maximum likelihood estimator (PMLE) of (α,β) and the maximum likelihood predictor (MLP) of $Y = R_s$ are obtained by maximizing the logarithm of the predictive likelihood function in equation (2.50) with respect to mentioned parameters.

2.3.3.2. Bayesian Prediction Approach

The prediction and prediction interval of future records based on a Bayesian approach are considered under the SE and the LINEX loss functions. The conditional density function of $Y = R_s, s > m$ given the past m records is

$$f(y|\,\mathbf{r},\alpha,\beta) = \frac{[H(y) - H(r_m)]^{s-m-1}}{\Gamma(s-m)} \frac{f(y)}{F(r_m)}$$

$$= \sum_{j=0}^{s-m-1} \binom{s-m-1}{j} \frac{(-1)^j f(y) [\ln F(y)]^j}{\Gamma(s-m) F(r_m) [\ln F(r_m)]^{-s+m+1+j}}, \tag{2.51}$$

where $0 < y < r_m$. The Bayes predictive density function Y is given by

$$h(y|\,\mathbf{r},\mathbf{k}) = \int_0^\infty \int_0^\infty f(y|\,\mathbf{r},\alpha,\beta)\pi(\alpha,\beta|\,\mathbf{r},\mathbf{k})d\alpha d\beta. \tag{2.52}$$

It is clear that $h(y|\,\mathbf{r},\mathbf{k})$ cannot be expressed in closed form and hence it cannot be computed analytically.

The consistent estimator of $h(y|\,\mathbf{r},\mathbf{k})$ is constructed by using the hybrid Metropolis-Hastings and Gibbs sampling procedure as described in MCMC case. Suppose that $\{(\alpha_i,\beta_i),\ i=1,...,N\}$ are MCMC samples obtained from $\pi(\alpha,\beta|\,\mathbf{r},\mathbf{k})$ using the hybrid Metropolis -Hastings and Gibbs sampling technique. The consistent estimator of $h(y|\,\mathbf{r},\mathbf{k})$ based on the simulation can be obtained as

$$\widehat{h}(y|\,\mathbf{r},\mathbf{k}) = \frac{1}{N}\sum_{i=1}^{N} f(y|\,\mathbf{r},\alpha_i,\beta_i), \tag{2.53}$$

and a consistent estimator of the predictive distribution of $Y = R_s$ based on the simulation, say $H(y|\,\mathbf{r},\mathbf{k})$, can be obtained as

$$\widehat{H}(y|\,\mathbf{r},\mathbf{k}) = \frac{1}{N}\sum_{i=1}^{N} F^*(y|\,\mathbf{r},\alpha_i,\beta_i), \tag{2.54}$$

and $F^*(y|\,\mathbf{r},\alpha,\beta)$ denotes the distribution function corresponding to the density function $f(y|\,\mathbf{r},\alpha,\beta)$ and

$$F^*(y|\,\mathbf{r},\alpha,\beta) = \int_0^y f(t|\,\mathbf{r},\alpha,\beta)dt$$

$$= \sum_{j=0}^{s-m-1} \binom{s-m-1}{j} \frac{[\ln F(r_m)]^{s-m-1-j}\,\Gamma(j+1,-\ln F(y))}{\Gamma(s-m)F(r_m)}, \tag{2.55}$$

where $\Gamma(x,y)$ is an incomplete Gamma function, i.e. $\Gamma(x,y) = \int_y^\infty t^{x-1}e^{-t}dt$. It should be noted that the same MCMC samples $\{(\alpha_i,\beta_i),\ i=1,...,N\}$ can be used to compute $\widehat{h}(y|\,\mathbf{r},\mathbf{k})$ or $\widehat{H}(y|\,\mathbf{r},\mathbf{k})$ for all y.

Then, the point predictor of $Y = R_s$ under the SE loss function is

$$\widehat{Y}_S = \int_0^{r_m} y\widehat{h}(y|\,\mathbf{r},\mathbf{k})dy = \frac{1}{N}\sum_{i=1}^{N}\int_0^{r_m} yf(y|\,\mathbf{r},\alpha_i,\beta_i)dy$$

$$= \frac{1}{N}\sum_{i=1}^{N}\int_0^{r_m} \frac{yf(y|\,;\alpha_i,\beta_i)}{\Gamma(s-m)F(r_{m;}\alpha_i,\beta_i)}\left[\ln\left(\frac{F(r_m;\alpha_i,\beta_i)}{F(y;\alpha_i,\beta_i)}\right)\right]^{s-m-1} dy. \qquad (2.56)$$

The point predictor of $Y = R_s$ under the LINEX loss function is

$$\widehat{Y}_L = -\frac{1}{v}\ln\left\{\frac{1}{N}\sum_{i=1}^{N}\int_0^{r_m} \frac{e^{-vy}f(y;\alpha_i,\beta_i)}{\Gamma(s-m)F(r_{m;}\alpha_i,\beta_i)}\right.$$

$$\left.\left(\ln\left(\frac{F(r_m;\alpha_i,\beta_i)}{F(y;\alpha_i,\beta_i)}\right)\right)^{s-m-1} dy\right\}. \qquad (2.57)$$

For a special case, when $s = m+1$, the conditional density function of $Y = R_s, s > m$ given $\mathbf{r}$ is $f(y|\,\mathbf{r},\alpha,\beta) = f(y)/F(r_m)$. Hence, the distribution function of $f(y|\,\mathbf{r},\alpha,\beta)$ is given by

$$F^*(y|\,\mathbf{r},\alpha,\beta) = \left(\frac{1+r_m^\alpha}{1+y^\alpha}\right)^\beta \frac{(1+y^\alpha)^\beta - 1}{(1+r_m^\alpha)^\beta - 1}. \qquad (2.58)$$

Therefore, $\widehat{h}(y|\,\mathbf{r},\mathbf{k})$, $\widehat{H}(y|\,\mathbf{r},\mathbf{k})$, $\widehat{Y}_S$ and $\widehat{Y}_L$ are obtained from equations (2.53), (2.54), (2.56) and (2.57), respectively by using equations (2.51) and (2.55).

Moreover, a symmetric $100\gamma\%$ prediction interval for Y, can be obtained by solving the following non-linear equations, for the lower bound L and upper bound U,

$$\frac{1+\gamma}{2} = P(Y > L|\,\mathbf{r},\mathbf{k}) = 1 - H(L|\,\mathbf{r},\mathbf{k}) \Rightarrow H(L|\,\mathbf{r},\mathbf{k}) = \frac{1-\gamma}{2}, \qquad (2.59)$$

$$\frac{1-\gamma}{2} = P(Y > U|\,\mathbf{r},\mathbf{k}) = 1 - H(U|\,\mathbf{r},\mathbf{k}) \Rightarrow H(U|\,\mathbf{r},\mathbf{k}) = \frac{1+\gamma}{2}. \qquad (2.60)$$

These equations can be easily solved by using the Newton-Raphson method.

2.3.4. Simulation Study

In this section, some numerical results are presented to compare the performance of the different methods for different sample sizes and different priors. The performances of the point estimators and predictors are compared by using the estimated risk (ER) and mean square predictor error (MSPE), respectively. The performances of the confidence, credible and prediction intervals are compared by using average confidence lengths and cps. The ER of θ, when θ is estimated by $\widehat{\theta}$, is given by

$$ER(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\theta}_i - \theta_i\right)^2, \tag{2.61}$$

under the SE loss function. Moreover, the estimated risk of θ under the LINEX loss function is given by

$$ER(\theta) = \frac{1}{N}\sum_{i=1}^{N}\left(e^{v\left(\widehat{\theta}_i-\theta_i\right)} - v\left(\widehat{\theta}_i - \theta_i\right) - 1\right), \tag{2.62}$$

where N is the number of replication. Similarly, the MSPEs can be computed with respect to SE and LINEX loss functions. All of the computations are performed by using Matlab R2010a. All the results are based on 5000 replications.

In Table 2.1, the ML and Bayes estimates under the SE and the LINEX ($v = -2, -1, 1$ and 2) loss functions with their corresponding ERs are listed for β when α is known ($\alpha = 2$), $\beta = 2.0092$ and the prior parameters of β are chosen to be $(a_1, b_1) = (3, 2)$. Since the exact distribution of the MLE of β is known, the 95% exact confidence intervals are easily constructed. Moreover, the 95% Bayesian credible interval for β which is obtained by using the posterior distribution of β are listed. From Table 2.1, the average ERs of β decrease as the sample size increases in all cases, as expected. The Bayes estimates under the SE and the LINEX loss functions have smaller ER than that of MLEs. The average lengths of the intervals decrease as the sample size increases. The lengths of the Bayesian credible intervals are smaller than that of exact confidence intervals. Also, the coverage probabilities are quite close to nominal level 95%.

In Tables 2.2 and 2.3, the ML and Bayes estimates under the SE and LINEX loss functions with their corresponding ERs are listed for α and β when $(\alpha, \beta) = (5.0381, 1.0564)$, $(5.9303, 1.4606)$, respectively. In the Bayesian case, two different bivariate priors are considered as follows: Prior 1: $(a_1, b_1) = (0.2, 2)$, $(a_2, b_2) = (5, 1)$ and Prior 2: $(a_1, b_1) = (1, 3)$, $(a_2, b_2) = (6, 1)$.The Bayes estimates are computed by using Lindley's approximation and MCMC method under SE and LINEX ($v = -2, -1, 1$ and 2) loss functions for different prior parameters. Moreover, the 95% asymptotic and HPD credible intervals with their coverage probabilities are listed. From Tables 2.2 and 2.3, the average ERs of α and β generally decrease as the sample size increases. The ERs of Bayes estimates under the SE loss function are smaller than that of MLEs. But under the LINEX loss function ERs of the Bayes and ML estimates can not be compared. On the other hand, the ERs of the Bayes estimates for α and β based on the Lindley's approximation and MCMC methods are close to each other under the SE and the LINEX loss functions when $v > 0$. The ERs of the Bayes estimates under the LINEX loss function close each other as the sample size increases when $v < 0$. The average lengths of the intervals decrease as the sample size increases. Furthermore, the average lengths of the Bayesian credible intervals are smaller than that of the asymptotic confidence intervals.

Table 2.1: Results for the true value of $\beta = 2.0092$, $(a_1, b_1) = (3, 2)$ and α is known.

	Bayes estimates						
		LINEX				Exact MLE	Bayesian
MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	confidence interval	credible interval
			$m=5$				
2.8388	2.1144	3.0982	2.4603	1.8802	1.7063	(0.927,5.8147)	(0.9668,3.7032)
5.4084	0.5554	2.0372	0.3545	0.2307	0.8004	0.9406	0.9438
			$m=10$				
2.3723	2.0860	2.5902	2.2948	1.9255	1.7961	(1.1376,4.0530)	(1.1404,3.3124)
1.0663	0.3572	0.9697	0.2046	0.1586	0.5731	0.9432	0.9464
			$m=15$				
2.1987	2.0400	2.3663	2.1838	1.9216	1.8215	(1.2306,3.4431)	(1.2282,3.0544)
0.5330	0.2654	0.7007	0.1497	0.1198	0.4372	0.9464	0.9496
			$m=20$				
2.1454	2.0295	2.2761	2.1415	1.9338	1.8505	(1.3105,3.1828)	(1.3003,2.9184)
0.3249	0.1890	0.4378	0.1007	0.0886	0.3318	0.9464	0.9528

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

In the MCMC case, five MCMC chains are run with fairly different initial values and generated 10000 iterations for each chain. To diminish the effect of the starting distribution, the first half of each sequence are discarded and focus on the second half. To provide relatively independent samples for improvement of prediction accuracy, the Bayesian MCMC estimates are calculated by the means of every 5^{th} sampled values after discarding the first half of the chains (see [Gelman et al., 2003]). In our case, the scale factor value of the MCMC estimators are found below 1.1 which is an acceptable value for their convergency.

In Tables 2.4-2.7, the point predictors for $Y = R_{m+1}$ which are computed (based on 5000 replications) by using non-Bayesian and Bayesian (with respect to the SE and LINEX loss functions) methods and the 95% prediction intervals are listed when $(\alpha, \beta) = (5, 1)$. In the Bayesian case, two different bivariate priors are considered as follows: Prior 1: $(a_1, b_1) = (0.2, 2)$, $(a_2, b_2) = (5, 1)$ and Prior 3: $(a_1, b_1) = (0.005, 2.4/1.005)$, $(a_2, b_2) = (2, 2.5)$. To observe the sensitivity of the predictors with respect to different informative priors, these priors are chosen with same means but different variances. Notice that the variances of the Prior 1 are smaller than that of Prior 3. Moreover, to observe the effect of the inter-record times, the point predictors and prediction intervals are also obtained based on only lower record values (without taking inter-record times into consideration). The results based on only lower record values are given in Tables 2.5 and 2.7. From Tables 2.4-2.7, the average MSPEs of the point predictors decrease as the sample size increases in all cases. Also, the average lengths of the prediction intervals decrease as the sample size increases and their coverage probabilities are quite close to nominal level 95%. Moreover, the MSPEs of the Bayesian point predictors which are obtained by using lower record values and lower record values with their inter-record times are almost the same. However, the average lengths of the prediction intervals based on lower record values with their inter-record times are smaller than the one based on only lower record values. Furthermore, Prior 1 can be considered as a good informative prior, because its variance smaller than Prior 3. The MSPEs and the average lengths of the prediction intervals using Prior 1 in almost all cases are smaller than those using Prior 3 based on lower record values with their inter-record times. Similar results based on just lower record values are observed only for sufficiently large m ($m \geq 12$). Therefore,

Table 2.2: Results for the true values of $(\alpha, \beta) = (5.0381, 1.0564)$ using Prior 1.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic	HPD
			LINEX					LINEX				confidence interval	credible interval
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$		
							$m=5$						
α	5.1749	4.2383	4.5078	4.4710	3.8270	3.6228	4.3111	4.9824	4.6699	4.0024	3.7770	-	(3.0485,5.7693)
	2.8453	1.8980	8223.1	5.2258	0.7183	2.1517	1.8668	5552.6	4.9275	0.6593	2.2144	-	0.7920
β	6.2200	0.9462	1.0089	0.9924	0.8552	0.7722	0.9715	1.3718	1.1475	0.8707	0.7995	-	(0.2715,1.8043)
	10534	0.2602	10.267	0.2337	0.1021	0.3490	0.2705	0.7314	1.1669	0.1067	0.3549	-	0.9132
							$m=10$						
α	4.9910	4.6617	4.7748	4.7401	4.5432	4.4414	4.7068	4.9012	4.8025	4.6216	4.5491	-	(4.0597,5.4217)
	0.4365	0.4504	6.7544	0.3584	0.2007	0.7138	0.4329	2.2028	0.2897	0.1925	1.4041	-	0.8094
β	1.6571	0.8832	0.9326	0.9193	0.8116	0.7409	0.9212	1.2279	1.0524	0.8412	0.7833	-	(0.3136,1.6547)
	30.1092	0.2519	23.0779	0.2647	0.1002	0.3454	0.2616	0.7221	0.1489	0.1117	0.4307	-	0.8770
							$m=15$						
α	5.0425	4.7917	4.8919	4.8565	4.7047	4.6283	4.8183	4.9673	4.8914	4.7520	4.6942	(4.1775,5.9076)	(4.2387,5.4485)
	0.2640	0.2586	1.5974	0.1654	0.1226	0.4558	0.2482	0.7816	0.1437	0.1129	0.4188	0.9418	0.8562
β	1.3302	0.9335	0.9884	0.9714	0.8679	0.8083	0.9556	1.1891	1.0587	0.8893	0.8396	(0.1664,2.5825)	(0.3994,1.6087)
	1.9021	0.1901	2.8452	0.1488	0.0791	0.2764	0.1967	0.5146	0.1117	0.0970	0.4921	0.9554	0.9008
							$m=18$						
α	5.0549	4.8349	4.9286	4.8940	4.7587	4.6913	4.8556	4.9885	4.9209	4.7960	4.7435	(4.2629,5.8469)	(4.3042,5.4505)
	0.2087	0.2086	1.4620	0.1283	0.1001	0.3762	0.2014	0.4820	0.1097	0.0936	0.3557	0.9456	0.8662
β	1.2815	0.9552	1.0106	0.9926	0.8939	0.8397	0.9747	1.1785	1.0659	0.9143	0.8682	(0.2487,2.3559)	(0.4428,1.5922)
	1.0519	0.1693	2.2039	0.1290	0.0714	0.2515	0.1785	0.4812	0.1050	0.0787	0.2926	0.9610	0.9080

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 2.3: Results for the true values of $(\alpha, \beta) = (5.9303, 1.4606)$ using Prior 2.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic	HPD
			LINEX					LINEX				confidence interval	credible interval
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$		
								$m=5$					
α	6.1491	5.3199	5.6354	5.6006	4.8277	4.6169	5.3937	6.1484	5.8093	5.0229	4.7581	-	(4.0094,6.9581)
	3.6422	1.7173	1801.5	3.4968	0.6863	2.0739	1.7406	623.6481	3.0135	0.6710	2.6727	-	0.8036
β	10.6931	1.4365	1.5424	1.5220	1.2383	1.1029	1.4741	2.2781	1.8449	1.2796	1.1514	-	(0.4617,2.6580)
	83613	0.5273	2.6416	0.2684	0.1691	0.5760	0.5563	2.2168	0.4304	0.1916	0.6156	-	0.9506
								$m=10$					
α	5.9252	5.6696	5.8057	5.7656	5.5243	5.4066	5.7157	5.9469	5.8328	5.6070	5.5128	-	(4.9749,6.5171)
	0.6111	0.3720	2.0430	0.2368	0.1802	0.6610	0.3874	1.1007	0.2148	0.1911	0.9145	-	0.8344
β	2.6745	1.3502	1.4351	1.4186	1.1855	1.0613	1.4048	2.0486	1.6951	1.2450	1.1373	-	(0.5091,2.4648)
	152.2277	0.4945	4.4929	0.2909	0.1654	0.5646	0.5026	1.1362	0.3024	0.2227	0.9157	-	0.9242
								$m=15$					
α	5.9424	5.7606	5.8821	5.8409	5.6512	5.5592	5.7844	5.9691	5.8775	5.6975	5.6209	(4.9221,6.9627)	(5.1146,6.5056)
	0.3557	0.2505	1.0702	0.1483	0.1223	0.4618	0.2549	0.6632	0.1402	0.1214	0.4824	0.9438	0.8736
β	1.9264	1.3771	1.4663	1.4444	1.2358	1.1333	1.4137	1.9075	1.6396	1.2819	1.1897	(0.1766,3.9052)	(0.6081,2.3483)
	5.7273	0.4036	2.3761	0.2258	0.1424	0.5022	0.4156	1.3688	0.4049	0.1732	0.7145	0.9506	0.9206
								$m=18$					
α	5.9540	5.7933	5.9076	5.8670	5.6967	5.6142	5.8131	5.9792	5.8967	5.7347	5.6650	(5.0210,6.8870)	(5.1752,6.4974)
	0.2855	0.2150	0.6945	0.1200	0.1057	0.4013	0.2166	0.4911	0.1136	0.1047	0.4124	0.9406	0.8794
β	1.8461	1.4002	1.4897	1.4663	1.2670	1.1730	1.4319	1.8695	1.6345	1.3099	1.2235	(0.2820,3.5072)	(0.6624,2.3178)
	3.3052	0.3506	1.9968	0.1881	0.1246	0.4355	0.3684	0.9759	0.2312	0.1459	0.5144	0.9616	0.9240

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 2.4: Predictions based on lower records with inter-record times using Prior 1.

r_m	Bayes point predictors					Prediction interval
		LINEX				
r_{m+1}	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	Length/cp
			$m=5$			
0.3738	0.304272	0.307671	0.306013	0.302443	0.300522	(0.157121,0.37166)
0.2755	0.006420	0.012191	0.003129	0.003290	0.013475	0.214545/0.9314
			$m=7$			
0.2518	0.206737	0.208197	0.207479	0.205972	0.205181	(0.111397,0.250464)
0.1867	0.003001	0.005747	0.001468	0.001534	0.006270	0.139067/0.9276
			$m=12$			
0.1635	0.135226	0.135801	0.135516	0.134929	0.134627	(0.075340,0.162595)
0.1238	0.001083	0.002114	0.000535	0.000548	0.002217	0.087255/0.9272
			$m=15$			
0.1607	0.133240	0.133782	0.133513	0.132961	0.132677	(0.074895,0.159849)
0.1242	0.000890	0.001741	0.000440	0.000450	0.001818	0.084953/0.9394

Notes: First column: The first row represents the average of the r_m th record values and second row represents the average of the true values (r_{m+1}) which we want to predict. Last column, the first row represents a 95% prediction interval (PI) and second row represents their lengths and cp's. For the others, the first row represents the average predictors and second row represents corresponding MSPEs for each choice of m.

from these results we can infer that using the record values with their corresponding inter-record times is preferable to the results based on only record values. On the other hand, we empirically see that the MLP of $Y = R_{m+1}$ is very big compared to the last record values which violates $\widehat{Y}_{MLP} < r_m$. Hence, the MLP of Y was not listed when the true values of $\alpha > 3$ and $\beta < 25$ (contains the case in Tables 2.4-2.7).

A real-life data set which the amount of rainfall (in inches) recorded at the Los Angeles Civic Center in February from 1943 to 2006 (see the website of Los Angeles Almanac: www.laalmanac.com/weather/we08aa.htm) are given in Table 2.8. To see if the underlying distribution follows the Burr Type XII, we compute the Kolmogorov-Smirnov distances between the empirical distribution and the fitted distribution functions based on the complete data set. These distances are 0.1112, 0.1572 and 0.1567 based on parameter estimations by using ML and Bayes (Lindley approximation and MCMC method under SE loss function) estimates, respectively. The associated p values for the Bayes cases are $0.1 < p < 0.2$ and for the ML case p value is greater than 0.2. This indicates that Burr Type XII model provides an adequate fit for data. The first 7 lower records (among 8 lower record values) with their corresponding inter-record times are used for the estimates of (α, β) and

Table 2.5: Predictions based on only lower records using Prior 1.

r_m / r_{m+1}	Bayes point predictors					Prediction interval length/cp
	SEL	LINEX				
		$v=-2$	$v=-1$	$v=1$	$v=2$	
			$m=5$			
0.3738	0.304120	0.307625	0.305937	0.302191	0.300212	(0.150699,0.371640)
0.2755	0.006759	0.012722	0.003278	0.003482	0.014346	0.220941/0.9315
			$m=7$			
0.2518	0.198408	0.200342	0.199436	0.197325	0.196321	(0.090187,0.250146)
0.1867	0.002964	0.005662	0.001448	0.001518	0.006216	0.159958/0.9480
			$m=12$			
0.1635	0.119685	0.120807	0.120257	0.119101	0.118519	(0.042573,0.161937)
0.1238	0.000999	0.001960	0.000495	0.000504	0.002036	0.119365/0.9610
			$m=15$			
0.1607	0.115616	0.116762	0.116192	0.115030	0.114429	(0.038611,0.159088)
0.1242	0.000896	0.001752	0.000443	0.000453	0.001831	0.120477/0.9648

Notes: First column: The first row represents the average of the r_m th record values and second row represents the average of the true values (r_{m+1}) which we want to predict. Last column, the first row represents a 95% PI and second row represents their lengths and cp's. For the others, the first row represents the average predictors and second row represents corresponding MSPEs for each choice of m.

Table 2.6: Predictions based on lower records with inter-record times using Prior 3.

r_m / r_{m+1}	Bayes point predictors					Prediction interval length/cp
	SEL	LINEX				
		$v=-2$	$v=-1$	$v=1$	$v=2$	
			$m=5$			
0.3738	0.303292	0.306801	0.305090	0.301402	0.299415	(0.153746,0.371641)
0.2755	0.006454	0.012241	0.003143	0.003310	0.013569	0.217895/0.9340
			$m=7$			
0.2518	0.206606	0.208076	0.207353	0.205835	0.205038	(0.110915,0.250461)
0.1867	0.003005	0.005753	0.001470	0.001536	0.006279	0.139546/0.9274
			$m=12$			
0.1635	0.135279	0.135853	0.135569	0.134984	0.134683	(0.075452,0.162597)
0.1238	0.001085	0.002118	0.000536	0.000549	0.002222	0.087145/0.9272
			$m=15$			
0.1607	0.133310	0.133850	0.133583	0.133033	0.132750	(0.075057,0.159852)
0.1242	0.000892	0.001745	0.000441	0.000451	0.001822	0.084794/0.9384

Notes: First column: The first row represents the average of the r_m th record values and second row represents the average of the true values (r_{m+1}) which we want to predict. Last column, the first row represents a 95% PI and second row represents their lengths and cp's. For the others, the first row represents the average predictors and second row represents corresponding MSPEs for each choice of m.

prediction of R_8. In the Bayesian case, we need to determine the hyperparameters. The method of moments are used as in Section 4.2.4 to obtain hyperparameters a_1, a_2,

Table 2.7: Predictions based on only lower records using Prior 3.

r_m / r_{m+1}	Bayes point predictors					Prediction interval length/cp
		LINEX				
	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	
			$m=5$			
0.3738	0.300983	0.304911	0.303292	0.298323	0.296371	(0.140628,0.371344)
0.2755	0.006793	0.12753	0.003289	0.003508	0.014474	0.230715/0.9348
			$m=7$			
0.2518	0.194859	0.197305	0.196724	0.192646	0.191954	(0.082255,0.249968)
0.1867	0.002953	0.005623	0.001434	0.001526	0.006230	0.167713/0.9502
			$m=12$			
0.1635	0.116909	0.118231	0.117754	0.116003	0.115493	(0.037944,0.161798)
0.1238	0.001039	0.002029	0.000510	0.000530	0.002133	0.123854/0.9616
			$m=15$			
0.1607	0.112941	0.114185	0.113578	0.112291	0.111652	(0.034418,0.158947)
0.1242	0.000956	0.001865	0.000472	0.000484	0.001959	0.124529/0.9644

Notes: First column: The first row represents the average of the r_m th record values and second row represents the average of the true values (r_{m+1}) which we want to predict. Last column, the first row represents a 95% PI and second row represents their lengths and cp's. For the others, the first row represents the average predictors and second row represents corresponding MSPEs for each choice of m.

b_1 and b_2. Therefore, the hyperparameters are obtained as $a_1 = 1.1685$, $a_2 = 0.9543$, $b_1 = 0.7750$ and $b_2 = 0.9985$. The Bayes estimates of the parameters, Bayes point predictors and Bayesian prediction interval of R_8 are obtained by using MCMC method. The findings based on these approaches are listed in Table 2.9. It can be observed that the prediction of R_8 are satisfactory under Bayesian approach.

Table 2.8: Record data from a set of rainfall data during February from 1943 to 2006.

i	1	2	3	4	5	6	7	8
R_i	3.07	1.52	0.86	0.63	0.33	0.15	0.11	0.08
K_i	3	1	5	1	8	6	1	-

2.3.5. Conclusions

In this section, firstly the non-Bayesian and Bayesian point estimates as well as confidence intervals for the unknown parameters of Burr Type XII distribution are considered based on the lower record values with their corresponding inter-record times. The ML estimates of the unknown parameters are derived under the inverse

Table 2.9: Results by using bivariate prior for α and β.

Method	α	β	Point Predictor of R_8
MLE	1.4089	0.5991	-
PMLE/MLP	1.3910	0.6743	0.3164
SEL	1.4524	0.6655	0.0634
LINEX($v = -2$)	1.6127	0.7191	0.0643
LINEX($v = -1$)	1.5257	0.6910	0.0638
LINEX($v = 1$)	1.3877	0.6423	0.0629
LINEX($v = 2$)	1.3293	0.6211	0.0624
HPD credible Int.	(0.7186,2.1370)	(0.2921,1.1150)	-
Prediction Int.	-	-	(0.0058,0.1080)

sampling scheme. The Lindley's approximation and MCMC methods are used to get the Bayes estimates under the SE and LINEX loss function for bivariate prior case. Monte Carlo simulation reveals out that the ERs of the Bayes estimates are smaller than that of MLEs under the SE loss function. However, the ERs for the LINEX loss function depend on the asymmetry parameter v. The average length of the HPD credible intervals are smaller than that of the asymptotic intervals.

Secondly, non-Bayesian and Bayesian point predictors as well as prediction intervals for the future lower record values are considered. The point predictors and prediction intervals of the future lower record values are computed based on only the lower record values and the lower record values with their corresponding inter-record times. Therefore, we can see the effect of considering the inter-record times for the predictors. It is observed that using the inter-record times in the prediction case decrease the average lengths of the prediction intervals with reasonable coverage probabilities. On the other hand, the MSPEs of the point predictors are almost the same for both cases. As a result, using the record values with their corresponding inter-record times instead of just using the record values is suggested.

2.4. Estimation of The Reliability Based on Record Values

In the literature, many papers are available for an estimate of the reliability based on a random sample or records sample. When the X and Y are independent and follow the Burr Type III, X and XII, generalized exponential, Weibull, generalized logistic and Kumaraswamy distributions, the estimation of R based on a random sample were studied by [Mokhlis, 2005], [Ahmad et al., 1997], [Awad and Gharraf, 1986], [Kundu and Gupta, 2005], [Kundu and Gupta, 2006], [Asgharzadeh et al., 2013], [Nadar et al., 2014], respectively. When the X and Y are independent and follow the one and two parameters generalized exponential, Weibull, exponentiated gumbel, one and two parameters exponential distributions, the classical and Bayesian estimates of R based on records were considered by [Baklizi, 2008], [Asgharzadeh et al., 2014], [Baklizi, 2012], [Tarvirdizade, 2013], [Baklizi, 2014], respectively.

The ML, uniformly minimum variance unbiased (UMVU) and Bayes estimates of the stress-strength reliability based on complete sample were obtained by [Awad and Gharraf, 1986] when the second shape parameter is common. They used the gamma priors for the first shape parameters and constant number for the common shape parameter in the Bayesian case. Recently, the ML, UMVU and Bayes estimates of the stress-strength reliability were discussed by [Panahi and Asadi, 2010] when the second shape parameter is common and known. The multicomponent stress-strength reliability was considered by [Web 1, 2015]. However, the statistical inference for the stress-strength reliability of the Burr Type XII distribution based on record values has not been considered up to now.

The main purpose of this section is to improve the inference procedures for the stress-strength reliability based on upper record values while the measurements follow the two-parameter Burr Type XII distribution when the first shape parameters are common. When the first shape parameter α is unknown, the ML and Bayes estimates, as well as asymptotic confidence and HPD credible intervals are derived. When α is known, different estimates, namely ML, UMVU, Bayes and empirical Bayes estimates, are obtained. The Bayes estimates of R under the SE and LINEX loss functions are derived in closed forms for informative and non informative prior cases. It is also obtained by using Lindley's approximation and MCMC method. The exact and other Bayes estimates are compared in terms of ER by the Monte Carlo simulations. Also, the exact and asymptotic confidence intervals, as well as Bayesian, empirical Bayesian and HPD credible intervals are constructed for R.

2.4.1. Estimation of R When α Is Common and Unknown

The ML estimates, its existence and uniqueness, asymptotic confidence intervals, as well as Bayes estimates and Bayesian credible interval for R are obtained when the first shape parameter α is common for the distributions of X and Y.

2.4.1.1. ML Estimation of R

Let $X \sim Burr(\alpha, \beta_1)$ and $Y \sim Burr(\alpha, \beta_2)$ are independent random variables. Then, the reliability $R = P(X < Y)$ is

$$R = P(X < Y) = \int_0^{\infty} f_Y(y) P(X < Y \mid Y = y) dy$$

$$= \frac{\beta_1}{\beta_1 + \beta_2}. \tag{2.63}$$

The estimate of R are considered based on upper record data on both variables. Let $R_1, \ldots, R_n$ be a set of upper records from $Burr(\alpha, \beta_1)$ and $S_1, \ldots, S_m$ be a set of upper records from $Burr(\alpha, \beta_2)$ independently from the first sample. The likelihood functions based on records are given by, see [Arnold et al., 1998],

$$L_1(\beta_1, \alpha \,|\underline{r}) = f(r_n; \alpha, \beta_1) \prod_{i=1}^{n-1} \frac{f(r_i; \alpha, \beta_1)}{1 - F(r_i; \alpha, \beta_1)}, \; 0 < r_1 < \ldots < r_n, \tag{2.64}$$

$$L_2(\beta_2, \alpha \,|\underline{s}) = g(s_m; \alpha, \beta_2) \prod_{j=1}^{m-1} \frac{g(s_j; \alpha, \beta_2)}{1 - G(s_j; \alpha, \beta_2)}, \; 0 < s_1 < \ldots < s_m, \tag{2.65}$$

where $\underline{r} = (r_1, \ldots, r_n)$, $\underline{s} = (s_1, \ldots, s_m)$, f and F are the pdf and cdf of X follows $Burr(\alpha, \beta_1)$, respectively and g and G are the pdf and cdf of Y follows $Burr(\alpha, \beta_2)$, respectively. Then, the joint likelihood function of $(\beta_1, \beta_2, \alpha)$ given $(\underline{r}, \underline{s})$ is given by

$$L(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s}) = h_1(\underline{r}; \alpha) h_2(\underline{s}; \alpha) \alpha^{n+m} \beta_1^n \beta_2^m e^{-\beta_1 T_1(r_n;\alpha)} e^{-\beta_2 T_2(s_m;\alpha)}, \tag{2.66}$$

where

$$h_1(\underline{r};\alpha) = \prod_{i=1}^{n}\frac{r_i^{\alpha-1}}{1+r_i^{\alpha}},\ h_2(\underline{s};\alpha)=\prod_{j=1}^{m}\frac{s_j^{\alpha-1}}{1+s_j^{\alpha}}, \tag{2.67}$$

$$T_1(r_n;\alpha) = \ln(1+r_n^{\alpha}),\ T_2(s_m;\alpha)=\ln(1+s_m^{\alpha}). \tag{2.68}$$

The joint log-likelihood function is

$$l(\beta_1,\beta_2,\alpha\,|\underline{r},\underline{s}) = \ln h_1(\underline{r};\alpha)+\ln h_2(\underline{s};\alpha)+(n+m)\ln\alpha$$

$$+\,n\ln\beta_1+m\ln\beta_2-\beta_1T_1(r_n;\alpha)-\beta_2T_2(s_m;\alpha). \tag{2.69}$$

The ML estimates of β_1,β_2 and α, say $\widehat{\beta}_1$, $\widehat{\beta}_2$ and $\widehat{\alpha}$, are given by

$$\widehat{\beta}_1 = \frac{n}{T_1(r_n;\widehat{\alpha})}, \tag{2.70}$$

$$\widehat{\beta}_2 = \frac{m}{T_2(s_m;\widehat{\alpha})}, \tag{2.71}$$

and $\widehat{\alpha}$ is the solution of the following non-linear equation

$$\frac{n+m}{\alpha}+\sum_{i=1}^{n}\frac{\ln r_i}{1+r_i^{\alpha}}-\frac{nr_n^{\alpha}\ln r_n/(1+r_n^{a})}{\ln(1+r_n^{\alpha})}$$

$$+\sum_{j=1}^{m}\frac{\ln s_j}{1+s_j^{\alpha}}-\frac{ms_m^{\alpha}\ln s_m/(1+s_m^{\alpha})}{\ln(1+s_m^{\alpha})}=0. \tag{2.72}$$

Therefore, $\widehat{\alpha}$ can be obtained as a solution of the non-linear equation of the form $h(\alpha)=\alpha$ where

$$h(a) = -(n+m)\left[\sum_{i=1}^{n}\frac{\ln r_i}{1+r_i^{\alpha}}-\frac{nr_n^{\alpha}\ln r_n/(1+r_n^{a})}{\ln(1+r_n^{\alpha})}\right.$$

$$\left.+\sum_{j=1}^{m}\frac{\ln s_j}{1+s_j^{\alpha}}-\frac{ms_m^{\alpha}\ln s_m/(1+s_m^{\alpha})}{\ln(1+s_m^{\alpha})}\right]^{-1}. \tag{2.73}$$

Since, $\widehat{\alpha}$ is a fixed point solution of the non-linear equation (2.73), its value can be obtained using an iterative scheme as: $\alpha_{(j+1)} = h(\alpha_{(j)})$, where $\alpha_{(j)}$ is the j^{th} iterate of $\widehat{\alpha}$. The iteration procedure should be stopped when $\left|\alpha_{(j+1)} - \alpha_{(j)}\right|$ is sufficiently small. After $\widehat{\alpha}$ is obtained, $\widehat{\beta}_1$ and $\widehat{\beta}_2$ can be obtained from (2.70) and (2.71), respectively. Therefore, the MLE of R, say $\widehat{R}$, is given as

$$\widehat{R} = \frac{\widehat{\beta}_1}{\widehat{\beta}_1 + \widehat{\beta}_2}. \tag{2.74}$$

Next, the existence and uniqueness of the ML estimates of the parameters β_1, β_2 and α are established. We present the following lemma that will be used in proof of Theorem 2.2.

Lemma 2.2: Let

$$w(x) = [\ln(1+x)]^2 + \xi^2(x)\left[\frac{\ln(1+x)}{x} - 1\right], \tag{2.75}$$

where $\xi(x) = x\ln(x)/(1+x)$. Then $w(x) \geq 0$ for $x \geq 0$.

Proof 2.2: For a proof, one may refer to [Ghitany and Al-Awadhi, 2002]. ■

Theorem 2.2: The ML estimates of the parameters β_1, β_2 and α are unique, with $\widehat{\beta}_1 = n/T_1(r_n;\widehat{\alpha})$, $\widehat{\beta}_2 = m/T_2(s_m;\widehat{\alpha})$ where $\widehat{\alpha}$ is the solution of the non-linear equation

$$\begin{aligned} G(\alpha) \equiv \frac{n+m}{\alpha} + \sum_{i=1}^{n}\frac{\ln r_i}{1+r_i^{\alpha}} - \frac{nr_n^{\alpha}\ln r_n/(1+r_n^{a})}{\ln(1+r_n^{\alpha})} \\ + \sum_{j=1}^{m}\frac{\ln s_j}{1+s_j^{\alpha}} - \frac{ms_m^{\alpha}\ln s_m/(1+s_m^{\alpha})}{\ln(1+s_m^{\alpha})} = 0, \end{aligned} \tag{2.76}$$

if at least one of the r_i, $i = 1, ..., n$ (or s_j, $j = 1, ..., m$) is less than unity.

Proof 2.2: We have, $G(0) \equiv \lim_{\alpha\to 0} G(\alpha)$,

$$G(0) = \lim_{\alpha\to 0}\frac{n+m}{\alpha} + \sum_{i=1}^{n}\frac{\ln r_i}{2} + \sum_{j=1}^{m}\frac{\ln s_j}{2} - \frac{n\ln r_n}{2\ln 2} - \frac{m\ln s_m}{2\ln 2} = \infty. \tag{2.77}$$

Let

$$G_1(\alpha;\underline{r}) = \frac{n}{\alpha} + \sum_{i=1}^{n} \frac{\ln r_i}{1+r_i^\alpha} - \frac{n r_n^\alpha \ln r_n/(1+r_n^a)}{\ln(1+r_n^\alpha)}, \tag{2.78}$$

and

$$G_2(\alpha;\underline{s}) = \frac{m}{\alpha} + \sum_{j=1}^{m} \frac{\ln s_j}{1+s_j^\alpha} - \frac{m s_m^\alpha \ln s_m/(1+s_m^a)}{\ln(1+s_m^\alpha)}. \tag{2.79}$$

Then, $G(\alpha) = G_1(\alpha;\underline{r}) + G_2(\alpha;\underline{s})$. *Firstly, the limit of* $G_1(\alpha;\underline{r})$ *as* $\alpha \to \infty$ *is considered.*

i) If r_n *is less than unity, that is* $r_i < 1$, $i = 1, ..., n$, *then,* $G_1(\infty;\underline{r}) \equiv \lim_{\alpha\to\infty} G_1(\alpha;\underline{r})$,

$$G_1(\infty;\underline{r}) = \lim_{\alpha\to\infty}\left(\frac{n}{\alpha} + \sum_{i=1}^{n} \frac{\ln r_i}{1+r_i^\alpha} - \frac{n \ln r_n/(1+r_n^\alpha)}{\ln(1+r_n^\alpha)/r_n^\alpha}\right) \tag{2.80}$$

$$= \sum_{i=1}^{n} (\ln r_i - \ln r_n) < 0. \tag{2.81}$$

ii) If only r_n *is greater than or equal to unity, that is* $r_n \geq 1$ *and* $r_i < 1$, $i = 1, ..., n-1$, *then*

$$G_1(\infty;\underline{r}) = \lim_{\alpha\to\infty}\left(\frac{n}{\alpha} + \sum_{i=1}^{n-1} \frac{\ln r_i}{1+r_i^\alpha} + \frac{\ln r_n}{1+r_n^\alpha} - \frac{n r_n^\alpha \ln r_n/(1+r_n^\alpha)}{\ln(1+r_n^\alpha)}\right) \tag{2.82}$$

$$= \sum_{i=1}^{n-1} \ln r_i < 0. \tag{2.83}$$

iii) If r_n *and some* r_i *record values are greater than unity and some* r_i *record values are less than unity, that is* $r_n > 1$ *and* $r_i > 1$, $i = p, ..., t$, $1 < p \leq t < n$, *then*

$$G_1(\infty;\underline{r}) = \lim_{\alpha\to\infty}\left\{\frac{n}{\alpha} + \sum_{i=1 (r_i<1)}^{n} \frac{\ln r_i}{1+r_i^\alpha} + \sum_{i=1 (r_i>1)}^{n} \frac{\ln r_i}{1+r_i^\alpha}\right.$$

$$\left. + \sum_{i=1 (r_i>1)}^{n} \frac{\ln r_i}{1+r_i^\alpha} - \frac{n r_n^\alpha \ln r_n/(1+r_n^\alpha)}{\ln(1+r_n^\alpha)}\right\} = \sum_{i=1 (r_i<1)}^{n} \ln r_i < 0. \tag{2.84}$$

When the conditions given in i)-iii) holds for s_j, $j = 1, ..., m$, $G_2(\alpha; \underline{s}) < 0$ *as* $\alpha \to \infty$. *So that, the limit of* $G(\alpha) = G_1(\alpha; \underline{r}) + G_2(\alpha; \underline{s}) < 0$ *as* $\alpha \to \infty$ *when* r_i, $i = 1, ..., n$ *and* s_j, $j = 1, ..., m$ *satisfy any of the conditions given in i)-iii).*

Next, we need to show the limit of $G(\alpha) < 0$ *as* $\alpha \to \infty$ *for* $s_j > 1$, $j = 1, ..., m$ *and when the conditions given i)-iii) holds for* r_i, $i = 1, ..., n$ *(or* $r_i > 1, i = 1, ..., n$ *and when the conditions given i)-iii) holds for* s_j, $j = 1, ..., m$*). In particular, when* $s_j > 1$, $j = 1, ..., m$ *and the conditions given i) holds for* r_i, $i = 1, ..., n$*, we can take* α *large enough, such that* $G_2(\alpha; \underline{s}) \to 0^+$ *and* $G_1(\alpha; \underline{r}) + G_2(\alpha; \underline{s}) < 0$ *as* $\alpha \to \infty$. *Other cases can be obtained similarly.*

Finally, we need to show that there is no solution if all records are greater than unity, that is $r_i > 1$, $i = 1, ..., n$ *and* $s_j > 1$, $j = 1, ..., m$. *If* $r_i > 1$, $i = 1, ..., n$*, then*

$$G_1(\alpha; \underline{r}) < \frac{n}{\alpha} + n \ln r_n \left[\frac{1}{1 + r_1^\alpha} - \frac{r_n^\alpha}{(1 + r_n^\alpha)^2} \right] \to 0^+ \text{ as } \alpha \to \infty. \tag{2.85}$$

Similarly, $G_2(\alpha; \underline{s}) \to 0^+$ *as* $\alpha \to \infty$. *Therefore,* $G(\alpha) \to 0^+$ *as* $\alpha \to \infty$.

Except all records are greater than unity, we obtain that $\lim_{\alpha \to 0} G(\alpha) = \infty$ *and* $\lim_{\alpha \to \infty} G(\alpha) < 0$. *By the intermediate value theorem* $G(\alpha)$ *has at least one root in* $(0, \infty)$. *If it can be shown that* $G(\alpha)$ *is decreasing, then the proof will be completed. It is easily obtained that*

$$\begin{aligned} \frac{\partial G_1(\alpha; \underline{r})}{\partial \alpha} &= -\frac{1}{\alpha^2} \left[n + \sum_{i=1}^{n} \frac{\xi^2(r_i^\alpha)}{r_i^\alpha} + \frac{n \xi^2(r_n^\alpha)}{\ln(1 + r_n^\alpha)} \left(\frac{1}{r_n^\alpha} - \frac{1}{\ln(1 + r_n^\alpha)} \right) \right] \\ &= -\frac{1}{\alpha^2} \left[\sum_{i=1}^{n} \frac{\xi^2(r_i^\alpha)}{r_i^\alpha} + \frac{n}{(\ln(1 + r_n^\alpha))^2} w(r_n^\alpha) \right]. \end{aligned} \tag{2.86}$$

Similarly,

$$\frac{\partial G_2(\alpha; \underline{s})}{\partial \alpha} = -\frac{1}{\alpha^2} \left[\sum_{j=1}^{m} \frac{\xi^2(s_j^\alpha)}{s_j^\alpha} + \frac{n}{(\ln(1 + s_m^\alpha))^2} w(s_m^\alpha) \right]. \tag{2.87}$$

It is clear that $\partial G_1(\alpha; \underline{r}) / \partial \alpha < 0$ *and* $\partial G_2(\alpha; \underline{s}) / \partial \alpha < 0$ *by using Lemma 2.2. Therefore,* $\partial G(\alpha) / \partial \alpha < 0$.

Finally, we will show that the ML estimates of $(\beta_1, \beta_2, \alpha)$ *maximizes the log-likelihood function* $l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s})$. *Let* $H(\beta_1, \beta_2, \alpha)$ *be the Hessian matrix of* $l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s})$ *at* $(\beta_1, \beta_2, \alpha)$. *It is clear that if* $\det(H) \neq 0$ *for the critical point* $(\beta_1, \beta_2, \alpha)$ *and* $\det(H_1) < 0$, $\det(H_2) > 0$ *and* $\det(H_3) < 0$ *at* $(\beta_1, \beta_2, \alpha)$ *then it is a local maximum of* $l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s})$*, where*

$$H_1 = \frac{\partial^2 l}{\partial \beta_1^2},\ H_2 = \begin{pmatrix} \frac{\partial^2 l}{\partial \beta_1^2} & \frac{\partial^2 l}{\partial \beta_1 \partial \beta_2} \\ \frac{\partial^2 l}{\partial \beta_2 \partial \beta_1} & \frac{\partial^2 l}{\partial \beta_2^2} \end{pmatrix},\ H_3 = H \text{ and } l = l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s}). \tag{2.88}$$

It can be easily seen that

$$\det(H_1(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})) = -\frac{\left(\ln(1 + r_n^{\widehat{\alpha}})\right)^2}{n} < 0, \tag{2.89}$$

$$\det(H_2(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})) = \frac{\left(\ln(1 + r_n^{\widehat{\alpha}})\right)^2}{n} \frac{\left(\ln(1 + s_m^{\widehat{\alpha}})\right)^2}{m} > 0, \tag{2.90}$$

and

$$\det(H_2(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})) = \frac{\partial G(\widehat{\alpha})}{\partial \alpha} \frac{\left(\ln(1 + r_n^{\widehat{\alpha}})\right)^2}{n} \frac{\left(\ln(1 + s_m^{\widehat{\alpha}})\right)^2}{m} < 0. \tag{2.91}$$

Hence, $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})$ *is the local maximum of* $l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s})$. *Since there is no singular point of* $l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s})$ *and it has a single critical point then, it is enough to show that the absolute maximum of the function is indeed the local maximum. Assume that there exist a* $\widehat{\alpha}_0$ *in the domain in which* $l^*(\widehat{\alpha}_0) > l^*(\widehat{\alpha})$, *where* $l^*(\widehat{\alpha}) = l(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha} \,|\underline{r}, \underline{s})$. *Since* $\widehat{\alpha}$ *is the local maximum there should be some point* α_1 *in the neighborhood of* $\widehat{\alpha}$ *such that* $l^*(\widehat{\alpha}) > l^*(\alpha_1)$. *Let* $k(\alpha) = l^*(\alpha) - l^*(\widehat{\alpha})$ *then* $k(\widehat{\alpha}_0) > 0$, $k(\alpha_1) < 0$ *and* $k(\widehat{\alpha}) = 0$. *This implies that* α_1 *is a local minimum of the* $l^*(\alpha)$, *but* $\widehat{\alpha}$ *is the only critical point so it is a contradiction. Therefore,* $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})$ *is the absolute maximum of* $l(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s})$. ■

Remark 2.1: In case all records are greater than one, we can still get a unique solution of the parameters when we divide the record values, say by r_n *(or by* s_m *or divide* r_i *by* r_n *and divide* s_j *by* s_m *) as long as the transformed observations follow from Burr Type XII.*

2.4.1.2. Asymptotic Distribution and Confidence Intervals For R

The Fisher information matrix $I \equiv I(\beta_1, \beta_2, \alpha)$ is given by

$$I = -\begin{pmatrix} E(\frac{\partial^2 l}{\partial \beta_1^2}) & E(\frac{\partial^2 l}{\partial \beta_1 \partial \beta_2}) & E(\frac{\partial^2 l}{\partial \beta_1 \partial \alpha}) \\ E(\frac{\partial^2 l}{\partial \beta_2 \partial \beta_1}) & E(\frac{\partial^2 l}{\partial \beta_2^2}) & E(\frac{\partial^2 l}{\partial \beta_2 \partial \alpha}) \\ E(\frac{\partial^2 l}{\partial \alpha \partial \beta_1}) & E(\frac{\partial^2 l}{\partial \alpha \partial \beta_2}) & E(\frac{\partial^2 l}{\partial \alpha^2}) \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}, \tag{2.92}$$

where $I_{11} = n/\beta_1^2$, $I_{22} = m/\beta_2^2$,

$$I_{13} = E\left(\frac{R_n \ln R_n}{1 + R_n^\alpha}\right) = \frac{\beta_1^n \psi_1(n, \beta_1)}{\alpha \Gamma(n)}, \tag{2.93}$$

$$I_{23} = E\left(\frac{S_m \ln S_m}{1 + S_m^\alpha}\right) = \frac{\beta_2^m \psi_1(m, \beta_2)}{\alpha \Gamma(m)}, \tag{2.94}$$

$$I_{33} = \frac{n+m}{\alpha^2} + \sum_{i=1}^{n} \frac{\beta_1^i \psi_2(i, \beta_1)}{\alpha^2 \Gamma(i)} + \sum_{j=1}^{m} \frac{\beta_2^j \psi_2(j, \beta_2)}{\alpha^2 \Gamma(j)}$$
$$+ \frac{\beta_1^{n+1} \psi_2(n, \beta_1)}{\alpha^2 \Gamma(n)} + \frac{\beta_2^{m+1} \psi_2(m, \beta_2)}{\alpha^2 \Gamma(m)}, \tag{2.95}$$

$$\psi_1(a, b) = \int_0^\infty \frac{x \ln x (\ln(1+x))^{a-1}}{(1+x)^{b+2}} dx, \tag{2.96}$$

$$\psi_2(a, b) = \int_0^\infty \frac{x (\ln x)^2 (\ln(1+x))^{a-1}}{(1+x)^{b+3}} dx. \tag{2.97}$$

By the asymptotic properties of the MLE, $\widehat{R}$ is asymptotically normal with mean R and asymptotic variance

$$\sigma_R^2 = \sum_{j=1}^{3} \sum_{i=1}^{3} \frac{\partial R}{\partial \beta_i} \frac{\partial R}{\partial \beta_j} I_{ij}^{-1}, \tag{2.98}$$

where $\beta_3 \equiv \alpha$ and I_{ij}^{-1} is the (i, j)th element of the inverse of the $I(\beta_1, \beta_2, \alpha)$, see [Rao, 1965]. Then,

$$\sigma_R^2 = \left(\frac{\partial R}{\partial \beta_1}\right)^2 I_{11}^{-1} + 2 \frac{\partial R}{\partial \beta_1} \frac{\partial R}{\partial \beta_2} I_{12}^{-1} + \left(\frac{\partial R}{\partial \beta_2}\right)^2 I_{22}^{-1}, \tag{2.99}$$

where $\partial R/\partial\beta_1 = \beta_2/(\beta_1+\beta_2)^2$ and $\partial R/\partial\beta_2 = -\beta_1/(\beta_1+\beta_2)^2$. Therefore, the asymptotic $100(1-\gamma)\%$ confidence interval of R is

$$\left(\widehat{R} - z_{\gamma/2}\widehat{\sigma}_R, \widehat{R} + z_{\gamma/2}\widehat{\sigma}_R\right), \tag{2.100}$$

where z_γ is the upper γth quantile of the standard normal distribution and $\widehat{\sigma}_R$ is the value of σ_R at the MLE of the parameters.

If the likelihood equations have a unique solution $\widehat{\theta}_n$, then $\widehat{\theta}_n$ is consistent, asymptotically normal and efficient, see [Lehmann and Casella, 1998]. When the likelihood equations have a unique solution, the observed information matrix $J_m(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})/m$ is a consistent estimator for $I_m(\beta_1, \beta_2, \alpha)/m$ (see Appendix C in [Lawless, 2003]). The observed information matrix $J(\beta_1, \beta_2, \alpha)$ is given by

$$J(\beta_1,\beta_2,\alpha) = -\begin{pmatrix} \frac{\partial^2 l}{\partial\beta_1^2} & \frac{\partial^2 l}{\partial\beta_1\partial\beta_2} & \frac{\partial^2 l}{\partial\beta_1\partial\alpha} \\ \frac{\partial^2 l}{\partial\beta_2\partial\beta_1} & \frac{\partial^2 l}{\partial\beta_2^2} & \frac{\partial^2 l}{\partial\beta_2\partial\alpha} \\ \frac{\partial^2 l}{\partial\alpha\partial\beta_1} & \frac{\partial^2 l}{\partial\alpha\partial\beta_2} & \frac{\partial^2 l}{\partial\alpha^2} \end{pmatrix} = \begin{pmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{pmatrix}, \tag{2.101}$$

where

$$J_{11} = \frac{n}{\beta_1^2},\; J_{12} = J_{21} = \frac{r_n^\alpha \ln r_n}{1+r_n^\alpha},\; J_{22} = \frac{m}{\beta_2^2},\; J_{23} = J_{32} = \frac{s_m^\alpha \ln s_m}{1+s_m^\alpha}, \tag{2.102}$$

$$J_{33} = \frac{n+m}{\alpha^2} + \sum_{i=1}^n r_i^\alpha \left(\frac{\ln r_i}{1+r_i^\alpha}\right)^2 + \sum_{j=1}^m s_j^\alpha \left(\frac{\ln s_j}{1+s_j^\alpha}\right)^2$$

$$+ \beta_1 r_n^\alpha \left(\frac{\ln r_n}{1+r_n^\alpha}\right)^2 + \beta_2 s_m^\alpha \left(\frac{\ln s_m}{1+s_m^\alpha}\right)^2. \tag{2.103}$$

Therefore, an asymptotic $100(1-\gamma)\%$ confidence interval of R can be obtained following from equation (2.100) by replacing I with J in equation (2.99).

2.4.1.3. Bayes Estimation of R

We assume that all parameters β_1, β_2 and α are unknown and have independent gamma prior distributions with parameters (a_i, b_i), $i = 1, 2, 3$, respectively. The density function of a gamma random variable X with parameters (a, b) is

$$f(x) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-xb},\ x > 0,\ a, b > 0. \tag{2.104}$$

Then, the joint posterior density function of β_1, β_2 and α is

$$\pi\left(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s}\right) = I(\underline{r}, \underline{s}) h_1(\underline{r}; \alpha) h_2(\underline{s}; \alpha) \alpha^{n+m+a_3-1} \beta_1^{n+a_1-1} \beta_2^{m+a_2-1}$$

$$\exp\left\{-\alpha b_3 - \beta_1 \left(b_1 + T_1(r_n; \alpha)\right) - \beta_2 \left(b_2 + T_2(s_m; \alpha)\right)\right\}, \tag{2.105}$$

where

$$\left[I(\underline{r}, \underline{s})\right]^{-1} = \int_0^{\infty} \frac{\Gamma(n+a_1)\Gamma(m+a_2) h_1(\underline{r};\alpha) h_2(\underline{s};\alpha) \alpha^{n+m+a_3-1} e^{-\alpha b_3}}{(b_1+T_1(r_n;\alpha))^{n+a_1} (b_2+T_2(s_m;\alpha))^{m+a_2}} d\alpha. \tag{2.106}$$

Then, the Bayes estimate of a given measurable function of β_1, β_2 and α, say $u(\beta_1, \beta_2, \alpha)$ under the SE loss function is

$$\widehat{u}_B = \int_0^{\infty} \int_0^{\infty} \int_0^{\infty} u(\beta_1, \beta_2, \alpha) \pi(\beta_1, \beta_2, \alpha \,|\underline{r}, \underline{s}) d\beta_1 d\beta_2 d\alpha. \tag{2.107}$$

It is not possible to compute equation (2.107) analytically. Two approaches can be applied to approximate equation (2.107), namely Lindley's approximation and MCMC method.

- Lindley's approximation

For the three parameter case $(\beta_1, \beta_2, \alpha)$, we have $L_{11} = -n/\beta_1^2$, $L_{22} = -m/\beta_2^2$, $L_{13} = L_{31} = -r_n^{\alpha} \ln r_n/(1 + r_n^{a})$, $L_{23} = L_{32} = -s_m^{\alpha} \ln s_m/(1 + s_m^{\alpha})$

$$L_{33} = -\frac{n+m}{\alpha^2} - \sum_{i=1}^{n} r_i^{\alpha} \left(\frac{\ln r_i}{1 + r_i^{\alpha}}\right)^2 - \sum_{j=1}^{m} s_j^{\alpha} \left(\frac{\ln s_j}{1 + s_j^{\alpha}}\right)^2$$

$$- \beta_1 r_n^{\alpha} \left(\frac{\ln r_n}{1 + r_n^{a}}\right)^2 - \beta_2 s_m^{\alpha} \left(\frac{\ln s_m}{1 + s_m^{\alpha}}\right)^2, \tag{2.108}$$

$\rho_1 = ((a_1 - 1)/\beta_1) - b_1$, $\rho_2 = ((a_2 - 1)/\beta_2) - b_2$, $\rho_3 = ((a_3 - 1)/\alpha) - b_3$, $\sigma_{ij}, i, j = 1, 2, 3$ are obtained by using L_{ij}, $i, j = 1, 2, 3$ and $L_{111} = 2n/\beta_1^3$, $L_{222} = 2m/\beta_2^3$,

$$L_{133} = L_{331} = -r_n^{\alpha}\left(\frac{\ln r_n}{1+r_n^{a}}\right)^2, \; L_{233} = L_{322} = -s_m^{\alpha}\left(\frac{\ln s_m}{1+s_m^{\alpha}}\right)^2, \tag{2.109}$$

$$L_{333} = \frac{2(n+m)}{\alpha^3} - \sum_{i=1}^{m}\frac{r_i^{\alpha}(1-r_i^{\alpha})(\ln r_i)^3}{(1+r_i^{\alpha})^3} - \sum_{j=1}^{m}\frac{s_j^{\alpha}(1-s_j^{\alpha})(\ln s_j)^3}{(1+s_j^{\alpha})^3}$$

$$- \frac{\beta_1 r_n^{\alpha}(1-r_n^{\alpha})(\ln r_n)^2}{(1+r_n^{\alpha})^2} - \frac{\beta_2 s_m^{\alpha}(1-s_m^{\alpha})(\ln s_m)^2}{(1+s_m^{\alpha})^2}. \tag{2.110}$$

Moreover, $A = \sigma_{11}L_{111} + \sigma_{33}L_{331}$, $B = \sigma_{22}L_{222} + \sigma_{33}L_{332}$ and $C = 2\sigma_{13}L_{133} + 2\sigma_{23}L_{233} + \sigma_{33}L_{333}$. To obtain the Bayes estimate of R under the SE loss function, we take $u(\beta_1,\beta_2,\alpha) = R = \beta_1/(\beta_1+\beta_2)$. Then, $u_3 = u_{13} = u_{23} = u_{33} = 0$, $u_1 = \beta_2/(\beta_1+\beta_2)^2$, $u_2 = -\beta_1/(\beta_1+\beta_2)^2$, $u_{12} = u_{21} = (\beta_1-\beta_2)/(\beta_1+\beta_2)^3$, $u_{11} = -2\beta_2/(\beta_1+\beta_2)^3$, $u_{22} = 2\beta_1/(\beta_1+\beta_2)^3$ and $c_4 = u_{12}\sigma_{12}$, $c_5 = (u_{11}\sigma_{11} + u_{22}\sigma_{22})/2$. Hence, the Bayes estimate of R under the SE loss function is

$$\widehat{R}_{BS,Lindley} = R + [u_1c_1 + u_2c_2 + c_4 + c_5] + \frac{1}{2}\left\{A\left[u_1\sigma_{11} + u_2\sigma_{12}\right]\right.$$

$$\left. B\left[u_1\sigma_{21} + u_2\sigma_{22}\right] + C\left[u_1\sigma_{31} + u_2\sigma_{32}\right]\right\}. \tag{2.111}$$

Notice that all parameters are evaluated at $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})$.

For the Bayes estimate of R under the LINEX loss function, $u(\beta_1,\beta_2,\alpha) = e^{-vR}$ are taken. Then, $u_3 = u_{13} = u_{23} = u_{33} = 0$, $u_1 = -v\beta_2 e^{-vR}/(\beta_1+\beta_2)^2$, $u_2 = v\beta_1 e^{-vR}/(\beta_1+\beta_2)^2$

$$u_{12} = -\frac{v^2 e^{-vR}\beta_1\beta_2}{(\beta_1+\beta_2)^4} - \frac{v e^{-vR}(\beta_1-\beta_2)}{(\beta_1+\beta_2)^3}, \tag{2.112}$$

$$u_{11} = \frac{v e^{-vR}(v\beta_2^2 + 2\beta_1\beta_2 + 2\beta_2^2)}{(\beta_1+\beta_2)^4}, \; u_{22} = \frac{v e^{-vR}(v\beta_1^2 - 2\beta_1\beta_2 - 2\beta_1^2)}{(\beta_1+\beta_2)^4}, \tag{2.113}$$

and $c_4 = u_{12}\sigma_{12}$, $c_5 = \frac{1}{2}(u_{11}\sigma_{11} + u_{22}\sigma_{22})$. Then, the Bayes estimate of R under the LINEX loss function is $\widehat{R}_{BL,Lindley} = -\left(\ln E(e^{-vR})\right)/v$ where

$$E(e^{-vR}) = e^{-vR} + [u_1c_1 + u_2c_2 + c_4 + c_5] + \frac{1}{2}\{A[u_1\sigma_{11} + u_2\sigma_{12}]$$

$$B[u_1\sigma_{21} + u_2\sigma_{22}] + C[u_1\sigma_{31} + u_2\sigma_{32}]\}. \quad (2.114)$$

Notice that all parameters are evaluated at $(\widehat{\beta}_1, \widehat{\beta}_2, \widehat{\alpha})$.

- MCMC method

In the previous section, the Bayes estimate of R are obtained by using the Lindley's approximation under the SE and the LINEX loss functions. Since the exact probability distribution of R are not known, it is difficult to evaluate Bayesian credible interval of R. For this reason, the MCMC method are used to compute the Bayes estimate R under the SE and the LINEX loss functions as well as the HPD credible interval.

The MCMC method are considered to generate samples from the posterior distributions and then compute the Bayes estimate of R under the SE and the LINEX loss functions. The joint posterior density of α and β is given by equation (2.105). It is easy to see that the posterior density functions of β_1, β_2 and α are

$$\beta_1 \,|\alpha, \underline{r}, \underline{s} \sim Gamma(n + a_1, b_1 + T_1(r_n; \alpha)), \quad (2.115)$$

$$\beta_2 \,|\alpha, \underline{r}, \underline{s} \sim Gamma(m + a_2, b_2 + T_2(s_m; \alpha)), \quad (2.116)$$

and

$$\pi(\alpha\,|\beta_1, \beta_2, \underline{r}, \underline{s}) \propto \alpha^{n+m+a_3-1} \exp\{-\alpha b_3 - \beta_1 T_1(r_n; \alpha) - \beta_2 T_2(s_m; \alpha)\}$$

$$\exp\left\{-\sum_{i=1}^{n} \ln(1 + r_i^{\alpha}) + \alpha\left(\sum_{i=1}^{n} \ln r_i +_{j=1}^{m} \ln s_j\right) - \sum_{j=1}^{m} \ln(1 + s_j^{\alpha})\right\}. \quad (2.117)$$

Therefore, samples of β_1 and β_2 can be generated by using the gamma distribution. However, the posterior distribution of α cannot be reduced analytically to well known distribution and therefore it is not possible to sample directly by standard methods.

If the posterior density of α is unimodal and roughly symmetric then it is often convenient to approximate it by a normal distribution (see [Gelman et al., 2003]). Since the posterior density of α is log-concave density (so unimodal) and it is roughly symmetric (by experimentation), we use the Metropolis-Hasting algorithm with the normal proposal distribution to generate a random sample from the posterior density of α. The hybrid Metropolis-Hastings and Gibbs sampling algorithm, which will be used to solve our problem, is suggested by [Tierney, 1994]. This algorithm combines the Metropolis-Hastings with Gibbs sampling scheme under the normal proposal distribution.

- Step 1: Start with initial guess $\alpha^{(0)}$.
- Step 2: Set $i = 1$.
- Step 3: Generate $\beta_1^{(i)}$ from $Gamma(n + a_1, T_1(r_n; \alpha^{(i-1)}) + b_1)$.
- Step 4: Generate $\beta_2^{(i)}$ from $Gamma(m + a_2, T_2(s_m; \alpha^{(i-1)}) + b_2)$.
- Step 5: Generate $\alpha^{(i)}$ from $\pi(\alpha\,|\beta_1, \beta_2, \underline{r}, \underline{s})$ using the Metropolis-Hastings algorithm with the proposal distribution $q(\alpha) \equiv N(\alpha^{(i-1)}, 1)$:

  -Step 5.1: Let $v = \alpha^{(i-1)}$.

  -Step 5.2: Generate w from the proposal distribution q.

  -Step 5.3: Let $p(v, w) = \min\left\{1, \dfrac{\pi(w\left|\beta_1^{(i)}, \beta_2^{(i)}, \underline{r}, \underline{s}\right.)\, q(v)}{\pi(v\left|\beta_1^{(i)}, \beta_2^{(i)}, \underline{r}, \underline{s}\right.)\, q(w)}\right\}$.

  -Step 5.4: Generate u from $Uniform(0, 1)$. If $u \leq p(v, w)$ then accept the proposal and set $\alpha^{(i)} = w$; otherwise, set $\alpha^{(i)} = v$.
- Step 6: Compute the $R^{(i)} = \beta_1^{(i)}/(\beta_1^{(i)} + \beta_2^{(i)})$.
- Step 7. Set $i = i + 1$.
- Step 8. Repeat Steps 2-7, N times, and obtain the posterior sample $R^{(i)}$, $i = 1, ..., N$.

This sample are used to compute the Bayes estimate and to construct the HPD credible interval for R. The Bayes estimate of R under the SE and the LINEX loss function are given as

$$\widehat{R}_{BS,MCMC} = \frac{1}{N-M} \sum_{i=M+1}^{N-M} R^{(i)}, \tag{2.118}$$

$$\widehat{R}_{BL,MCMC} = -\frac{1}{v}\ln E(e^{-vR}) = -\frac{1}{v}\ln\left(\frac{1}{N-M}\sum_{i=M+1}^{N-M} e^{-vR^{(i)}}\right), \qquad (2.119)$$

where M is the burn-in period.

The HPD $100(1-\gamma)\%$ credible interval of R is obtained by the method of [Chen and Shao, 1999]. From MCMC, the sequence $R^{(1)},\ldots,R^{(N)}$, are obtained, and ordered as $R_{(1)} < \ldots < R_{(N)}$. The credible intervals are constructed as $\left(R_{(j)}, R_{(j+[N(1-\gamma)])}\right)$ for $j = 1,...,N-[N(1-\gamma)]$ where $[x]$ denotes the largest integer less than or equal to x. Then, the HPD credible interval of R is that interval which has the shortest length.

2.4.2. Estimation of R When α Is Common and Known

The estimation of R are considered when α is known, say $\alpha = \alpha_0$. Let $R_1,\ldots,R_n$ be a set of upper records from $Burr(\alpha_0,\beta_1)$ and $S_1,\ldots,S_m$ be an independent set of upper records from $Burr(\alpha_0,\beta_2)$.

2.4.2.1. ML Estimation and Confidence Intervals of R

Based on the above samples, the MLE of R, say $\widehat{R}_{MLE}$, is

$$\widehat{R}_{MLE} = \frac{\widehat{\beta}_1}{\widehat{\beta}_1+\widehat{\beta}_2} = \frac{nT_2(s_m;\alpha_0)}{nT_2(s_m;\alpha_0)+mT_1(r_n;\alpha_0)}, \qquad (2.120)$$

where $T_1(r_n;\alpha_0) = \ln(1+r_n^{\alpha_0})$, $T_2(s_m;\alpha_0) = \ln(1+s_m^{\alpha_0})$.

It is easy to see that $2\beta_1\ln(1+r_n^{\alpha_0}) \sim \chi^2(2n)$ and $2\beta_2\ln(1+s_m^{\alpha_0}) \sim \chi^2(2m)$. Therefore,

$$F^* = \left(\frac{R}{1-R}\right)\left(\frac{1-\widehat{R}_{MLE}}{\widehat{R}_{MLE}}\right) \qquad (2.121)$$

is an F distributed random variable with $(2n,2m)$ degrees of freedom. The pdf of $\widehat{R}_{MLE}$ is as follows;

$$f_{\widehat{R}_{MLE}}(r) = \frac{1}{r^2 B(m,n)} \left(\frac{n\beta_1}{m\beta_2}\right)^n \frac{\left(\frac{1-r}{r}\right)^{n-1}}{\left(1+\frac{n\beta_1(1-r)}{m\beta_2 r}\right)^{n+m}}, \tag{2.122}$$

where $0 < r < 1$. The $100(1-\gamma)\%$ exact confidence interval for R can be obtained as

$$\left(\frac{1}{1+F_{2m,2n;\frac{\gamma}{2}}\left(\frac{1-\widehat{R}_{MLE}}{\widehat{R}_{MLE}}\right)}, \frac{1}{1+F_{2m,2n;1-\frac{\gamma}{2}}\left(\frac{1-\widehat{R}_{MLE}}{\widehat{R}_{MLE}}\right)}\right), \tag{2.123}$$

where $F_{2m,2n;\frac{\gamma}{2}}$ and $F_{2m,2n;1-\frac{\gamma}{2}}$ are the lower and upper $\frac{\gamma}{2}$th percentile points of a F distribution with $(2m, 2n)$ degrees of freedom.

On the other hand, the approximate confidence interval of R can be easily obtained by using the Fisher information matrix. The Fisher information matrix of (β_1, β_2) is

$$I = -\begin{pmatrix} E\left(\frac{\partial^2 l}{\partial\beta_1^2}\right) & E\left(\frac{\partial^2 l}{\partial\beta_1\partial\beta_2}\right) \\ E\left(\frac{\partial^2 l}{\partial\beta_1\partial\beta_2}\right) & E\left(\frac{\partial^2 l}{\partial\beta_2^2}\right) \end{pmatrix} = \begin{pmatrix} n/\beta_1^2 & 0 \\ 0 & m/\beta_2^2 \end{pmatrix}. \tag{2.124}$$

By the asymptotic properties of the MLE, $\widehat{R}_{MLE}$ is asymptotically normal with mean R and asymptotic variance

$$\sigma_R^2 = \sum_{j=1}^{2}\sum_{i=1}^{2} \frac{\partial R}{\partial\beta_i}\frac{\partial R}{\partial\beta_j} I_{ij}^{-1} \tag{2.125}$$

where I_{ij}^{-1} is the (i,j) th element of the inverse of the I, (see [Rao, 1965]). Then, the asymptotic $100(1-\gamma)\%$ confidence interval for R is

$$\left(\widehat{R}_{MLE} - z_{\gamma/2}\widehat{\sigma}_R, \widehat{R}_{MLE} + z_{\gamma/2}\widehat{\sigma}_R\right), \tag{2.126}$$

where z_γ is the upper γth percentile points of a standard normal distribution, $\sigma_R^2 = R^2(1-R)^2\left(1/n + 1/m\right)$ and $\widehat{\sigma}_R$ is the value of σ_R at the MLE of the parameters.

2.4.2.2. UMVUE of R

When the first shape parameter α is known, $(T_1(r_n;\alpha_0), T_2(s_m;\alpha_0))$ is a sufficient statistics for (β_1, β_2). It can be shown that it is also a complete sufficient statistic by using Theorem 10-9 in [Arnold, 1990]. Let us define

$$\phi(R_1, S_1) = \begin{cases} 1 & \text{if } R_1 < S_1 \\ 0 & \text{if } R_1 \geq S_1 \end{cases} . \tag{2.127}$$

Then $E(\phi(R_1, S_1)) = R$ so it is an unbiased estimator of R. Let $P_1 = \ln(1 + R_1^{\alpha_0})$ and $P_2 = \ln(1 + S_1^{\alpha_0})$. The UMVUE of R, say $\widehat{R}_U$, can be obtained by using the Rao-Blackwell and the Lehmann-Scheffe's Theorems, (see [Arnold, 1990]),

$$\widehat{R}_U = E(\phi(P_1, P_2) \mid (T_1, T_2))$$

$$= \int_{P_2}\int_{P_1} \phi(P_1, P_2) f_{P_1|T_1}(p_1 \left| T_1 \right.) f_{P_2|T_2}(p_2 \left| T_2 \right.) dp_1 dp_2, \tag{2.128}$$

where $(T_1, T_2) = (T_1(r_n;\alpha_0), T_2(s_m;\alpha_0))$, $f(p_1, p_2 \left| T_1, T_2 \right.)$ is the conditional pdf of (P_1, P_2) given (T_1, T_2). Using the joint pdf of (R_1, R_n) and (S_1, S_m) and after making a simple transformation, we obtain the $f_{P_1|T_1}(p_1 \left| T_1 \right.)$ and $f_{P_2|T_2}(p_2 \left| T_2 \right.)$, and are given by

$$f_{P_1|T_1}(p_1 \mid T_1) = (n-1)\frac{(t_1 - p_1)^{n-2}}{t_1^{n-1}},\ 0 < p_1 < t_1, \tag{2.129}$$

$$f_{P_2|T_2}(p_2 \mid T_2) = (m-1)\frac{(t_2 - p_2)^{m-2}}{t_2^{m-1}},\ 0 < p_2 < t_2. \tag{2.130}$$

Therefore,

$$\widehat{R}_U = \iint\limits_{P_1 < P_2} f_{P_1|T_1}(p_1 \mid T_1) f_{P_2|T_2}(p_2 \mid T_2) dp_1 dp_2$$

$$= \begin{cases} \int_0^{t_1}\int_{p_1}^{t_2} \dfrac{(n-1)(m-1)(t_1-p_1)^{n-2}(t_2-p_2)^{m-2}}{t_1^{n-1}t_2^{m-1}} dp_2 dp_1 & \text{if } t_2 \geq t_1 \\ \int_0^{t_2}\int_0^{p_2} \dfrac{(n-1)(m-1)(t_1-p_1)^{n-2}(t_2-p_2)^{m-2}}{t_1^{n-1}t_2^{m-1}} dp_1 dp_2 & \text{if } t_2 < t_1 \end{cases} \tag{2.131}$$

$$= \begin{cases} {}_2F_1(1, 1-m; n; t_1/t_2) & \text{if } t_2 \geq t_1 \\ 1 -_2 F_1(1, 1-n; m; t_2/t_1) & \text{if } t_2 < t_1 \end{cases},$$

where ${}_2F_1(.,.;.;.)$ is Gauss hypergeometric function, (see formula 3.196(1) in [Gradshteyn and Ryzhik, 1994]).

2.4.2.3. Bayes Estimation of R

Assume that the parameters β_1 and β_2 are random variables and have independent gamma prior distributions with parameters (a_i, b_i), $i = 1, 2$, respectively. Then, the joint posterior density function of β_1 and β_2 is

$$\pi\left(\beta_1, \beta_2 \left|\alpha_0, \underline{r}, \underline{s}\right.\right) = \frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)} \beta_1^{\delta_1-1}\beta_2^{\delta_2-1} e^{-\beta_1\lambda_1} e^{-\beta_2\lambda_2}, \tag{2.132}$$

where $\lambda_1 = b_1 + T_1(r_n; \alpha_0)$, $\lambda_2 = b_2 + T_2(s_m; \alpha_0)$, $\delta_1 = n + a_1$, $\delta_2 = m + a_2$. The posterior pdf of R can be obtained by using the joint posterior density function and is given by

$$f_R(r) = \frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{B(\delta_1, \delta_2)} \frac{r^{\delta_1-1}(1-r)^{\delta_2-1}}{(r\lambda_1 + (1-r)\lambda_2)^{\delta_1+\delta_2}},\ 0 < r < 1. \tag{2.133}$$

After making suitable transformations and simplifications by using formula 3.197(3) of [Gradshteyn and Ryzhik, 1994], the Bayes estimate of R, say $\widehat{R}_{BS}$, under the SE loss function is

$$\widehat{R}_{BS} = \begin{cases} c\left(\frac{\lambda_1}{\lambda_2}\right)^{\delta_1} {}_2F_1(c^*, \delta_1+1; c^*+1; 1-\frac{\lambda_1}{\lambda_2}) & \text{if } \lambda_1 < \lambda_2 \\ c\left(\frac{\lambda_2}{\lambda_1}\right)^{\delta_2} {}_2F_1(c^*, \delta_2; c^*+1; 1-\frac{\lambda_2}{\lambda_1}) & \text{if } \lambda_2 \leq \lambda_1 \end{cases}, \tag{2.134}$$

where $c = \delta_1/c^*$ and $c^* = \delta_1 + \delta_2$.

The Bayes estimate of R under the LINEX loss function, say $\widehat{R}_{BL}$, is $\widehat{R}_{BL} = -\left\{\ln E_R(e^{-vR})\right\}/v$, where $E_R(.)$ denotes posterior expectation with respect to the posterior density of R. It can be easily obtained that

$$E(e^{-vR}) = \begin{cases} \left(\frac{\lambda_1}{\lambda_2}\right)^{\delta_1} \Phi_1(\delta_1, c^*, c^*, 1-\frac{\lambda_1}{\lambda_2}, -v) & \text{if } \lambda_1 < \lambda_2 \\ \left(\frac{\lambda_2}{\lambda_1}\right)^{\delta_2} e^{-v}\Phi_1(\delta_2, c^*, c^*, 1-\frac{\lambda_2}{\lambda_1}, v) & \text{if } \lambda_2 \leq \lambda_1 \end{cases}, \tag{2.135}$$

where $\Phi_1(.,.,.,.,.)$ is confluent hypergeometric series of two variables, (see formulas 3.385 and 9.261(1) in [Gradshteyn and Ryzhik, 1994]). Therefore,

$$\widehat{R}_{BL} = \begin{cases} -\frac{1}{v}\left(c_1 + \ln\left[\Phi_1(\delta_1, c^*, c^*, 1-\frac{\lambda_1}{\lambda_2}, -v)\right]\right) & \text{if } \lambda_1 < \lambda_2 \\ -\frac{1}{v}\left(c_2 + \ln\left[\Phi_1(\delta_2, c^*, c^*, 1-\frac{\lambda_2}{\lambda_1}, v)\right]\right) & \text{if } \lambda_2 \leq \lambda_1 \end{cases}, \tag{2.136}$$

where $c_1 = \delta_1 \ln(\frac{\lambda_1}{\lambda_2})$ and $c_2 = \delta_2 \ln(\frac{\lambda_2}{\lambda_1}) - v$.

If we use the Jeffrey's non informative prior, is given by $\sqrt{\det I}$, then the joint prior density function is $\pi(\beta_1, \beta_2) \propto 1/\beta_1\beta_2$. Therefore, the joint posterior density function of β_1 and β_2 is

$$\pi(\beta_1, \beta_2\,|\alpha_0, \underline{r}, \underline{s}) = \frac{T_1^n {T_2}^m}{\Gamma(n)\Gamma(m)}\beta_1^{n-1}\beta_2^{m-1}e^{-\beta_1 T_1}e^{-\beta_2 T_2}, \tag{2.137}$$

and the posterior pdf of R is given by

$$f_R(r) = \frac{T_1^n {T_2}^m}{B(n,m)}\frac{r^{n-1}(1-r)^{m-1}}{(rT_1 + (1-r)T_2)^{n+m}},\ 0 < r < 1, \tag{2.138}$$

where $T_1 = T_1(r_n; \alpha_0)$ and $T_2 = T_2(s_m; \alpha_0)$. The Bayes estimate of R under the SE and the LINEX loss function, say $\widehat{R}^*_{BS}$ and $\widehat{R}^*_{BL}$ respectively, are

$$\widehat{R}^*_{BS} = \begin{cases} c_3\left(\frac{T_1}{T_2}\right)^n {}_2F_1(c_3^*, n+1; c_3^*+1; 1-\frac{T_1}{T_2}) & \text{if } T_1 < T_2 \\ c_3\left(\frac{T_2}{T_1}\right)^m {}_2F_1(c_3^*, m; c_3^*+1; 1-\frac{T_2}{T_1}) & \text{if } T_2 \leq T_1 \end{cases}, \tag{2.139}$$

and

$$\widehat{R}^*_{BL} = \begin{cases} -\dfrac{1}{v}\left(c_4 + \ln\left[\Phi_1(n, c_3^*, c_3^*, 1 - \dfrac{T_1}{T_2}, -v)\right]\right) & \text{if } T_1 < T_2 \\ -\dfrac{1}{v}\left(c_5 - v + \ln\left[\Phi_1(m, c_3^*, c_3^*, 1 - \dfrac{T_2}{T_1}, v)\right]\right) & \text{if } T_2 \leq T_1 \end{cases}, \qquad (2.140)$$

where $c_3 = n/c_3^*$, $c_3^* = n + m$, $c_4 = n\ln(T_1/T_2)$ and $c_5 = m\ln(T_2/T_1)$.

The Bayes estimates are not always derived in the closed forms. However, for our case the Bayes estimates are obtained in the closed form. These estimates can be obtained by using alternative methods such as Lindley's approximation and MCMC method. The purpose of applying all these two methods is to see how good the approximate methods compared with the exact one. If these result are close, then it will be encouraging to use the approximate methods when the exact form can not be obtained as in the case of α unknown. These estimators will be compared in the simulation study section. Next, the Bayes estimates of R are given by using Lindley's approximation and MCMC method.

- Lindley's approximation

The approximate Bayes estimate of R under the SE and the LINEX loss functions for the informative prior case, say $\widehat{R}_{BS,Lindley}$ and $\widehat{R}_{BL,Lindley}$ respectively, are

$$\widehat{R}_{BS,Lindley} = \widetilde{R}\left(1 + \frac{(1-\widetilde{R})^2}{n + a_1 - 1} - \frac{\widetilde{R}(1-\widetilde{R})}{m + a_2 - 1}\right), \qquad (2.141)$$

and

$$\widehat{R}_{BL,Lindley} = \widetilde{R} - \frac{1}{v}\ln\left(1 + \frac{\widetilde{R}_1(1-\widetilde{R})(v\widetilde{R} - 2)}{2(n + a_1 - 1)} + \frac{\widetilde{R}_1\widetilde{R}(v - v\widetilde{R} + 2)}{2(m + a_2 - 1)}\right), \quad (2.142)$$

where $\widetilde{R} = \widetilde{\beta}_1/(\widetilde{\beta}_1 + \widetilde{\beta}_2)$, $\widetilde{R}_1 = v\widetilde{R}(1-\widetilde{R})$, $\widetilde{\beta}_1 = (n + a_1 - 1)/(b_1 + T_1(r_n;\alpha_0))$ and $\widetilde{\beta}_2 = (m + a_2 - 1)/(b_2 + T_2(s_m;\alpha_0))$.

If we use the Jeffrey's non informative prior, the approximate Bayes estimate of R under the SE and the LINEX loss functions, say $\widehat{R}^*_{BS,Lindley}$ and $\widehat{R}^*_{BL,Lindley}$ respectively, are

$$\widehat{R}^*_{BS,Lindley} = \widetilde{R}\left(1 + \frac{(1-\widetilde{R})^2}{n-1} - \frac{\widetilde{R}(1-\widetilde{R})}{m-1}\right), \tag{2.143}$$

and

$$\widehat{R}^*_{BL,Lindley} = \widetilde{R} - \frac{1}{v}\ln\left(1 + \frac{\widetilde{R}_1(1-\widetilde{R})(v\widetilde{R}-2)}{2(n-1)} + \frac{\widetilde{R}_1\widetilde{R}(v-v\widetilde{R}+2)}{2(m-1)}\right), \tag{2.144}$$

where $\widetilde{R} = \widetilde{b}_1/(\widetilde{b}_1 + \widetilde{b}_2)$, $\widetilde{R}_1 = v\widetilde{R}(1-\widetilde{R})$, $\widetilde{b}_1 = (n-1)/T_1(r_n;\alpha_0)$ and $\widetilde{b}_2 = (m-1)/T_2(s_m;\alpha_0)$.

- MCMC method

It is clear from equation (2.132), the marginal posterior densities of β_1 and β_2 are gamma distribution with the parameters (δ_1,λ_1) and (δ_2,λ_2), respectively. A samples are generate by using Gibss sampling from these distributions. The following algorithm are used.

- Step 1: Set $i=1$.
- Step 2: Generate $\beta_1^{(i)}$ from $Gamma(\delta_1,\lambda_1)$.
- Step 3: Generate $\beta_2^{(i)}$ from $Gamma(\delta_2,\lambda_2)$.
- Step 4: Compute the $R^{(i)} = \beta_1^{(i)}/(\beta_1^{(i)} + \beta_2^{(i)})$.
- Step 5: Set $i = i+1$.
- Step 6: Repeat Steps 2-5, N times, and obtain the posterior sample $R^{(i)}$, $i = 1,...,N$.

This sample is used to compute the Bayes estimate and to construct the HPD credible interval for R. The Bayes estimate of R under the SE and the LINEX loss functions are given as

$$\widehat{R}_{BS,MCMC} = \frac{1}{N}\sum_{i=1}^{N} R^{(i)}, \tag{2.145}$$

$$\widehat{R}_{BL,MCMC} = -\frac{1}{v}\ln E(e^{-vR}) = -\frac{1}{v}\ln\left(\frac{1}{N}\sum_{i=1}^{N} e^{-vR^{(i)}}\right). \tag{2.146}$$

The HPD $100(1-\gamma)\%$ credible interval of R can be obtained by the method of [Chen and Shao, 1999].

2.4.2.4. Empirical Bayes Estimation of R

The Bayes estimates of R are obtained by using three different ways. It is clear that these estimates depend on the prior parameters. However, the Bayes estimates can be also obtained independently of the prior parameters.

These prior parameters could be estimated by means of an empirical Bayes procedure, (see [Lindley, 1969], [Awad and Gharraf, 1986]). Let $R_1, \ldots, R_n$ and $S_1, \ldots, S_m$ be two independent random samples from $Burr(\alpha_0, \beta_1)$ and $Burr(\alpha_0, \beta_2)$, respectively. For fixed $\underline{r}$, the function $L_1(\beta_1 \,|\alpha_0, \underline{r})$ of β_1 can be considered as a gamma density with parameters $(n+1, T_1(r_n;\alpha_0))$. Therefore, it is proposed to estimate the prior parameters α_1 and β_1 from the samples as $n+1$ and $T_1(r_n;\alpha_0)$, respectively. Similarly, α_2 and β_2 could be estimated from the samples as $m+1$ and $T_2(s_m;\alpha_0)$, respectively. Hence, the empirical Bayes estimate of R with respect to SE and LINEX loss functions, say $\widehat{R}_{EBS}$ and $\widehat{R}_{EBL}$, respectively, could be given as

$$\widehat{R}_{EBS} = \begin{cases} c_6 c_7 \, {}_2F_1(c_{13}, 2n+2; c_{13}+1; c_9) & \text{if } T_1 < T_2 \\ c_6 c_8 \, {}_2F_1(c_{13}, 2m+1; c_{13}+1; c_{10}) & \text{if } T_2 \leq T_1 \end{cases}, \tag{2.147}$$

and

$$\widehat{R}_{EBL} = \begin{cases} -\frac{1}{v}\left((2n+1)\ln(\frac{T_1}{T_2}) + \ln c_{11}\right) & \text{if } T_1 < T_2 \\ -\frac{1}{v}\left((2m+1)\ln(\frac{T_2}{T_1}) - v + \ln c_{12}\right) & \text{if } T_2 \leq T_1 \end{cases}. \tag{2.148}$$

where $c_6 = (2n+1)/c_{13}$, $c_7 = (T_1/T_2)^{2n+1}$, $c_8 = (T_2/T_1)^{2m+1}$, $c_9 = 1-(T_1/T_2)$, $c_{10} = 1-(T_2/T_1)$, $c_{11} = \Phi_1(2n+1, c_{13}, c_{13}, c_9, -v)$ and $c_{12} = \Phi_1(2m+1, c_{13}, c_{13}, c_{10}, v)$, $c_{13} = 2n+2m+2$.

2.4.2.5. Bayesian Credible Intervals For R

It is known that $\beta_1 \,|\alpha_0, \underline{r} \sim Gamma(\delta_1, \lambda_1)$ and $\beta_2 \,|\alpha_0, \underline{s} \sim Gamma(\delta_2, \lambda_2)$. Then, $2\lambda_1\beta_1 \,|\alpha_0, \underline{r} \sim \chi^2(2(n+a_1))$ and $2\lambda_2\beta_2 \,|\alpha_0, \underline{s} \sim \chi^2(2(m+a_2))$. Therefore,

$$W = \frac{2\lambda_2\beta_2 \,|\alpha_0, \underline{s} \,/2(m+a_2)}{2\lambda_1\beta_1 \,|\alpha_0, \underline{r} \,/2(n+a_1)} \tag{2.149}$$

is an F distributed random variable with $(2(m+a_2), 2(n+a_1))$ degrees of freedom and the $100(1-\gamma)\%$ Bayesian credible interval for R can be obtained as

$$\left(\frac{1}{1+CF_{2(m+a_2),2(n+a_1);\frac{\gamma}{2}}}, \frac{1}{1+CF_{2(m+a_2),2(n+a_1);1-\frac{\gamma}{2}}}\right) \tag{2.150}$$

where $C = \delta_2\lambda_1/\delta_1\lambda_2$, $F_{2(m+a_2),2(n+a_1);\frac{\gamma}{2}}$ and $F_{2(m+a_2),2(n+a_1);1-\frac{\gamma}{2}}$ are the lower and upper $\frac{\gamma}{2}$th percentile points of a F distribution with $(2(m+a_2), 2(n+a_1))$ degrees of freedom.

Moreover, this interval can be obtained independently of these parameters by using the empirical method. In this case, the posterior distributions of β_1 and β_2 have gamma distributions with parameters $(2n+1, 2T_1(r_n;\alpha_0))$ and $(2m+1, 2T_2(s_m;\alpha_0))$, respectively and the $100(1-\gamma)\%$ Bayesian credible interval for R can be obtained as

$$\left(\frac{1}{1+C_1F_{(4m+2),(4n+2);\frac{\gamma}{2}}}, \frac{1}{1+C_1F_{(4m+2),(4n+2);1-\frac{\gamma}{2}}}\right) \tag{2.151}$$

where $C_1 = ((4m+2)T_1(r_n;\alpha_0)) \,/\, ((4n+2)T_2(s_m;\alpha_0))$, $F_{(4m+2),(4n+2);\frac{\gamma}{2}}$ and $F_{(4m+2),(4n+2);1-\frac{\gamma}{2}}$ are the lower and upper $\frac{\gamma}{2}$th percentile points of a F distribution with $(4m+2, 4n+2)$ degrees of freedom.

2.4.3. Numerical Experiments

In this section, firstly the Monte Carlo simulations for the comparison of the derived estimates are presented, then two real life data sets are analysed.

2.4.3.1. Simulation Study

In this section, some numerical results are presented to compare the performance of the different estimates for different sample sizes and different priors. The performances of the point estimates are compared by using ERs. The performances of the confidence and credible intervals are compared by using average interval lengths and cps. All of the computations are performed by using MATLAB R2010a. All the results are based on 3000 replications.

We consider two cases separately to draw inference on R, namely when the common first shape parameter α is unknown and known. In both cases, the upper record values are generated from the Burr Type XII distribution with the sample sizes; $(n, m) = (5, 5)$, $(8, 8)$, $(10, 10)$, $(12, 12)$, $(15, 15)$.

In Table 2.10, the ML and Bayes estimates of R and their corresponding ERs are listed when α is unknown. The Bayes estimates are computed by using Lindley's approximation and MCMC method under the SE and the LINEX ($v = -1$ and 1) loss functions for different prior parameters. In the Bayesian case, Prior 1: $(a_1, b_1) = (4, 2)$, $(a_2, b_2) = (4, 2)$, $(a_3, b_3) = (3, 3)$, Prior 2: $(a_1, b_1) = (5, 1)$, $(a_2, b_2) = (3, 3/2)$, $(a_3, b_3) = (3, 3/2)$ and Prior 3: $(a_1, b_1) = (5, 1/2)$, $(a_2, b_2) = (3, 3)$, $(a_3, b_3) = (3, 3/2)$, are used for $R = 0.5006$, 0.7145 and 0.9095, respectively. Moreover, the 95% asymptotic confidence intervals, which are computed based on Fisher information and observation matrices, and HPD credible intervals with their cps are listed. From Table 2.10, the ERs of all estimates decrease as the sample sizes increase in all cases, as expected. The Bayes estimates under the SE and LINEX loss functions generally have smaller ER than that of ML estimates. Moreover, the ERs of the Bayes estimates based on Lindley's approximation are smaller than that of MCMC method. These estimates are close to each other as the sample sizes increase. The average lengths of the intervals decrease as the sample sizes increase. The asymptotic confidence intervals based on Fisher information and observation matrices are very similar, as expected. The average lengths of the HPD Bayesian credible intervals are smaller than that of the asymptotic confidence intervals.

In Tables 2.11 and 2.12, the ML, UMVU and Bayesian estimates of R and their corresponding ERs are listed when α is known ($\alpha = 3$). In this case, the Bayes estimates are evaluated analytically under the SE and the LINEX ($v = -1$ and 1) loss functions for different prior parameters. Moreover, it is also computed by using Lindley's approximation and MCMC method. In the Bayesian case, Prior 4: $(a_1, b_1) = (6, 5/2)$, $(a_2, b_2) = (4, 2)$, Prior 5: $(a_1, b_1) = (12, 2)$, $(a_2, b_2) = (3, 3/2)$ and Prior 6: $(a_1, b_1) = (15, 5/4)$, $(a_2, b_2) = (2, 2)$ are used for $R = 0.5484$, 0.7506 and 0.9165, respectively. In addition, the empirical Bayes estimates are obtained. All point estimates of R are listed in Table 2.11. The exact and asymptotic

Table 2.10: Estimates of R using the Priors 1-3 when α is unknown.

			Bayes under the SE		Bayes estimates under the LINEX $v=-1$		Bayes estimates under the LINEX $v=1$		Asymptotic C.I.	Asymptotic C.I.	HPD Credible
(n,m)	R	$\widehat{R}_{MLE}$	$\widehat{R}_{BS,Lindley}$	$\widehat{R}_{BS,MCMC}$	$\widehat{R}_{BL,Lindley}$	$\widehat{R}_{BL,MCMC}$	$\widehat{R}_{BL,Lindley}$	$\widehat{R}_{BL,MCMC}$	based on Fisher	based on Obser.	Interval
(5,5)	0.5006	0.5016	0.4976	0.5005	0.4947	0.5069	0.5018	0.4941	(0.2224,0.7808)	(0.2221,0.7819)	(0.2852,0.7157)
		0.0263	0.0053	0.0070	0.0136	0.0035	0.0046	0.0035	0.5584/0.8640	0.5598/0.8653	0.4305/0.9917
(8,8)		0.5002	0.5007	0.5002	0.5053	0.5050	0.4961	0.4954	(0.2700,0.7304)	(0.2700,0.7304)	(0.3127,0.6877)
		0.0155	0.0039	0.0065	0.0020	0.0033	0.0020	0.0033	0.4604/0.8980	0.4604/0.8977	0.3750/0.9743
(10,10)		0.5035	0.5018	0.5024	0.5064	0.5065	0.4972	0.4982	(0.2940,0.7129)	(0.2940,0.7129)	(0.3276,0.6773)
		0.0112	0.0037	0.0054	0.0019	0.0027	0.0019	0.0027	0.4189/0.9233	0.4189/0.9230	0.3498/0.9780
(12,12)		0.4990	0.4996	0.4993	0.5038	0.5029	0.4954	0.4956	(0.3060,0.6920)	(0.3060,0.6920)	(0.3347,0.6638)
		0.0089	0.0036	0.0047	0.0018	0.0023	0.0018	0.0023	0.3860/0.9390	0.3860/0.9390	0.3291/0.9820
(15,15)		0.5011	0.5007	0.5008	0.5044	0.5039	0.4971	0.4977	(0.3261,0.6761)	(0.3261,0.6761)	(0.3484,0.6529)
		0.0055	0.0025	0.0032	0.0013	0.0016	0.0013	0.0016	0.3500/0.9647	0.3500/0.9647	0.3045/0.9890
(5,5)	0.7145	0.7116	0.6738	0.7051	0.6598	0.7099	0.6813	0.7001	(0.4753,0.9415)	(0.4728,0.9527)	(0.5125,0.8805)
		0.0203	0.0085	0.0053	0.0193	0.0026	0.0033	0.0027	0.4662/0.8367	0.4799/0.8410	0.3680/0.9920
(8,8)		0.7085	0.7013	0.7054	0.7042	0.7090	0.6983	0.7018	(0.5143,0.9027)	(0.5139,0.9032)	(0.5401,0.8584)
		0.0120	0.0029	0.0050	0.0014	0.0025	0.0015	0.0025	0.3884/0.8963	0.3893/0.8990	0.3182/0.9750
(10,10)		0.7096	0.7050	0.7067	0.7083	0.7097	0.7018	0.7036	(0.5346,0.8845)	(0.5346,0.8846)	(0.5546,0.8483)
		0.0087	0.0031	0.0043	0.0015	0.0021	0.0016	0.0022	0.3499/0.9073	0.3500/0.9077	0.2937/0.9717
(12,12)		0.7087	0.7059	0.7064	0.7089	0.7090	0.7030	0.7037	(0.5481,0.8692)	(0.5481,0.8692)	(0.5643,0.8395)
		0.0071	0.0031	0.0039	0.0015	0.0019	0.0016	0.0020	0.3212/0.9267	0.3211/0.9267	0.2752/0.9713
(15,15)		0.7092	0.7066	0.7069	0.7091	0.7091	0.7041	0.7047	(0.5651,0.8532)	(0.5652,0.8532)	(0.5766,0.8294)
		0.0056	0.0030	0.0035	0.0015	0.0017	0.0015	0.0018	0.2881/0.9247	0.2881/0.9247	0.2527/0.9673
(5,5)	0.9095	0.9112	0.8523	0.9001	0.8347	0.9011	0.8615	0.8990	(0.7964,0.9988)	(0.7918,0.9989)	(0.8131,0.9563)
		0.0043	0.0147	0.0010	0.0504	0.0005	0.0052	0.0005	0.2024/0.7897	0.2070/0.8033	0.1432/0.9705
(8,8)		0.9089	0.8965	0.9032	0.8964	0.9040	0.8965	0.9024	(0.8216,0.9935)	(0.8203,0.9939)	(0.8249,0.9673)
		0.0026	0.0008	0.0009	0.0004	0.0005	0.0004	0.0005	0.1719/0.8363	0.1737/0.8433	0.1424/0.9873
(10,10)		0.9055	0.8990	0.9018	0.8995	0.9025	0.8985	0.9011	(0.8266,0.9843)	(0.8260,0.9848)	(0.8301,0.9621)
		0.0019	0.0006	0.0009	0.0003	0.0004	0.0003	0.0005	0.1577/0.8900	0.1588/0.8927	0.1320/0.9853
(12,12)		0.9049	0.9001	0.9017	0.9007	0.9023	0.8995	0.9012	(0.8335,0.9762)	(0.8333,0.9764)	(0.8357,0.9581)
		0.0014	0.0006	0.0008	0.0003	0.0004	0.0003	0.0004	0.1427/0.9153	0.1431/0.9167	0.1224/0.9863
(15,15)		0.9000	0.8975	0.8983	0.8980	0.8988	0.8970	0.8978	(0.8344,0.9656)	(0.8343,0.9657)	(0.8371,0.9517)
		0.0011	0.0006	0.0007	0.0003	0.0004	0.0003	0.0004	0.1312/0.9563	0.1313/0.9563	0.1146/0.9847

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of (n,m). But, for the last three columns, the first row represents 95% confidence interval and the second row represents their lengths and cp's.

confidence intervals are computed from equations (2.123) and (2.126). The Bayesian and empirical Bayesian credible intervals are computed from equations (2.150) and (2.151). The HPD credible interval is constructed by using the MCMC samples. All interval estimates of R are listed in Table 2.12.

From Table 2.11, the ERs of all estimates decrease as the sample sizes increase in all cases, as expected. The Bayes estimates with their corresponding ERs based on Lindley's approximation and MCMC method are very close to the exact values. The ERs of the ML, UMVU, Bayes and empiric Bayes (under the SE loss function) estimates are ordered as $ER(\widehat{R}_{BS}) < ER(\widehat{R}_{EBS}) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_{U})$ when $R = 0.5484$, 0.7506 and $ER(\widehat{R}_{BS}) < ER(\widehat{R}_{U}) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_{EBS})$ when $R = 0.9165$. Moreover, the ERs of the Bayes estimates under the LINEX loss function have smaller than that of ML estimates. From Table 2.12, the average lengths of the intervals decrease as the sample sizes increase. The average lengths of the empirical Bayesian credible intervals are smallest, but their cps are not preferable. The HPD Bayesian credible intervals are more suitable than others in terms of the average lengths and cps.

In the MCMC case, three MCMC chains are run with fairly different initial values and generated 10000 iterations for each chain. To diminish the effect of the starting distribution, the first half of each sequence are discarded and focus on the second half. To provide relatively independent samples for improvement of prediction accuracy, the Bayesian MCMC estimates are calculated by the means of every 5^{th} sampled values after discarding the first half of the chains (see [Gelman et al., 2003]). In our case, the scale factor value of the MCMC estimates are found below 1.1 which is an acceptable value for their convergency.

In Table 2.13, the ML, UMVU and Bayesian estimates of R and their corresponding ERs are listed when α is known ($\alpha = 3$). In this case, the Bayes estimates are evaluated analytically and by using the Lindley's approximation under the SE and the LINEX ($v = -1$ and 1) loss functions for the non informative prior. Moreover, the exact and asymptotic confidence intervals are computed from equations (2.123) and (2.126). The point and interval estimates are computed for $R = 0.25$, 0.33, 0.50, $0, 70$, 0.90 and 0.92 when $(\beta_1, \beta_2) = (2, 6)$, $(2, 4)$, $(2, 2)$, $(7, 3)$, $(18, 2)$ and $(23, 2)$, respectively. From Table 2.13, the ERs of all estimates decrease as the sample

Table 2.11: Estimates of R using the Priors 4-6 when α is known ($\alpha = 3$).

				Bayes estimates under the SE				Bayes estimates under the LINEX							
								$v=-1$				$v=1$			
(n,m)	R	$\widehat{R}_{MLE}$	$\widehat{R}_U$	$\widehat{R}_{BS}$	$\widehat{R}_{BS,Lindley}$	$\widehat{R}_{BS,MCMC}$	$\widehat{R}_{EBS}$	$\widehat{R}_{BL}$	$\widehat{R}_{BL,Lindley}$	$\widehat{R}_{BL,MCMC}$	$\widehat{R}_{EBL}$	$\widehat{R}_{BL}$	$\widehat{R}_{BL,Lindley}$	$\widehat{R}_{BL,MCMC}$	$\widehat{R}_{EBL}$
(5,5)	0.5484	0.5475	0.5511	0.5508	0.5503	0.5508	0.5461	0.5562	0.5565	0.5561	0.5502	0.5453	0.5442	0.5453	0.5421
		0.0191	0.0222	0.0108	0.0108	0.0108	0.0181	0.0054	0.0054	0.0054	0.0091	0.0054	0.0054	0.0054	0.0091
(8,8)		0.5477	0.5499	0.5502	0.5500	0.5502	0.5468	0.5543	0.5545	0.5543	0.5495	0.5461	0.5455	0.5461	0.5440
		0.0127	0.0139	0.0085	0.0085	0.0085	0.0122	0.0043	0.0043	0.0042	0.0061	0.0043	0.0043	0.0043	0.0061
(10,10)		0.5427	0.5442	0.5462	0.5461	0.5462	0.5420	0.5497	0.5499	0.5497	0.5443	0.5427	0.5422	0.5427	0.5397
		0.0100	0.0107	0.0072	0.0072	0.0072	0.0097	0.0036	0.0036	0.0036	0.0048	0.0036	0.0036	0.0036	0.0048
(12,12)		0.5446	0.5460	0.5472	0.5471	0.5472	0.5439	0.5503	0.5504	0.5503	0.5459	0.5441	0.5438	0.5441	0.5420
		0.0085	0.0091	0.0064	0.0064	0.0064	0.0083	0.0032	0.0032	0.0032	0.0042	0.0032	0.0032	0.0032	0.0042
(15,15)		0.5433	0.5444	0.5456	0.5455	0.5456	0.5428	0.5483	0.5484	0.5483	0.5444	0.5430	0.5427	0.5430	0.5412
		0.0069	0.0072	0.0056	0.0056	0.0056	0.0067	0.0028	0.0028	0.0028	0.0034	0.0028	0.0028	0.0028	0.0034
(5,5)	0.7506	0.7347	0.7496	0.7513	0.7502	0.7513	0.7291	0.7543	0.7534	0.7543	0.7321	0.7482	0.7471	0.7482	0.7261
		0.0140	0.0153	0.0062	0.0062	0.0062	0.0137	0.0031	0.0031	0.0031	0.0070	0.0031	0.0031	0.0031	0.0068
(8,8)		0.7426	0.7520	0.7522	0.7516	0.7522	0.7387	0.7545	0.7540	0.7545	0.7406	0.7498	0.7492	0.7498	0.7368
		0.0086	0.0090	0.0047	0.0047	0.0047	0.0085	0.0023	0.0023	0.0023	0.0043	0.0024	0.0023	0.0024	0.0042
(10,10)		0.7417	0.7493	0.7499	0.7494	0.7498	0.7384	0.7519	0.7516	0.7519	0.7400	0.7478	0.7474	0.7478	0.7368
		0.0067	0.0069	0.0041	0.0041	0.0041	0.0066	0.0020	0.0020	0.0020	0.0033	0.0021	0.0021	0.0021	0.0033
(12,12)		0.7404	0.7467	0.7484	0.7481	0.7484	0.7377	0.7502	0.7499	0.7502	0.7390	0.7465	0.7462	0.7465	0.7363
		0.0062	0.0064	0.0038	0.0038	0.0038	0.0062	0.0019	0.0019	0.0019	0.0031	0.0019	0.0019	0.0019	0.0031
(15,15)		0.7463	0.7514	0.7511	0.7509	0.7511	0.7440	0.7527	0.7525	0.7527	0.7450	0.7496	0.7494	0.7496	0.7429
		0.0045	0.0046	0.0032	0.0032	0.0032	0.0045	0.0016	0.0016	0.0016	0.0022	0.0016	0.0016	0.0016	0.0022
(5,5)	0.9165	0.9029	0.9155	0.9165	0.9162	0.9166	0.8977	0.9172	0.9167	0.9172	0.8986	0.9159	0.9156	0.9159	0.8967
		0.0042	0.0037	0.0012	0.0012	0.0012	0.0044	0.0006	0.0006	0.0006	0.0023	0.0006	0.0006	0.0006	0.0022
(8,8)		0.9080	0.9157	0.9156	0.9154	0.9156	0.9042	0.9161	0.9158	0.9161	0.9051	0.9151	0.9150	0.9151	0.9040
		0.0023	0.0021	0.0010	0.0010	0.0010	0.0024	0.0005	0.0005	0.0005	0.0012	0.0005	0.0005	0.0005	0.0012
(10,10)		0.9085	0.9146	0.9147	0.9145	0.9147	0.9056	0.9151	0.9149	0.9151	0.9060	0.9142	0.9141	0.9142	0.9052
		0.0017	0.0016	0.0008	0.0008	0.0008	0.0018	0.0004	0.0004	0.0004	0.0009	0.0004	0.0004	0.0004	0.0009
(12,12)		0.9077	0.9129	0.9132	0.9131	0.9132	0.9052	0.9136	0.9135	0.9136	0.9056	0.9129	0.9128	0.9128	0.9049
		0.0015	0.0014	0.0008	0.0008	0.0008	0.0016	0.0004	0.0004	0.0004	0.0008	0.0004	0.0004	0.0004	0.0008
(15,15)		0.9079	0.9121	0.9119	0.9118	0.9119	0.9059	0.9122	0.9121	0.9122	0.9062	0.9115	0.9115	0.9115	0.9057
		0.0012	0.0011	0.0007	0.0007	0.0007	0.0012	0.0004	0.0004	0.0004	0.0006	0.0004	0.0004	0.0004	0.0006

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of (n,m).

Table 2.12: Confidence intervals of R when α is known ($\alpha = 3$).

(n,m)	R	Exact C.I.	Asymptotic C.I.	Bayes Credible I.	HPD Bayes C.I.	Empiric Bayes C.I.
(5,5)	0.5484	(0.2855,0.7877)	(0.2895,0.8055)	(0.3461,0.7476)	(0.3508,0.7483)	(0.3704,0.7145)
		0.5023/0.9500	0.5160/0.8940	0.4014/0.9473	0.3975/0.9450	0.3441/0.8173
(8,8)		(0.3351,0.7456)	(0.3386,0.7568)	(0.3722,0.7220)	(0.3761,0.7227)	(0.4019,0.6864)
		0.4104/0.9427	0.4182/0.9023	0.3498/0.9440	0.3466/0.9407	0.2845/0.8067
(10,10)		(0.3507,0.7240)	(0.3531,0.7323)	(0.3808,0.7066)	(0.3840,0.7068)	(0.4102,0.6700)
		0.3732/0.9527	0.3792/0.9240	0.3257/0.9497	0.3228/0.9463	0.2598/0.8203
(12,12)		(0.3683,0.7110)	(0.3709,0.7182)	(0.3922,0.6973)	(0.3953,0.6979)	(0.4226,0.6617)
		0.3427/0.9463	0.3473/0.9223	0.3051/0.9457	0.3026/0.9470	0.2391/0.8207
(15,15)		(0.3839,0.6948)	(0.3861,0.7005)	(0.4026,0.6845)	(0.4050,0.6847)	(0.4326,0.6502)
		0.3109/0.9460	0.3144/0.9303	0.2819/0.9440	0.2797/0.9437	0.2175/0.8207
(5,5)	0.7506	(0.4760,0.8997)	(0.5253,0.9441)	(0.5877,0.8828)	(0.6015,0.8910)	(0.5712,0.8554)
		0.4237/0.9497	0.4189/0.8860	0.2951/0.9460	0.2895/0.9497	0.2842/0.8043
(8,8)		(0.5429,0.8789)	(0.5762,0.9090)	(0.6105,0.8681)	(0.6215,0.8748)	(0.6129,0.8424)
		0.3360/0.9463	0.3328/0.9070	0.2576/0.9470	0.2533/0.9493	0.2295/0.8180
(10,10)		(0.5632,0.8678)	(0.5905,0.8929)	(0.6177,0.8593)	(0.6272,0.8652)	(0.6244,0.8339)
		0.3046/0.9533	0.3024/0.9190	0.2416/0.9483	0.2380/0.9417	0.2095/0.8160
(12,12)		(0.5791,0.8579)	(0.6020,0.8789)	(0.6243,0.8521)	(0.6329,0.8574)	(0.6336,0.8261)
		0.2787/0.9423	0.2769/0.9200	0.2278/0.9390	0.2246/0.9377	0.1925/0.8077
(15,15)		(0.6035,0.8524)	(0.6226,0.8700)	(0.6373,0.8470)	(0.6448,.08518)	(0.6511,0.8237)
		0.2489/0.9470	0.2474/0.9333	0.2098/0.9433	0.2070/0.9417	0.1726/0.8297
(5,5)	0.9165	(0.7433,0.9697)	(0.8033,0.9952)	(0.8394,0.9676)	(0.8508,0.9733)	(0.8115,0.9540)
		0.2264/0.9540	0.1918/0.8933	0.1282/0.9613	0.1225/0.9527	0.1425/0.8057
(8,8)		(0.7945,0.9631)	(0.8309,0.9851)	(0.8494,0.9614)	(0.8582,0.9662)	(0.8395,0.9498)
		0.1686/0.9447	0.1541/0.9227	0.1121/0.9473	0.1080/0.9500	0.1103/0.8150
(10,10)		(0.8105,0.9594)	(0.8394,0.9776)	(0.8531,0.9582)	(0.8610,0.9625)	(0.8481,0.9469)
		0.1489/0.9497	0.1382/0.9363	0.1050/0.9493	0.1016/0.9547	0.0988/0.8303
(12,12)		(0.8199,0.9558)	(0.8438,0.9715)	(0.8551,0.9551)	(0.8620,0.9591)	(0.8527,0.9439)
		0.1360/0.9410	0.1277/0.9457	0.1001/0.9350	0.0970/0.9473	0.0912/0.8127
(15,15)		(0.8315,0.9523)	(0.8506,0.9653)	(0.8578,0.9516)	(0.8640,0.9552)	(0.8594,0.9412)
		0.1208/0.9250	0.1147/0.9423	0.0937/0.8953	0.0912/0.9097	0.0818/0.7570

Notes: The first row represents a 95% confidence interval and the second row represents their lengths and cp's.

sizes increase in all cases, as expected. The Bayes estimates under the SE loss function with their corresponding ERs are close to their response in the ML case. Moreover, the Bayes estimates with their corresponding ERs based on Lindley's approximation are very close the exact values. The ERs of the ML, UMVU and Bayes (under the SE loss function) estimates are ordered as $ER(\widehat{R}^*_{BS}) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_U)$ when $R = 0.25, 0.33, 0.50, 0.70$ and $ER(\widehat{R}_U) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}^*_{BS})$ when $R = 0.90$, 0.92. The ERs of ML and Bayes estimates have larger values when the true value of R is around 0.5 and it decreases as the true value of R approaches the extremes. Furthermore, the average lengths of the intervals decrease as the sample sizes increase. When $R = 0.25$, 0.90 and 0.92 the lengths of the asymptotic confidence intervals are smaller than that of exact confidence intervals, but for $R = 0.33$, 0.50 and $0,70$ it is other way around.

Table 2.13: Estimates of R for the non informative prior when α is known ($\alpha = 3$).

						Bayes estimates under the LINEX					
				Bayes estimates under the SE		$v=-1$		$v=1$		Exact C.I.	Asymptotic C.I.
(n,m)	R	$\widehat{R}_{MLE}$	$\widehat{R}_U$	$\widehat{R}^*_{BS}$	$\widehat{R}^*_{BS,Lindley}$	$\widehat{R}^*_{BL}$	$\widehat{R}^*_{BL,Lindley}$	$\widehat{R}^*_{BL}$	$\widehat{R}_{BL,Lindley}$	length/cp	length/cp
(5,5)	0.25(0.33)	0.2604(0.3376)	0.2421(0.3231)	0.2750(0.3487)	0.2789(0.3519)	0.2822(0.3575)	0.2870(0.3624)	0.2681(0.3403)	0.2705(0.3412)	0.4442(0.4956)	0.4388(0.5056)
		0.0157(0.0197)	0.0170(0.0228)	0.0149(0.0177)	0.0150(0.0173)	0.0075(0.0089)	0.0077(0.0078)	0.0075(0.0088)	0.0072(0.0084)	0.9457(0.9473)	0.8807(0.8813)
	0.50(0.70)	0.5004(0.6828)	0.5005(0.6984)	0.5003(0.6707)	0.5003(0.6673)	0.5104(0.6788)	0.5133(0.6776)	0.4903(0.6623)	0.4873(0.6573)	0.5396(0.4851)	0.5658(0.4919)
		0.0218(0.0185)	0.0264(0.0207)	0.0187(0.0171)	0.0178(0.0168)	0.0093(0.0084)	0.0090(0.0080)	0.0093(0.0087)	0.0089(0.0088)	0.9497(0.9507)	0.8943(0.8907)
	0.90(0.92)	0.8790(0.9059)	0.8950(0.9196)	0.8641(0.8926)	0.8614(0.8905)	0.8672(0.8948)	0.8644(0.8925)	0.8609(0.8902)	0.8585(0.8885)	0.2825(0.2353)	0.2515(0.2029)
		0.0054(0.0036)	0.0044(0.0029)	0.0066(0.0046)	0.0070(0.0048)	0.0032(0.0022)	0.0033(0.0023)	0.0034(0.0024)	0.0036(0.0025)	0.9460(0.9487)	0.8937(0.8940)
(8,8)	0.25(0.33)	0.2598(0.3401)	0.2482(0.3310)	0.2699(0.3477)	0.2715(0.3490)	0.2745(0.3535)	0.2764(0.3555)	0.2654(0.3420)	0.2664(0.3424)	0.3628(0.4109)	0.3600(0.4176)
		0.0088(0.0114)	0.0091(0.0125)	0.0086(0.0107)	0.0086(0.0106)	0.0044(0.0054)	0.0044(0.0054)	0.0043(0.0053)	0.0042(0.0052)	0.9583(0.9563)	0.9193(0.9167)
	0.50(0.70)	0.5006(0.6895)	0.5006(0.6998)	0.5006(0.6809)	0.5006(0.6794)	0.5073(0.6862)	0.5085(0.6855)	0.4939(0.6754)	0.4926(0.6734)	0.4472(0.3963)	0.4627(0.4000)
		0.0139(0.0101)	0.0157(0.0107)	0.0125(0.0097)	0.0123(0.0097)	0.0062(0.0047)	0.0061(0.0047)	0.0063(0.0050)	0.0062(0.0050)	0.9563(0.9573)	0.9137(0.9220)
	0.90(0.92)	0.8909(0.9121)	0.9005(0.9205)	0.8817(0.9039)	0.8809(0.9032)	0.8832(0.9050)	0.8823(0.9042)	0.8801(0.9028)	0.8795(0.9023)	0.2030(0.1713)	0.1857(0.1540)
		0.0025(0.0016)	0.0022(0.0014)	0.0029(0.0020)	0.0030(0.0020)	0.0014(0.0010)	0.0015(0.0010)	0.0015(0.0010)	0.0015(0.0010)	0.9517(0.9567)	0.9137(0.9150)
(10,10)	0.25(0.33)	0.2556(0.3393)	0.2463(0.3321)	0.2639(0.3455)	0.2649(0.3464)	0.2675(0.3502)	0.2687(0.3515)	0.2603(0.3409)	0.2610(0.3412)	0.3233(0.3717)	0.3210(0.3767)
		0.0072(0.0093)	0.0074(0.0100)	0.0071(0.0088)	0.0072(0.0088)	0.0036(0.0045)	0.0037(0.0045)	0.0035(0.0044)	0.0035(0.0043)	0.9557(0.9590)	0.9237(0.9220)
	0.50(0.70)	0.4997(0.6887)	0.4997(0.6968)	0.4997(0.6817)	0.4997(0.6808)	0.5053(0.6860)	0.5061(0.6856)	0.4942(0.6773)	0.4934(0.6760)	0.4066(0.3574)	0.4183(0.3599)
		0.0114(0.0092)	0.0126(0.0096)	0.0104(0.0089)	0.0103(0.0089)	0.0052(0.0044)	0.0052(0.0044)	0.0052(0.0045)	0.0052(0.0045)	0.9617(0.9580)	0.9220(0.9220)
	0.90(0.92)	0.8900(0.9135)	0.8976(0.9200)	0.8826(0.9070)	0.8820(0.9066)	0.8838(0.9078)	0.8832(0.9074)	0.8814(0.9062)	0.8809(0.9059)	0.1803(0.1484)	0.1678(0.1360)
		0.0023(0.0014)	0.0020(0.0013)	0.0026(0.0017)	0.0027(0.0017)	0.0013(0.0008)	0.0013(0.0008)	0.0013(0.0008)	0.0013(0.0009)	0.9557(0.9440)	0.9197(0.9177)
(12,12)	0.25(0.33)	0.2545(0.3353)	0.2468(0.3292)	0.2616(0.3407)	0.2623(0.3413)	0.2646(0.3446)	0.2654(0.3455)	0.2586(0.3369)	0.2591(0.3371)	0.2962(0.3401)	0.2942(0.3438)
		0.0060(0.0081)	0.0061(0.0085)	0.0059(0.0077)	0.0059(0.0076)	0.0030(0.0039)	0.0030(0.0039)	0.0029(0.0038)	0.0029(0.0038)	0.9530(0.9520)	0.9197(0.9257)
	0.50(0.70)	0.4963(0.6869)	0.4961(0.6935)	0.4964(0.6810)	0.4964(0.6804)	0.5011(0.6846)	0.5017(0.6843)	0.4917(0.6773)	0.4912(0.6764)	0.3753(0.3290)	0.3846(0.3311)
		0.0096(0.0084)	0.0105(0.0087)	0.0090(0.0081)	0.0089(0.0081)	0.0045(0.0040)	0.0044(0.0040)	0.0045(0.0041)	0.0044(0.0041)	0.9570(0.9440)	0.9283(0.9227)
	0.90(0.92)	0.8928(0.9135)	0.8990(0.9189)	0.8867(0.9081)	0.8863(0.9078)	0.8876(0.9087)	0.8872(0.9084)	0.8857(0.9074)	0.8855(0.9072)	0.1599(0.1343)	0.1503(0.1247)
		0.0019(0.0011)	0.0017(0.0010)	0.0021(0.0013)	0.0021(0.0013)	0.0010(0.0006)	0.0010(0.0007)	0.0010(0.0007)	0.0011(0.0007)	0.9447(0.9550)	0.9157(0.9300)
(15,15)	0.25(0.33)	0.2534(0.3350)	0.2472(0.3301)	0.2592(0.3394)	0.2597(0.3398)	0.2616(0.3426)	0.2622(0.3432)	0.2568(0.3363)	0.2572(0.3364)	0.2663(0.3075)	0.2647(0.3103)
		0.0043(0.0060)	0.0043(0.0063)	0.0042(0.0058)	0.0043(0.0058)	0.0022(0.0029)	0.0022(0.0029)	0.0021(0.0029)	0.0021(0.0029)	0.9647(0.9603)	0.9393(0.9347)
	0.50(0.70)	0.4971(0.6951)	0.4970(0.7006)	0.4972(0.6900)	0.4972(0.6896)	0.5010(0.6929)	0.5014(0.6927)	0.4934(0.6871)	0.4930(0.6866)	0.3401(0.2939)	0.3472(0.2951)
		0.0074(0.0058)	0.0079(0.0060)	0.0070(0.0057)	0.0070(0.0057)	0.0035(0.0028)	0.0035(0.0028)	0.0035(0.0029)	0.0035(0.0029)	0.9630(0.9483)	0.9350(0.9293)
	0.90(0.92)	0.9004(0.9152)	0.9052(0.9195)	0.8956(0.9109)	0.8954(0.9108)	0.8962(0.9114)	0.8960(0.9112)	0.8950(0.9104)	0.8948(0.9103)	0.1342(0.1168)	0.1270(0.1099)
		0.0010(0.0008)	0.0009(0.0008)	0.0010(0.0009)	0.0010(0.0010)	0.0005(0.0005)	0.0005(0.0005)	0.0005(0.0005)	0.0005(0.0005)	0.9490(0.9630)	0.9337(0.9423)

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of (n, m). But, for the last two columns, the first row represents the average length and second row represents their cp's. The corresponding results are reported within bracket in each cell for $R = 0.33$, 0.70 and 0.92.

2.4.3.2. Real Examples

The two real life data sets, lifetime data for insulation specimens and lifetime data for steel specimens, are considered to illustrate the use of the methods proposed in this paper.

- Lifetime data for insulation specimens

The results of a life test experiment in which specimens of a type of electrical insulating fluid were subjected to a constant voltage stress was given in [Nelson, 1972]. The length of time until each specimen failed, or "broke down," was observed. The results for seven groups of specimens, tested at voltages ranging from 26 to 38 kilovolts (kV) were presented. The data sets for 36kV and 38 kV, reported in [Lawless, 2003], are considered and corresponding upper record values are given in Table 2.14. We fit the Burr Type XII distribution to the two data sets. The Kolmogorov-Smirnov (K-S) distances between the fitted and the empirical distribution functions and corresponding p-values, the parameters and the reliability (R) estimates are computed. All these results are presented in Table 2.15. It is observed that the Burr Type XII distribution provides an adequate fit for both the data sets.

Table 2.14: Upper record values from 36kV and 38kV data sets.

i	1	2	3	4
r	1.97	2.58	2.71	25.50
s	0.47	0.73	1.40	2.38

- Lifetime data for steel specimens

The lifetimes of steel specimens tested at 14 different stress levels was given in [Crowder, 2000]. The data sets for 38.5 and 36 stress levels are considered and corresponding upper record values are given in Table 2.16. Since all record values are greater than unity, we encounter the problem for the uniqueness of the ML estimates of the parameters. To overcome this situation, these data sets are divided by the corresponding maximum values. Then, we compute the K-S distances between the fitted and the empirical distribution functions. The K-S and the corresponding p-values, the parameters and the reliability (R) estimates are presented in Table 2.17.

Table 2.15: K-S values and estimates for Table 2.14 when α is common.

Kolmogorov-Smirnov and corresponding p values				
Data Set	K-S(MLE)	p-value	K-S(Lindley)	p-value
r	0.6111	>0.05	0.4796	>0.2
s	0.3879	>0.2	0.4180	>0.2
Parameter and reliability estimates				
Parameter	MLE		Lindley(SEL)	
β_1	0.5468		0.4227	
β_2	1.9134		0.4736	
α	2.2587		1.9249	
R	0.2222		0.3311	

Table 2.16: Upper record values from 38.5 and 36 stress levels.

i	1	2	3	4	5
r	60	83	140	–	–
s	173	218	288	394	585

It is observed that the Burr Type XII distribution provides an adequate fit for both the data sets.

Table 2.17: K-S values and estimates when the first shape parameters α is common.

Kolmogorov-Smirnov and corresponding p values				
Data Set	K-S(MLE)	p-value	K-S(Lindley)	p-value
r	0.5104	>0.2	0.4464	>0.2
s	0.4431	>0.2	0.3098	>0.2
Parameter and reliability estimates				
Parameter	MLE		Lindley(SEL)	
β_1	4.3281		15.1596	
β_2	7.2135		14.3937	
α	2.0278		4.3117	
R	0.3750		0.7283	

2.4.4. Conclusion

In this section, the estimates of the stress-strength reliability based on upper record values are derived when the stress and strength variables follow the Burr Type XII distribution under the non-Bayesian and Bayesian frameworks. The first shape parameters of the distributions of the measurements are assumed to be the same. When the first shape parameters are unknown, the ML and Bayes estimates are obtained by using Lindley's approximation and MCMC method. It is observed that the performance of the Bayes estimates are better than ML estimates. When the first shape parameters are known, the Bayes estimates are obtained exactly and approximately by using the Lindley and MCMC methods for the informative prior case. It is observed that the performance of the Bayes estimates are better than ML and UMVU. Moreover, for the non informative prior case, it is observed that the performance of the Bayes estimates are better than others when the true values of the stress-strength reliability is not close to the extremes $(0$ or $1)$, while near the extremes the UMVU and ML estimates are better than the Bayes estimates. It is observed that the performance of the HPD Bayesian credible interval are better than others in all cases. When the first shape parameter is unknown, it is encouraging to see that the estimates of the stress-strength reliability are very close for the exact and approximate methods when it is known. Furthermore, the Bayes estimates based on the Lindley's approximation and MCMC method are close to each other. Since the cost of time for the MCMC method is more than the Lindley's approximation, the Bayes estimates based on the Lindley's approximation are recommended.

To obtain the point and interval estimates of the stress-strength reliability are difficult due to lack of explicit form of the reliability when the measurements follow from the Burr Type XII distribution with no common parameters. More work is needed along that direction.

3. STATISTICAL ANALYSIS FOR THE GENERALIZED EXPONENTIAL DISTRIBUTION

3.1. Introduction

The distributions from the gamma and the Weibull families are commonly used for analyzing a lifetime data as well as a skewed data. The pros and cons of these distributions were discussed by [Gupta and Kundu, 1999]. They introduced the generalized exponential distribution and pointed out that many of the properties of this distribution are similar to those of gamma and the Weibull families.

If a random variable X follows a two-parameter generalized exponential (GE) distribution, denoted by $GE(\alpha,\lambda)$, then its pdf and cdf are given by

$$F(x;\alpha,\lambda) = (1-e^{-\lambda x})^{\alpha},\ x>0 \tag{3.1}$$

$$f(x;\alpha,\lambda) = \alpha\lambda e^{-\lambda x}(1-e^{-\lambda x})^{\alpha-1},\ x>0 \tag{3.2}$$

where $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters, respectively. The mean and variance of a two-parameter generalized exponential distribution are given by

$$E(X) = \frac{1}{\lambda}\left\{\psi(\alpha+1)-\psi(1)\right\}, \tag{3.3}$$

and

$$Var(X) = \frac{1}{\lambda^2}\left\{\psi^{'}(1)-\psi^{'}(\alpha+1)\right\}, \tag{3.4}$$

where $\psi(x)$ is the digamma function, $\psi^{'}(x)$ its derivative and $\psi(x) = d\ln\Gamma(x)/dx = \Gamma^{'}(x)/\Gamma(x)$.

The density functions of the GE distribution can take different shapes. For $\alpha \leq 1$, it is a decreasing function and for $\alpha > 1$, it is a unimodal, skewed, right tailed similar to the Weibull or gamma density function. It is observed that even for very large shape parameter, it is not symmetric. For $\lambda = 1$, the mode is at $\ln\alpha$ for $\alpha > 1$ and for $\alpha \leq 1$, the mode is at 0.

The GE distribution has a wide range of applications. A comprehensive amount of research has been done over the years from both frequentist and Bayesian perspectives. The GE distribution has been studied extensively by many authors. The properties of the ML estimates of the GE distribution were studied by [Gupta and Kundu, 1999]. The ML etimates were compared with the other estimators like method of moment estimators, estimators based on percentiles, least squares estimators, weighted least squares estimators and the estimators based on the linear combinations of order statistics by [Gupta and Kundu, 2001]. The ML estimates of the unknown parameters of the GE distribution for complete sample as well as censored sample were considered by [Gupta and Kundu, 2002]. An extensive survey of some recent developments for the GE distribution based on a complete random sample was provided by [Gupta and Kundu, 2007]. The Bayes estimators of the unknown parameters of the GE distribution under the assumptions of gamma priors on both the shape and scale parameters were derived by [Kundu and Gupta, 2008]. The statistical inference of the unknown parameters of the GE distribution in presence of progressive censoring were considered by [Pradhan and Kundu, 2009]. The analysis of the hybrid censored data was considered by [Kundu and Pradhan, 2009] when the lifetime distribution of the individual item was the GE distribution.

3.2. Estimation of The Parameters Based on Record Values

Exact expressions for single and product moments of record statistics and the best linear unbiased estimators of the location and scale parameters of the GE distribution were obtained by [Raqab, 2002]. The ML, Bayes and the empirical Bayes estimates of the shape parameter based on lower record values with known scale parameter were derived by [Jaheen, 2004]. Also, prediction bounds for future lower record values was obtained by using Bayes and empirical Bayes techniques. The Bayes estimates of the shape and scale parameters and Bayesian prediction for future lower record values were considered by [Madi and Raqab, 2007]. The Bayesian estimates of the parameters with respect to quadratic loss function using uniform priors for both parameters were obtained by [Sarhan and Tadj, 2008]. Recently, the frequentist and Bayesian estimation of the parameters based on lower record values were obtained by [Dey et al., 2013]. The Bayesian estimates were derived by using symmetric and asymmetric loss functions when the parameters have gamma priors. Also, the Bayesian interval and Bayesian prediction intervals of the future record values were discussed.

3.3. Estimation of The Parameters Based on Records and Inter-Record Times

When the underlying distribution is exponential, estimation of the mean parameter by using record values and their corresponding inter-record times was obtained by [Samaniego and Whitaker, 1986] under random sampling and inverse sampling scheme. The optimal random sampling plan and associated cost analysis for exponential distribution were studied by [Doostparast and Balakrishnan, 2010]. Non-Bayesian and Bayesian estimates were derived by [Doostparast, 2009] for the two parameters exponential distribution based on record values and their corresponding inter-record times under the inverse sampling scheme. The optimal confidence intervals and uniformly most powerful tests for the one-sided alternatives were derived by [Doostparast and Balakrishnan, 2011] when the underlying distribution is two parameter exponential. Also, they obtained the generalized likelihood ratio test, uniformly unbiased and invariant tests for a two-sided alternative. The optimal statistical procedure including point and interval estimation as well as most powerful tests based on record data from a two-parameter Pareto model were obtained by [Doostparast and Balakrishnan, 2013]. When the underlying distribution is lognormal, non-Bayesian and Bayesian point estimates as well as asymptotic confidence intervals for the unknown parameters were obtained by [Doostparast et al., 2013].

In this section, the parameter estimations for GE distribution are obtained by using upper record values and their corresponding inter-record times under the classical and Bayesian frameworks. The Lindley approximation and MCMC technique are proposed to obtain the Bayesian estimates under different loss functions. Moreover, the estimates of the parameters only by using the upper record values (without considering inter-record times) are also obtained. Finally, the two approaches are compared by using Monte Carlo simulations to see the effect of the inter-record times in the estimation.

3.3.1. ML Estimation

Let $X_1, X_2, \ldots$ be i.i.d. random variables, coming from a population with cdf and pdf $F(.)$ and $f(.)$, respectively. Then the likelihood function associated with the sequence $\{R_1, K_1, \ldots, R_m, K_m\}$ is given by [Hofmann and Nagaraja, 2003] as

$$L(\mathbf{r}, \mathbf{k}) = \prod_{i=1}^{m} f(r_i) \{F(r_i)\}^{k_i - 1} I_{(r_{i-1}, \infty)}(r_i), \tag{3.5}$$

where $r_0 \equiv -\infty$, $k_m \equiv 1$ and $I_A(x)$ is the indicator function of the set A. From equations (3.1), (3.2)and(3.5), we have

$$L(\alpha, \lambda; \mathbf{r}, \mathbf{k}) = \alpha^m \lambda^m \exp\left\{ -\lambda \sum_{i=1}^{m} r_i + \sum_{i=1}^{m} (\alpha k_i - 1) \ln(1 - e^{-\lambda r_i}) \right\}, \tag{3.6}$$

where $-\infty < r_1 < \ldots < r_m$ and so the log-likelihood function is

$$l(\alpha, \lambda; \mathbf{r}, \mathbf{k}) = m(\ln \alpha + \ln \lambda) - \lambda \sum_{i=1}^{m} r_i + \sum_{i=1}^{m} (\alpha k_i - 1) \ln(1 - e^{-\lambda r_i}). \tag{3.7}$$

The ML estimates of α and λ are given by

$$\widehat{\alpha} = \frac{m}{U_{\widehat{\lambda}}}, \tag{3.8}$$

where $U_\lambda = -\sum_{i=1}^{m} K_i \ln(1 - e^{-\lambda R_i})$ and $\widehat{\lambda}$ is the solution of the following non-linear equation

$$\frac{m}{\lambda} - \sum_{i=1}^{m} r_i + \sum_{i=1}^{m} \frac{r_i e^{-\lambda r_i}}{1 - e^{-\lambda r_i}} \left(k_i \frac{m}{U_\lambda} - 1 \right) = 0. \tag{3.9}$$

Therefore, $\widehat{\lambda}$ can be obtained as the solution of the non-linear equation of the form $h(\lambda) = \lambda$ where

$$h(\lambda) = m \left[\sum_{i=1}^{m} \frac{r_i}{1 - e^{-\lambda r_i}} - \sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1 - e^{-\lambda r_i}} \frac{m}{U_\lambda} \right]^{-1}. \tag{3.10}$$

Since, $\widehat{\lambda}$ is a fixed point solution of the non-linear equation (3.10), its value can be obtained using an iterative scheme as $\lambda_{(j+1)} = h(\lambda_{(j)})$ where $\lambda_{(j)}$ is the jth iterate of $\widehat{\lambda}$. The iteration procedure should stopped when $\left|\lambda_{(j)} - \lambda_{(j+1)}\right|$ is sufficiently small.

Next, the existence and uniqueness of the ML estimates of the parameters are proved. Following limits and inequalities are used in the proof.

Lemma 3.1: For $\lambda > 0$

$$\left(\sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1-e^{-\lambda r_i}}\right)^2 + \left(\sum_{i=1}^{m} k_i \ln(1-e^{-\lambda r_i})\right)\left(\sum_{i=1}^{m} \frac{k_i r_i^2 e^{-\lambda r_i}}{(1-e^{-\lambda r_i})^2}\right) < 0. \qquad (3.11)$$

Proof 3.1: It is known that the Cauchy-Schwarz inequality is $\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$. *Let* $x_i = \sqrt{k_i} e^{-\frac{\lambda ri}{2}}$ *and* $y_i = \sqrt{k_i} r_i e^{-\frac{\lambda ri}{2}}/(1-e^{-\lambda r_i})$ *then we by the Cauchy-Schwarz inequality we have*

$$\left(\sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1-e^{-\lambda r_i}}\right)^2 \leq \left(\sum_{i=1}^{m} k_i e^{-\lambda r_i}\right)\left(\sum_{i=1}^{m} \frac{k_i r_i^2 e^{-\lambda r_i}}{(1-e^{-\lambda r_i})^2}\right). \qquad (3.12)$$

It can be shown that $k_i e^{-\lambda r_i} < -k_i \ln(1-e^{-\lambda r_i}),\ i = 1, \ldots, m$ *for* $\lambda > 0$. *Let* $f(\lambda) = k_i e^{-\lambda r_i} + k_i \ln(1-e^{-\lambda r_i})$. *It is clear that* $\lim_{\lambda \to 0} f(\lambda) = -\infty$, $\lim_{\lambda \to \infty} f(\lambda) = 0$ *and* $f'(\lambda) > 0$. *Then,* f *is an increasing function and therefore* $f(\lambda) < 0$ *for every* $\lambda > 0$. *The proof is completed by using this inequality in equation (3.12).* ■

The existence and uniqueness of the ML estimates of the parameters of a general class of exponentiated distributions based on a complete sample are proved by [Ghitany et al., 2013]. The following results are used from [Ghitany et al., 2013] to prove the existence and uniqueness of the ML estimates.

Lemma 3.2: $\lim_{t \to 0}(1/t - 1/(1-e^{-t})) = -1/2$, $\lim_{t \to 0} te^{-t}/(1-e^{-t}) = 1$, $\lim_{t \to 0} t\left|\ln(1-e^{-t})\right| = 0$ *and* $\lim_{t \to 0} \left|\ln(1-e^{-t})\right|/e^{-t} = 1$.

Lemma 3.3:

$i)$ *For all* $t > 0,\ t^k e^{-t} < (1-e^{-t})^k,\ k = 1, 2,$

$ii)$ *For all* $-1 < a_i < 1,$ *and* $-\infty < b_i < \infty,\ i = 1, 2, ..., n$

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq \sum_{i=1}^{n} \left|\ln(1-a_i^2)\right| \sum_{i=1}^{n} b_i^2. \tag{3.13}$$

Theorem 3.1: The ML estimates of the parameters α and λ are unique and are given by $\widehat{\alpha} = m/U_{\widehat{\lambda}}$ where $\widehat{\lambda}$ is the solution of the non-linear equation:

$$G(\lambda) = \frac{m}{\lambda} - \sum_{i=1}^{m} \frac{r_i}{1-e^{-\lambda r_i}} + \sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1-e^{-\lambda r_i}} \frac{m}{U_\lambda} = 0. \tag{3.14}$$

Proof 3.1: The limit of $G(\lambda)$ is considered as $\lambda \to 0$ and $\lambda \to \infty$. Let $t_i = \lambda r_i$, $i = 1, \ldots, m$. Then, by using parts of (i)-(iii) of Lemma 1 in [Ghitany et al., 2013],

$$\begin{aligned}
\lim_{\lambda\to 0} G(\lambda) = G(0) &= \sum_{i=1}^{m} r_i \lim_{t_i \to 0} \left(\frac{1}{t_i} - \frac{1}{1-e^{-t_i}}\right) \\
&+ \sum_{i=1}^{m} k_i \lim_{t_i\to 0} \left(\frac{t_i e^{-t_i}}{1-e^{-t_i}}\right) \frac{m}{\sum_{i=1}^{m} (k_i/r_i) \lim_{t_i\to 0} \left(-t_i \ln(1-e^{-t_i})\right)} \\
&= -\frac{1}{2}\sum_{i=1}^{m} r_i + \frac{m}{\sum_{i=1}^{m} (k_i/r_i) \lim_{t_i\to 0} \left(-t_i \ln(1-e^{-t_i})\right)} \sum_{i=1}^{m} k_i = \infty.
\end{aligned} \tag{3.15}$$

Moreover, using part of (iv) of Lemma 1 in [Ghitany et al., 2013], we have

$$\begin{aligned}
\lim_{\lambda\to\infty} G(\lambda) = G(\infty) &= \lim_{\lambda\to\infty} \left(\frac{m}{\lambda} - \sum_{i=1}^{m} \frac{r_i}{1-e^{-\lambda r_i}}\right) \\
&+ \lim_{\lambda\to\infty} \sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1-e^{-\lambda r_i}} \frac{m}{-\sum_{i=1}^{m} k_i \ln(1-e^{-\lambda r_i})} \\
&= -\sum_{i=1}^{m} r_i + m \lim_{\lambda\to\infty} \frac{\sum_{i=1}^{m} k_i r_i e^{-\lambda r_i}}{\sum_{i=1}^{m} k_i e^{-\lambda r_i}} = -\sum_{i=1}^{m} (r_i - r_1) < 0.
\end{aligned} \tag{3.16}$$

Hence, we obtain that $\lim_{\lambda\to 0} G(\lambda) = \infty$ and $\lim_{\lambda\to\infty} G(\lambda) < 0$. By the intermediate value theorem $G(\lambda)$ has at least one root in $(0, \infty)$. If it can be shown that $G^{'}(\lambda) < 0$ then the proof will be completed. It is easily obtain that

$$G'(\lambda) = -\frac{1}{\lambda^2}\left\{G_1(\lambda) - \frac{mG_2(\lambda)\lambda^2}{\left(\sum_{i=1}^m k_i \ln(1-e^{-\lambda r_i})\right)^2}\right\}, \tag{3.17}$$

where

$$G_1(\lambda) = m - \sum_{i=1}^m \frac{\lambda^2 r_i^2 e^{-\lambda r_i}}{(1-e^{-\lambda r_i})^2}, \tag{3.18}$$

$$G_2(\lambda) = \sum_{i=1}^m \frac{k_i r_i^2 e^{-\lambda r_i}}{(1-e^{-\lambda r_i})^2}\left(\sum_{i=1}^m k_i \ln(1-e^{-\lambda r_i})\right) + \left(\sum_{i=1}^m \frac{k_i r_i e^{-\lambda r_i}}{1-e^{-\lambda r_i}}\right)^2. \tag{3.19}$$

$G_1(\lambda) > 0$ *is obtained by using part (i) of Lemma 2 in [Ghitany et al., 2013] and* $G_2(\lambda) < 0$ *from Lemma 3.1. Therefore,* $G'(\lambda) < 0$.

Finally, we will show that the ML estimates of (α, λ) *maximizes the log-likelihood function* $l(\alpha, \lambda; \mathbf{r}, \mathbf{k})$. *Let* $H(\alpha, \lambda)$ *be the Hessian matrix of* $l(\alpha, \lambda; \mathbf{r}, \mathbf{k})$ *at* (α, λ). *It is clear that* $H_{11}(\widehat{\alpha}, \widehat{\lambda}) < 0$ *and the determinant of the Hessian matrix*

$$\begin{aligned} D(\widehat{\alpha}, \widehat{\lambda}) &= H_{11}(\widehat{\alpha}, \widehat{\lambda})H_{22}(\widehat{\alpha}, \widehat{\lambda}) - \left(H_{12}(\widehat{\alpha}, \widehat{\lambda})\right)^2 \\ &= \frac{m}{\widehat{\alpha}^2\widehat{\lambda}^2}G_1(\widehat{\lambda}) - G_2(\widehat{\lambda}) > 0. \end{aligned} \tag{3.20}$$

Hence, $(\widehat{\alpha}, \widehat{\lambda})$ *is the local maximum of* $l(\alpha, \lambda; \mathbf{r}, \mathbf{k})$. *Since there is no singular point of* $l(\alpha, \lambda; \mathbf{r}, \mathbf{k})$ *and it has a single critical point then, it is enough to show that the absolute maximum of the function is indeed the local maximum. Assume that there exist a* $\widehat{\lambda}_0$ *in the domain in which* $l^*(\widehat{\lambda}_0) > l^*(\widehat{\lambda})$, *where* $l^*(\widehat{\lambda}) = l(\widehat{\alpha}, \widehat{\lambda}; \mathbf{r}, \mathbf{k})$. *Since* $\widehat{\lambda}$ *is the local maximum there should be some point* λ_1 *in the neighborhood of* $\widehat{\lambda}$ *such that* $l^*(\widehat{\lambda}) > l^*(\lambda_1)$. *Let* $k(\lambda) = l^*(\lambda) - l^*(\widehat{\lambda})$ *then* $k(\widehat{\lambda}_0) > 0$, $k(\lambda_1) < 0$ *and* $k(\widehat{\lambda}) = 0$. *This implies that* λ_1 *is a local minimum of the* $l^*(\lambda)$, *but* $\widehat{\lambda}$ *is the only critical point so it is a contradiction. Therefore,* $(\widehat{\alpha}, \widehat{\lambda})$ *is the absolute maximum of* $l(\alpha, \lambda; \mathbf{r}, \mathbf{k})$. ■

3.3.1.1. ML Estimation When λ Is Known

Without loss of generality, the parameter λ is assumed to be known, say $\lambda = 1$. Then, from equation (3.6)

$$L(\alpha, 1; \mathbf{r}, \mathbf{k}) = \alpha^m \exp\left\{-\sum_{i=1}^{m} r_i + \sum_{i=1}^{m} (\alpha k_i - 1)\ln(1 - e^{-r_i})\right\}, \tag{3.21}$$

where $-\infty < r_1 < \ldots < r_m$. In this case, U_1 is a sufficient statistic for α and the MLE of α is $\widehat{\alpha}_{ML} = m/U_1$. The moment generating function of U_1 is $M(t) = 1/\left(1 - t/\alpha\right)^m$, $\alpha > t$. By the uniqueness of the moment generating function U_1, is distributed as $Gamma(m, \alpha)$ and its mean and variance are m/α and m/α^2, respectively. Therefore, $E(\widehat{\alpha}_{ML}) = m\alpha/(m-1)$ and an unbiased estimator for α is given by $\widehat{\alpha}_U = (m-1)/U_1$. Notice that, $MSE(\widehat{\alpha}_U) < MSE(\widehat{\alpha}_{ML})$ and $MSE(\widehat{\alpha}_{ML}) \to 0$ as $m \to \infty$ then $\widehat{\alpha}_{ML}$ and $\widehat{\alpha}_U$ converge to α in mean square.

When the scale parameter $\lambda = \lambda_0$ is known, then U_{λ_0} is a complete sufficient statistic for α. The confidence interval of α is constructed based on this statistic. The distribution of U_{λ_0} can be easily obtained. We have $U_{\lambda_0} \sim Gamma\,(m, \alpha)$ and $2\alpha U_{\lambda_0} \sim \chi^2_{2m}$. An equi-tailed $100(1-\gamma)\%$ confidence interval of the shape parameter α has the form

$$\left(\frac{\chi^2_{2m,\gamma/2}}{2U_{\lambda_0}}, \frac{\chi^2_{2m,1-\gamma/2}}{2U_{\lambda_0}}\right). \tag{3.22}$$

3.3.1.2. Asymptotic Confidence Interval

To obtain the exact confidence interval for the parameters is not easy in every case, so that their asymptotic behavior constitutes an appealing alternative. In practice, the observed information matrix is used as a consistent estimator of the Fisher information matrix. An asymptotic confidence intervals for the parameters α and λ based on the record values and their corresponding inter-record times are obtained by using the observed information matrix. The observed information matrix $\mathbf{J}_m(\alpha, \lambda)$ is given by

$$\mathbf{J}_m(\alpha, \lambda) = -\begin{bmatrix} \frac{\partial^2 l}{\partial \alpha^2} & \frac{\partial^2 l}{\partial \alpha \partial \lambda} \\ \frac{\partial^2 l}{\partial \lambda \partial \alpha} & \frac{\partial^2 l}{\partial \lambda^2} \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}, \tag{3.23}$$

where $J_{11} = m/\alpha^2$, $J_{12} = -\sum_{i=1}^{m}(k_i r_i e^{-\lambda r_i})/(1 - e^{-\lambda r_i})$ and $J_{22} = m/\lambda^2 - \sum_{i=1}^{m}(1 - \alpha k_i)(r_i^2 e^{-\lambda r_i})/\left(1 - e^{-\lambda r_i}\right)^2$.

By the asymptotic normality of the MLE, we have $\left[\sqrt{m}(\widehat{\alpha}-\alpha), \sqrt{m}(\widehat{\lambda}-\lambda)\right] \overset{a}{\sim} N_2(\mathbf{0}, \mathbf{I}^{-1})$ for large m, where $\overset{a}{\sim}$ means approximately distributed and $\mathbf{I}^{-1}$ is the inverse of the observed information matrix. The asymptotic behavior remains valid if $\mathbf{I} = \lim_{m\to\infty}(1/m)\mathbf{J}_m(\widehat{\alpha}, \widehat{\lambda})$. For large m (the number of record values) under inverse sampling or for large n (the number of observations) under random sampling scheme, we can construct the approximate $100(1-\gamma)\%$ equi-tailed confidence intervals for α and λ. These are given by

$$\left(\widehat{\alpha} \pm z_{1-\gamma/2}\sqrt{\frac{J_{22}}{J_{11}J_{22}-J_{12}^2}}\right) \text{ and } \left(\widehat{\lambda} \pm z_{1-\gamma/2}\sqrt{\frac{J_{11}}{J_{11}J_{22}-J_{12}^2}}\right), \tag{3.24}$$

where $z_{\gamma/2}$ is the upper $\gamma/2$th quantile of the standard normal distribution.

3.3.2. Bayesian Estimation

In this section, the Bayes estimates of the parameters GE distribution are obtained by using different loss functions for both λ is known and unknown cases under the inverse sampling scheme.

3.3.2.1. Bayesian Estimation When λ Is Known

It is assumed that α has a gamma prior with parameters (a_1, b_1). Then, the posterior density function of α is $\alpha|\left(\mathbf{r},\mathbf{k}\right) \sim Gamma\left(m+a_1, b_1+U_1\right)$. Then, the Bayes estimate of α under the SE loss function, $\widehat{\alpha}_{BS,1}$, is the mean of the $\alpha|\left(\mathbf{r},\mathbf{k}\right)$. Therefore,

$$\widehat{\alpha}_{BS,1} = \frac{m+a_1}{b_1+U_1}, \tag{3.25}$$

and the Bayes estimate of α under the LINEX loss function, $\widehat{\alpha}_{BL,1}$, is

$$\widehat{\alpha}_{BL,1} = -\frac{1}{v}\ln E_{\alpha|(\mathbf{r},\mathbf{k})}(e^{-v\alpha}) = \frac{m+a_1}{v}\ln\left(1+\frac{v}{b_1+U_1}\right). \tag{3.26}$$

If we use the Jeffrey's non-informative prior, that is $\pi(\alpha) = 1/\alpha$, then we have $\alpha|\left(\mathbf{r},\mathbf{k}\right) \sim Gamma\left(m, U_1\right)$. Therefore, the Bayes estimates of α under the SE and the LINEX loss functions are obtained as

$$\widehat{\alpha}_{BS,0} = \frac{m}{U_1} \text{ and } \widehat{\alpha}_{BL,0} = \frac{m}{v}\ln\left(1+\frac{v}{U_1}\right), \tag{3.27}$$

respectively. $\widehat{\alpha}_{BS,0}$ and $\widehat{\alpha}_{BL,0}$ are the limit of $\widehat{\alpha}_{BS,1}$ and $\widehat{\alpha}_{BL,1}$ as $a_1 \to 0$ and $b_1 \to 0$. Moreover, $\widehat{\alpha}_{BL,1} \to \widehat{\alpha}_{BS,1}$ as $v \to 0$ is satisfied.

Notice that, it is easily seen that if $mb_1 > a_1U_1$ then $\widehat{\alpha}_M > \widehat{\alpha}_{BS,1}$ and if $mb_1 < a_1U_1$ then $\widehat{\alpha}_M < \widehat{\alpha}_{BS,1}$. In the following proposition, the comparison of Bayes estimates are given under the SE and the LINEX loss functions.

Proposition 3.1:

i) $\widehat{\alpha}_{BL,1} \leq \widehat{\alpha}_{BS,1}$ *for* $v > 0$.

ii) $\widehat{\alpha}_{BL,1} \geq \widehat{\alpha}_{BS,1}$ *for* $-(b_1+U_1) < v < 0$.

Proof 3.1: It is known that

$$\ln(1+x) \leq x \text{ for every } x > -1. \tag{3.28}$$

i) Suppose $v > 0$. $v/(b_1+U_1) > 0$, *when* $b_1 > 0$ *and* $U_1 > 0$. *We have* $\ln\left(1+v/(b_1+U_1)\right) \leq v/(b_1+U_1)$ *by the inequality (3.28). Therefore,* $\widehat{\alpha}_{BL,1} \leq \widehat{\alpha}_{BS,1}$

ii) Suppose $v < 0$ *and* $-(b_1+U_1) < v$, *then* $v/(b_1+U_1) > -1$. *We have* $\ln\left(1+v/(b_1+U_1)\right) \leq v/(b_1+U_1)$ *by the inequality (3.28). Therefore,* $\widehat{\alpha}_{BL,1} \geq \widehat{\alpha}_{BS,1}$. ■

3.3.2.2. Bayesian Estimation When α and λ Are Unknown

Assume that the parameters α and λ have independent gamma priors with parameters (a_1, b_1) and (a_2, b_2), respectively and densities are denoted by $\pi(\alpha)$ and $\pi(\lambda)$. The joint posterior density function of α and λ is

$$\pi(\alpha,\lambda|\,\mathbf{r},\mathbf{k}) = \frac{L(\alpha,\lambda;\mathbf{r},\mathbf{k})\pi(\alpha)\pi(\lambda)}{\int_0^\infty\int_0^\infty L(\alpha,\lambda;\mathbf{r},\mathbf{k})\pi(\alpha)\pi(\lambda)d\alpha d\lambda}$$

$$= I(\mathbf{r},\mathbf{k})\alpha^{m+a_1-1}\lambda^{m+a_2-1} \tag{3.29}$$

$$\exp\left\{-\lambda\left(b_2+\sum_{i=1}^{m} r_i\right) - \alpha\,(b_1+U_\lambda) - \sum_{i=1}^{m}\ln(1-e^{-\lambda r_i})\right\},$$

where

$$\frac{[I(\mathbf{r},\mathbf{k})]^{-1}}{\Gamma(m+a_1)} = \int_0^\infty \lambda^{m+a_2-1}\,(b_1+U_\lambda)^{-m-a_1}$$

$$\exp\left\{-\lambda\left(b_2+\sum_{i=1}^{m} r_i\right) - \sum_{i=1}^{m}\ln(1-e^{-\lambda r_i})\right\}d\lambda. \tag{3.30}$$

The Bayes estimate of any function of α and λ, say $g(\alpha,\lambda)$ under the SE loss function is

$$\widehat{g}_{BS} = E_{\alpha,\lambda|\mathbf{r},\mathbf{k}}(g(\alpha,\lambda)) = \frac{\int_0^\infty\int_0^\infty g(\alpha,\lambda)L(\alpha,\lambda;\mathbf{r},\mathbf{k})\pi(\alpha)\pi(\lambda)d\alpha d\lambda}{\int_0^\infty\int_0^\infty L(\alpha,\lambda;\mathbf{r},\mathbf{k})\pi(\alpha)\pi(\lambda)d\alpha d\lambda}. \tag{3.31}$$

It is not possible to compute (3.31) analytically. Two approaches are suggested here to approximate equation (3.31), namely i) Lindley's approximation and ii) MCMC method.

- Lindley's approximation

  For the two parameter case (α,λ), we have from equation (3.29)

$$Q = \ln I(\mathbf{r},\mathbf{k}) + (m+a_1-1)\ln\alpha + (m+a_2-1)\ln\lambda - \lambda b_2$$

$$-\lambda\sum_{i=1}^{m} r_i - \alpha\,(b_1+U_\lambda) - \sum_{i=1}^{m}\ln(1-e^{-\lambda r_i}). \tag{3.32}$$

The joint posterior mode is obtained from the equations $\partial Q/\partial\alpha = 0$ and $\partial Q/\partial\lambda = 0$ as

$$\widetilde{\alpha} = \frac{m + a_1 - 1}{b_1 + U_{\widetilde{\lambda}}}, \tag{3.33}$$

and $\widetilde{\lambda}$ is the solution of the following nonlinear equation

$$\frac{m + a_2 - 1}{\lambda} - b_2 - \sum_{i=1}^{m} \frac{r_i}{1 - e^{-\lambda r_i}} + \frac{m + a_1 - 1}{b_1 + U_\lambda} \sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1 - e^{-\lambda r_i}} = 0. \tag{3.34}$$

It can be solved by using the same procedure in equation (3.10). The elements of the Q^* are given by

$$Q^*_{11} = -\frac{m + a_1 - 1}{\alpha^2},\ Q^*_{12} = Q^*_{21} = \sum_{i=1}^{m} \frac{k_i r_i e^{-\lambda r_i}}{1 - e^{-\lambda r_i}}, \tag{3.35}$$

$$Q^*_{22} = -\frac{m + a_2 - 1}{\lambda^2} + \sum_{i=1}^{m} (1 - \alpha k_i) \frac{r_i^2 e^{-\lambda r_i}}{(1 - e^{-\lambda r_i})^2}, \tag{3.36}$$

and $\tau_{ij},\ i, j = 1, 2$ are obtained by using equation (3.35) and (3.36). Moreover,

$$Q_{12} = -\sum_{i=1}^{m} \frac{k_i r_i^2 e^{-\lambda r_i}}{(1 - e^{-\lambda r_i})^2},\ Q_{21} = 0,\ Q_{30} = \frac{2\,(m + a_1 - 1)}{\alpha^3}, \tag{3.37}$$

$$Q_{03} = \frac{2\,(m + a_2 - 1)}{\lambda^3} - \sum_{i=1}^{m} (1 - \alpha k_i) \frac{r_i^3 (e^{-\lambda r_i} + e^{-2\lambda r_i})}{(1 - e^{-\lambda r_i})^3}. \tag{3.38}$$

Therefore, the approximate Bayes estimates of α and λ under the SE and LINEX loss functions are obtained as

$$\widehat{\alpha}_{BS,Lind} = \widetilde{\alpha} + \frac{1}{2} \left[Q_{30} \tau_{11}^2 + Q_{12} (\tau_{22} \tau_{11} + 2\tau_{21}^2) + Q_{03} \tau_{21} \tau_{22} \right], \tag{3.39}$$

$$\widehat{\alpha}_{BL,Lind} = \widetilde{\alpha} - \frac{1}{v} \ln \left[1 + \frac{v^2}{2} \tau_{11} - \frac{v}{2} Q_{12} (\tau_{22} \tau_{11} + 2\tau_{21}^2) - \frac{v}{2} Q_{30} \tau_{11}^2 - \frac{v}{2} Q_{03} \tau_{21} \tau_{22} \right], \tag{3.40}$$

$$\widehat{\lambda}_{BS,Lind} = \widetilde{\lambda} + \frac{1}{2}\left[Q_{30}\tau_{12}\tau_{11} + 3Q_{12}\tau_{22}\tau_{21} + Q_{03}\tau_{22}^{2}\right], \tag{3.41}$$

$$\widehat{\lambda}_{BS,Lind} = \widetilde{\lambda} - \frac{1}{v}\ln\left[1 + \frac{v}{2}\left(v\tau_{22} - Q_{30}\tau_{12}\tau_{11} - 3Q_{12}\tau_{22}\tau_{21} - Q_{03}\tau_{22}^{2}\right)\right]. \tag{3.42}$$

Notice that all approximate Bayes estimates are evaluated at $(\widetilde{\alpha}, \widetilde{\lambda})$.

If we use the non-informative prior density, that is $\pi(\alpha, \lambda) \propto 1/\alpha\lambda$, then the Bayes estimates of α and λ can be computed by using the Lindley's approximation. In this case, these estimates are easily obtained from equations (3.39)-(3.42) using $a_1 = a_2 = b_1 = b_2 = 0$.

- MCMC method

In the previous section, we obtain the Bayes estimates of α and λ using Lindley's approximation under the SE and LINEX loss functions. However, since the exact probability distributions of these estimates are not known it is difficult to evaluate HPD credible intervals of parameters. For this reason, the MCMC method is proposed to compute the Bayes estimates of α and λ under the SE and LINEX loss functions and their corresponding HPD credible intervals.

The MCMC method are considered to generate samples from the posterior distributions and then the Bayes estimates of α and λ under the SE and LINEX loss functions are computed. The joint posterior of α and λ is given in equation (3.29). It is easy to see that

$$\alpha|\, \lambda, \mathbf{r}, \mathbf{k} \sim Gamma\left(m + a_1, b_1 + U_\lambda\right), \tag{3.43}$$

and

$$\pi(\lambda|\, \alpha, \mathbf{r}, \mathbf{k}) \propto \lambda^{m+a_2+1} \exp\left\{-\lambda\left(b_2 + \sum_{i=1}^{m} r_i\right)\right\}$$
$$\exp\left\{-\alpha U_\lambda - \sum_{i=1}^{m}\ln(1 - e^{-\lambda r_i})\right\}. \tag{3.44}$$

Therefore, samples of α can be easily generated by using gamma distribution. However, the posterior distribution of λ cannot be reduced analytically to well known distribution and therefore it is not possible to sample directly by standard methods. It is observed that the plot of the posterior distribution of λ is similar to Gaussian distribution. The hybrid Metropolis -Hastings and Gibbs sampling algorithm, which will be used to solve our problem, is suggested by [Tierney, 1994]. This algorithm combines the Metropolis-Hastings with Gibbs sampling scheme under the Gaussian proposal distribution.

- Step 1: Take some initial guess of α and λ, say $\alpha^{(0)}$ and $\lambda^{(0)}$.
- Step 2: Set $t = 1$.
- Step 3: Generate $\lambda^{(t)}$ from $\pi(\lambda|\, \alpha, \mathbf{r}, \mathbf{k})$ using the Metropolis-Hastings algorithm with the proposal distribution $q(\lambda) \equiv N(\lambda^{(t-1)}, 1)$:

  -Step 3.1: Let $v = \lambda^{(t-1)}$.

  -Step 3.2: Generate w from the proposal distribution q.

  -Step 3.3: Let $p(v, w) = \min\left\{1, \dfrac{\pi(w|\, \alpha^{(t-1)}, \mathbf{r}, \mathbf{k})\; q(v)}{\pi(v|\, \alpha^{(t-1)}, \mathbf{r}, \mathbf{k})\; q(w)}\right\}$.

  -Step 3.4: Generate u from $Uniform(0, 1)$. If $u \leq p(v, w)$ then accept the proposal and set $\lambda^{(t)} = w$; otherwise, set $\lambda^{(t)} = v$.
- Step 4: Generate $\alpha^{(t)}$ from $Gamma\left(m + a_1, b_1 - \sum\limits_{i=1}^{m} K_i \ln(1 - e^{-\lambda^{(t)} R_i})\right)$.
- Step 5: Set $t = t + 1$.
- Step 6: Repeat Steps 3-5, N times, and obtain the posterior samples $(\alpha^{(i)}, \lambda^{(i)})$, $i = 1, \ldots, N$.

The samples obtained from the algorithm are used to compute the Bayes estimates and to construct the HPD credible intervals. The Bayes estimator of $g \equiv g(\alpha, \lambda)$ based on SE and LINEX loss function are given, respectively, by

$$\widehat{g}_{BS,MH} = E(g|\, \mathbf{r}, \mathbf{k}) = \frac{1}{N - M} \sum_{i=M+1}^{N-M} g(\alpha^{(i)}, \lambda^{(i)}), \tag{3.45}$$

and

$$\widehat{g}_{BL,MH} = -\frac{1}{v} \ln\left[\frac{1}{N - M} \sum_{i=M+1}^{N-M} \exp\left(-v\; g(\alpha^{(i)}, \lambda^{(i)})\right)\right], \tag{3.46}$$

where M is the burn-in period.

The HPD $100(1-\gamma)\%$ credible intervals of α and λ can be obtained by the method of [Chen and Shao, 1999]. In particular for α:

From MCMC, the sequence of α, $\alpha_1, \ldots, \alpha_N$, are obtained, then it is ordered as $\alpha_{(1)} < \ldots < \alpha_{(N)}$. The credible intervals are constructed as $\left(\alpha_{(j)}, \alpha_{(j+[N(1-\gamma)])}\right)$ for $j = 1, ..., N - [N(1-\gamma)]$ where $[x]$ denotes the largest integer less than or equal to x. Then, the HPD credible interval of α is that interval which has the shortest length. Similarly, the HPD credible interval of λ can also be constructed.

3.3.3. Simulation Study

In this section, in order to compare the proposed point and interval estimates for the Bayesian and ML cases, we perform a Monte Carlo simulation studying using different sample sizes and different priors. All the programs are written in MATLAB R2010a. All the results are based on 1000 replications. The ER of θ, when θ is estimated by $\widehat{\theta}$, is given by

$$ER(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left(\widehat{\theta}_i - \theta_i\right)^2, \tag{3.47}$$

under the SE loss function. Moreover, the estimated risk of θ under the LINEX loss function is given by

$$ER(\theta) = \frac{1}{N} \sum_{i=1}^{N} \left(e^{v\left(\widehat{\theta}_i - \theta_i\right)} - v\left(\widehat{\theta}_i - \theta_i\right) - 1\right), \tag{3.48}$$

where N is the number of replication.

A simulation study is carried out to investigate the performance of the point and interval estimation of the GE distribution parameters based on ML and Bayesian methods. If we use the non-informative prior density, it is known that the Bayes estimates of α and λ are similar to the ML estimates for a large number of observations. On the other hand, the informative priors are chosen to be gamma distribution with parameters Prior 1: $(a_1, b_1) = (5, 2)$, $(a_2, b_2) = (10, 5)$ and Prior 2: $(a_1, b_1) = (10, 3)$, $(a_2, b_2) = (7, 1.5)$ for the parameters $(\alpha, \lambda) = (2.5362, 2.0154)$ and $(3.3711, 4.7148)$ in Tables 3.1 and 3.2, respectively. The findings based on ML and Bayesian methods are given in the Tables 3.1-3.3 when the sample size m is $5, 8$ and 10.

In Tables 3.1 and 3.2, the ML estimates and their corresponding ERs are presented. Furthermore, a %95 HPD credible intervals and their coverage probabilities (cp) are presented. Moreover, the Bayes estimates are computed by using Lindley's approximation and MCMC method under SE and LINEX ($v = -2, -1, 1$ and 2) loss functions.

In the MCMC case, five MCMC are run chains with fairly different initial values and generated 5000 iterations for each chain. To diminish the effect of the starting distribution, the first half of each sequence are discarded and focus attention on the second half. To provide relatively independent samples for improvement of prediction accuracy, the Bayesian MCMC estimates are calculated by the means of every 5^{th} sampled values after discarding the first half of the chains (see [Gelman et al., 2003]). In our case, the scale factor value of the MCMC estimators are found below 1.1 which is an acceptable value for their convergency.

For convenience, the estimates based on the record values and their corresponding inter-record times are denoted by α and λ. On the other hand, estimates based on just the record values are denoted by α^* and λ^*.

To be able to compare the two methods, the ML and Bayes estimates of the parameters are also derived based on upper record values (without taking inter-record times into consideration). In this case, the Bayes estimates are obtained by using Lindley's approximation method. The estimates α^* and λ^* and their corresponding ERs are tabulated in Table 3.3.

The following results are obtained from Tables 3.1-3.3 under the SE loss function. The average ERs of α and λ decrease as the sample size increases in all cases, as expected. Moreover, the ERs of Bayes estimates under the SE loss function are smaller than that of ML estimates. Furthermore, the ERs of Bayes estimates for α and λ based on Lindley's approximation and MCMC samples are close to each other. In general, similar patterns are observed for α^* and λ^*. Finally, the ERs of estimates for α and λ are smaller than that of α^* and λ^*. It is quite natural to see such a result when more information is available.

The following results are obtained from Tables 3.1-3.3 under the LINEX loss function. The ERs are smaller than that of ML estimates only for some v, since the

Table 3.1: Results for the true values of $(\alpha, \lambda) = (2.5362, 2.0154)$ using Prior 1.

		Bayesian estimates using Lindley					Bayesian estimates using MCMC					HPD credible interval
			LINEX					LINEX				
	MLE	SE	$v=-2$	$v=-1$	$v=1$	$v=2$	SE	$v=-2$	$v=-1$	$v=1$	$v=2$	
							$m=5$					
α	2.1511	2.2639	2.4384	2.4136	1.9518	1.7025	2.2952	3.6509	2.8515	1.9718	1.7530	(0.8068,4.0379)
	1.7093	0.9133	28.4919	0.9535	0.3660	1.2572	0.8946	3.0446	0.6010	0.3596	1.2171	0.9440
λ	1.7093	1.8485	1.9100	1.8858	1.7991	1.7640	1.8673	1.9953	1.9285	1.8110	1.7587	(1.2574,2.5077)
	0.2874	0.1443	0.3810	0.0797	0.0687	0.2659	0.1366	0.2791	0.0693	0.0666	0.2580	0.9310
							$m=8$					
α	2.3291	2.2543	2.4286	2.4024	1.9498	1.7058	2.2731	3.5448	2.7950	1.9634	1.7512	(0.8170,3.9635)
	1.6100	0.8831	30.5882	0.9394	0.3561	1.2273	0.8726	2.6743	0.5782	0.3504	1.1863	0.9370
λ	1.8689	1.9189	1.9503	1.9367	1.8976	1.8745	1.9230	1.9712	1.9468	1.8999	1.8774	(1.5382,2.3196)
	0.0894	0.0532	0.1179	0.0276	0.0262	0.1046	0.0523	0.1058	0.0263	0.0260	0.1033	0.9250
							$m=10$					
α	2.3433	2.2517	2.4268	2.3982	1.9586	1.7241	2.2651	3.4469	2.7512	1.9697	1.7648	(0.8477,3.9056)
	1.5637	0.8649	25.4029	0.8589	0.3487	1.1969	0.8577	2.3813	0.5412	0.3443	1.1631	0.9260
λ	1.9009	1.9387	1.9590	1.9498	1.9259	1.9121	1.9405	1.9686	1.9544	1.9268	1.9133	(1.6444,2.2418)
	0.0532	0.0316	0.0671	0.0161	0.0157	0.0625	0.0312	0.0621	0.0156	0.0155	0.0619	0.9320

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. However, for the last columns, the first row represents the %95 confidence interval and the second row represents their cp's.

Table 3.2: Results for the true values of $(\alpha, \lambda) = (3.3711, 4.7148)$ using Prior 2.

		Bayesian estimates using Lindley					Bayesian estimates using MCMC					HPD credible interval
			LINEX				SE	LINEX				
	MLE	SE	$v=-2$	$v=-1$	$v=1$	$v=2$		$v=-2$	$v=-1$	$v=1$	$v=2$	
						$m=5$						
α	2.9144	3.1385	3.4138	3.3642	2.7511	2.5048	3.1937	4.6141	3.8051	2.7925	2.5054	(1.4641,5.1233)
	2.9420	0.9147	7.5367	0.6633	0.3947	1.3472	0.8917	2.2063	0.5018	0.3856	1.3541	0.9540
λ	4.0468	4.3697	4.5692	4.5194	4.1357	3.9722	4.4970	5.4027	4.9427	4.1497	3.8879	(3.0031,6.0987)
	1.3912	0.6696	8.1078	0.5151	0.3027	1.0632	0.5575	1.1714	0.2718	0.3025	1.1737	0.9610
						$m=8$						
α	3.0682	3.1328	3.4050	3.3539	2.7578	2.5170	3.1586	4.4458	3.7127	2.7863	2.5150	(1.4979,5.0038)
	2.6962	0.8964	6.8530	0.6523	0.3843	1.3139	0.8844	2.0717	0.4987	0.3776	1.3212	0.9340
λ	4.3847	4.5417	4.6505	4.6117	4.4484	4.3657	4.5701	4.8312	4.6953	4.4577	4.3565	(3.7349,5.4430)
	0.4750	0.2363	1.2247	0.1506	0.1093	0.4117	0.2235	0.4614	0.1179	0.1080	0.4232	0.9600
						$m=10$						
α	3.0538	3.1252	3.3937	3.3415	2.7622	2.5263	3.1408	4.3557	3.6622	2.7859	2.5246	(1.5246,4.9365)
	2.4358	0.8721	6.2661	0.6196	0.3786	1.3011	0.8640	1.9980	0.4768	0.3724	1.3053	0.9420
λ	4.4482	4.5790	4.6533	4.6240	4.5227	4.4689	4.5911	4.7307	4.6587	4.5277	4.4681	(3.9650,5.2391)
	0.3044	0.1438	0.3925	0.0792	0.0691	0.2692	0.1397	0.3094	0.0733	0.0680	0.2698	0.9300

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. However, for the last columns, the first row represents the %95 confidence interval and the second row represents their cp's.

Table 3.3: Estimates of α and λ based on only upper records.

True values of α and λ	MLE	Bayesian estimates using Lindley: SE	LINEX $v=-2$	LINEX $v=-1$	LINEX $v=1$	LINEX $v=2$	True values of α and λ	MLE	Bayesian estimates using Lindley: SE	LINEX $v=-2$	LINEX $v=-1$	LINEX $v=1$	LINEX $v=2$
							$m=5$						
α^*=2.5362	6.8768	2.3806	2.5888	2.5671	1.9776	1.6935	α^*=3.3711	10.6858	3.3175	3.6327	3.5816	2.8687	2.6146
	279.27	1.0592	43.3161	1.1310	0.4155	1.3965		741.53	0.9805	6.1501	0.6445	0.4160	1.4237
λ^*=2.0154	1.9165	1.8790	1.9941	1.9532	1.7767	1.6769	λ^*=4.7148	4.5156	4.9708	5.3249	5.3286	4.2221	3.9755
	0.3484	0.1672	0.4073	0.0874	0.0861	0.3536		1.8381	0.9435	1.5526	0.4112	0.4139	3.8238
							$m=8$						
α^*=2.5362	9.5029	2.4152	2.6273	2.6066	1.9989	1.7124	α^*=3.3711	10.8964	3.3173	3.6332	3.5818	2.8691	2.6162
	500.73	1.0371	36.9046	1.0394	0.4072	1.3617		636.48	0.9714	6.2601	0.6446	0.4106	1.4025
λ^*=2.0154	2.0217	1.9152	2.0254	1.9844	1.8241	1.7358	λ^*=4.7148	4.6548	5.0307	5.3948	5.3768	4.3806	4.1507
	0.2048	0.1148	0.2515	0.0572	0.0621	0.2659		0.9181	0.8588	1.1277	0.3603	0.5986	108.47
							$m=10$						
α^*=2.5362	7.9866	2.4230	2.6361	2.6155	2.0044	1.7173	α^*=3.3711	11.2986	3.3281	3.6454	3.5939	2.8776	2.6243
	306.07	1.0792	32.0324	1.0582	0.4176	1.3855		703.35	0.9976	6.2215	0.6526	0.4230	1.4659
λ^*=2.0154	2.0122	1.9265	2.0324	1.9921	1.8420	1.7605	λ^*=4.7148	4.7214	5.0040	5.3660	5.3353	4.4178	4.1969
	0.1106	0.0931	0.1943	0.0451	0.0519	0.2274		0.6918	0.6851	1.0880	0.3211	0.2693	2.1100

Note: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m.

ERs of Bayes estimates under the LINEX loss function depend on the values of v. Moreover, the ERs of Bayes estimates for α and λ based on Lindley's approximation and MCMC samples are close to each other for $v > 0$. On the other hand, the ERs of λ are close to each other for the two methods when $v < 0$ for sufficiently large sample sizes. Furthermore, the ERs for all estimators are decreasing while v is getting close to zero for positive values of v. On the contrary, they are increasing while v getting away from zero for negative values of v. Finally, the ERs of α and λ are generally smaller than that of α^* and λ^*.

In Tables 3.1 and 3.2, we observe that the average length of the HPD credible intervals decrease and their cp are comparable with the nominal values as the sample size increases.

In Table 3.4, the ML and Bayes estimates under SE loss functions and their corresponding ERs are presented when the true value of $\alpha = 2.4640$ and λ is known ($\lambda = 1$). Since the exact distribution of the MLE of α is known, a %95 confidence interval is easily constructed. For a comparison, a %95 confidence intervals of the α under both the asymptotic and exact distributions are tabulated. An approximate confidence interval for α is obtained by using the ML estimate of λ. Moreover, a %95 Bayesian credible interval for α is obtained by using the posterior distribution of α are also reported.

The following results are obtained from Table 3.4, the average ERs of α decrease as the sample size increases in all cases, as expected. Moreover, the ERs of Bayes estimates of α under the SE loss function are smaller than that of ML estimates. Also, it is observed that the average length of the intervals decrease as the sample size increases in all cases. Furthermore, the average length of the Bayesian credible intervals are smaller than that of the other intervals. Finally, cp values for confidence intervals are closer to the nominal value for m as large as 15.

To generate a large number of record values takes too much time. For this reason, an approximate confidence interval is reported under two different cases. Firstly, a %95 confidence intervals for α under both the asymptotic and exact distributions are tabulated for $m = 5, 10, 15, 20$ in Table 3.4. Secondly, we only generated a set of 30 records from $GE(2.6236, 1.5007)$ by using gamma prior with parameters $(a_1, b_1) = (5, 2)$ and $(a_2, b_2) = (10, 5)$. The record values and their inter-record times

Table 3.4: Results for the true value of $\alpha = 2.4640$, $(a_1, b_1) = (5, 2)$ and λ is known.

MLE	Bayes SE	Exact MLE C.I.	Bayesian credible interval	Approximate C.I.
		$m = 5$		
2.2313	2.1224	(0.9485,5.9836)	(1.1781,4.1971)	(0,6.4313)
1.5602	0.8830	0.9720	0.9550	0.8990
		$m = 10$		
2.6157	2.3958	(1.2529,4.4639)	(1.3499,3.7770)	(0.0214,5.7319)
1.3170	0.6401	0.9620	0.9580	0.9420
		$m = 15$		
2.5859	2.4313	(1.4453,4.0439)	(1.4939,3.6283)	(0.1871,5.0795)
0.9273	0.5379	0.9560	0.9500	0.9490
		$m = 20$		
2.4476	2.3529	(1.5185,3.6880)	(1.5525,3.4268)	(0.2856,4.6510)
0.7796	0.5199	0.9540	0.9520	0.9510

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. However, for the last three columns, the first row represents a %95 confidence interval and the second row represents their cp's.

Table 3.5: Data is generated from $GE(2.6236, 1.5007)$.

i	1	2	3	4	5	6	7	8	9	10
R_i	0.4440	0.7010	0.7709	0.8751	0.8851	1.1427	1.1933	1.4927	1.6651	1.8378
K_i	3	1	1	1	1	3	1	4	1	9
i	11	12	13	14	15	16	17	18	19	20
R_i	2.2633	2.8443	2.9103	3.1298	3.3530	3.7073	3.7633	3.8257	3.9191	4.0392
K_i	5	3	23	4	244	61	170	9	111	42
i	21	22	23	24	25	26	27	28	29	30
R_i	4.0528	4.3658	4.4601	4.5889	4.6356	4.9092	5.0858	5.2541	5.4127	7.5152
K_i	51	266	382	369	211	77	184	72	21	1

are tabulated in Table 3.5 and then the approximate and HPD credible intervals are reported in Table 3.6. In Tables 3.4 and 3.6 are mainly obtained to illustrate how the approximate confidence intervals perform when the number of record values get larger. From these tables, it can be observed that these confidence intervals and their cp values are comparable.

3.3.4. Conclusions

This section deals with the ML and Bayesian point estimates as well as confidence intervals for the unknown parameters when the underlying distribution

Table 3.6: Confidence intervals for α and λ.

	Approximate confidence interval	HPD credible interval
α	(0.7394,3.7993)	(1.1846,3.6846)
λ	(1.1909,1.6342)	(1.2455,1.6226)

is GE distribution. The ML estimates of the unknown parameters are derived under inverse sampling scheme. The Lindley's approximation and MCMC methods are used to get the Bayes estimates under the SE and the LINEX loss function. Monte Carlo simulation reveals out that the ERs of the Bayes estimates are smaller than that of ML estimates under the SE loss function. However, the ERs for the LINEX loss function depend on the asymmetry parameter v. In particular, the ERs of the estimates are decreasing while v is getting close to zero for positive values of v and are increasing while v getting away from zero for negative values of v. The average length of the HPD credible intervals are smaller than that of the confidence intervals with more reasonable cp. Finally, it is suggested to use record values and their corresponding inter-record times instead of just using record values to decrease the ERs of estimates under the SE and the LINEX loss function.

3.4. Estimation of The Reliability Based on Record Values

The likelihood and Bayesian estimation of the stress-strength reliability based on lower record values from the GE distribution with known scale parameter were considered by [Baklizi, 2008]. Confidence intervals, exact and approximate, as well as the Bayesian credible sets for the stress-strength reliability were also obtained. An interval estimates for the stress-strength reliability using lower record data from the GE distribution with known scale parameter was developed by [Wong and Wu, 2009]. Recently, the ML and Bayesian estimation of the stress-strength reliability based on lower record values from the GE distribution are considered by [Asgharzadeh et al., 2014].

4. STATISTICAL ANALYSIS FOR THE KUMARASWAMY DISTRIBUTION

4.1. Introduction

Kumaraswamy showed that the well known pdf such as the normal, log-normal, beta and empirical distributions such as Johson's and polynomial-transformed-normal, etc., do not fit well hydrological data, such as daily rainfall, daily stream flow, etc. and developed a new probability distribution function known as the sinepower pdf in [Kumaraswamy, 1976], [Kumaraswamy, 1978]. Furthermore, a more general pdf for double bounded random processes, which is known as Kumaraswamy's distribution was developed by [Kumaraswamy, 1980]. This distribution is applicable to many natural phenomena whose outcomes have lower and upper bounds, such as the heights of individuals, scores obtained on a test, atmospheric temperatures, hydrological data, etc. Also, this distribution could be appropriate in situations where scientists use probability distributions which have infinite lower and/or upper bounds to fit data, when in reality the bounds are finite.

If a random variable X follows a Kumaraswamy distribution, denoted by $Kum(a, b)$, then its cdf and pdf are given by

$$F(x; a, b) = 1 - (1 - x^a)^b, \ 0 < x < 1 \tag{4.1}$$

$$f(x; a, b) = abx^{a-1}(1 - x^a)^{b-1}, \ 0 < x < 1 \tag{4.2}$$

where $a, \ b > 0$ are the shape parameters. The mean and the variance of a Kumaraswamy distribution are given by

$$E(X) = bB\left(1 + \frac{1}{a}, b\right), \tag{4.3}$$

and

$$Var(X) = bB\left(1 + \frac{2}{a}, b\right) - \left\{bB\left(1 + \frac{1}{a}, b\right)\right\}^2, \tag{4.4}$$

where $B(x, y)$ is the Beta function and $B(x, y) = \Gamma(x)\Gamma(y)/\Gamma(x + y)$.

The Kumaraswamy distribution has the same basic shape properties as the beta distribution, namely: when $a > 1$ and $b > 1$ it is unimodal; when $a < 1$ and $b < 1$ it is uniantimodal; when $a > 1$ and $b \leq 1$ it is increasing; when $a \leq 1$ and $b > 1$ it is decreasing; when $a = b = 1$ it is constant.

In a discussion paper, [Nadarajah, 2008] has pointed out that many papers in the hydrological literature have used Kumaraswamy's distribution because it is deemed as a better alternative to the beta distribution, see [Koutsoyiannis and Xanthopoulos, 1989]. Over the years, this distribution has received considerable attention by scientists working in hydrology and related areas, see [Sundar and Subbiah, 1989], [Fletcher and Ponnamblam, 1996], [Seifi et al., 2000], [Ponnambalam et al., 2001], [Ganji et al., 2006]. The background and genesis of the Kumaraswamy distribution, and more importantly, made clear some similarities and differences between the beta and Kumaraswamy distributions were explored by [Jones, 2009]. He highlighted that the Kumaraswamy distribution has some advantages over the beta distribution in terms of tractability. For example, its cdf has a closed form, the quantile functions are easily obtainable and one can easily generate random variables from Kumaraswamy distribution. The generalized order statistics of Kumaraswamy distribution was considered by [Garg, 2009]. A modified ML estimators that are bias-free to second order were derived by [Lemonte, 2011] for the Kumaraswamy distribution. New properties of the Kumaraswamy distribution was derived by [Mitnik, 2013].The Bayesian and non-Bayesian estimation for the shape parameter of the Kumaraswamy distribution and the predictive intervals of a future observation under type-II censored sample was obtained by [Sindhu et al., 2013]. The ML and Bayesian estimation for the shape parameters, the reliability and the hazard rate functions of the Kumaraswamy distribution and the prediction for a new observation based on generalized order statistics were derived by [El-Deen et al., 2014]. Moreover, the Kumaraswamy distribution has used excessively to construct a new distributions, see [Corderio and Castro, 2011], [Paranaiba et al., 2013], [Gomes et al., 2014].

The rest of this chapter is organized as follows. In Section 4.1,the statistical inferences for the Kumaraswamy distribution based on record values are considered. In Section 4.2, the statistical inferences for the Kumaraswamy distribution based on record values with their corresponding inter-record times are considered. In Section 4.3, the statistical inferences for the stress-strength reliability of the Kumaraswamy distribution based on record values are considered.

4.2. Estimation of The Parameters Based on Record Values

When the underlying distribution is the exponential model, the statistical properties of record values were studied by [Ahsanullah, 1980], [Awad and Raqab, 2000], they studied four procedures to obtain prediction intervals for the future sth record value and by means of computer simulation they compared these procedures. Three types of estimators, ML, minimum variance unbiased and Bayesian estimators for the one parameter Burr type X distribution based on the record values were derived by [Ali Mousa, 2001]. Bayesian estimation for the exponential, Weibull, Pareto and Burr Type XII distribution based on record values were considered by [Ahmadi and Doostparast, 2006] when both of the parameters are considered as a random variables. Based on the record values from the two-parameter Pareto distribution, ML and Bayes estimators for the unknown parameters and point and interval prediction for the future record values were obtained by [Raqab et al., 2007]. Statistical analysis of record values from the geometric distribution was done by [Doostparast and Ahmadi, 2006]. Furthermore, they derived estimators for the unknown parameter and also considered the problem of predicting the future record values based on past record values from a non-Bayesian and Bayesian point of view. Bayes estimators of the parameter, reliability function and hazard rate for the Rayleigh distribution based on upper record values were obtained by [Hendi et al., 2007]. The prediction of k-records from a general class of distributions under balanced type loss functions was studied by [Ahmadi et al., 2009].

In this section, the estimates for the parameters of the Kumaraswamy distribution and the prediction of the future record values are obtained under the classical and Bayesian frameworks. In the Bayesian case, the shape parameters are assumed to be random variables and estimates of the parameters and for the future sth record value are obtained under the SE and the LINEX loss function. Finally, the findings are illustrated with actual and computer generated data.

4.2.1. ML Estimation

Let $R_1, \ldots, R_m$ be the first m upper record values observed from $Kum(a, b)$. The likelihood function associated with the sequence $\{R_1, ..., R_m\}$ is given by [Arnold et al., 1998] as

$$L(\mathbf{r}) = f(r_m) \prod_{i=1}^{m-1} \frac{f(r_i)}{1 - F(r_i)}, \tag{4.5}$$

where $0 < r_1 < ... < r_m$. From equations (4.1), (4.2) and (4.5), we have

$$L(a, b; \mathbf{r}) = a^m b^m (1 - r_m^a)^b \prod_{i=1}^{m} \frac{r_i^{a-1}}{1 - r_i^a}, \tag{4.6}$$

and so the log-likelihood function is

$$l(a, b; \mathbf{r}) = m(\ln a + \ln b) + b \ln(1 - r_m^a) + (a - 1) \sum_{i=1}^{m} \ln r_i - \sum_{i=1}^{m} \ln(1 - r_i^a). \tag{4.7}$$

Then, the MLEs of a and b are given by

$$\widehat{b} = -\frac{m}{\ln(1 - r_m^{\widehat{a}})}, \tag{4.8}$$

and $\widehat{a}$ is a solution of the following non-linear equation

$$\frac{m}{a} + \frac{m r_m^a \ln r_m}{(1 - r_m^a) \ln(1 - r_m^a)} + \sum_{i=1}^{m} \frac{\ln r_i}{1 - r_i^a} = 0. \tag{4.9}$$

Therefore, $\widehat{a}$ can be obtained as a solution of the non-linear equation of the form $h(a) = a$ where

$$h(a) = -m \left[\frac{m r_m^a \ln r_m}{(1 - r_m^a) \ln(1 - r_m^a)} + \sum_{i=1}^{m} \frac{\ln r_i}{1 - r_i^a} \right]^{-1}. \tag{4.10}$$

Since, $\widehat{a}$ is a fixed point solution of the non-linear equation(4.10), its value can be obtained using an iterative scheme as like: $a_{(j+1)} = h(a_{(j)})$, where $a_{(j)}$ is the jth iterate of $\widehat{a}$. The iteration procedure should be stopped when $\left|a_{(j+1)} - a_{(j)}\right|$ is sufficiently small. After $\widehat{a}$ is obtained, $\widehat{b}$ can be obtained from equation (4.8).

Next, the existence and uniqueness of the ML estimates of the parameters a and b of the Kumaraswamy distribution based on upper record values are proved.

Theorem 4.1: The ML estimates of the parameters a and b are unique and are given by

$$\widehat{b} = -\frac{m}{\ln(1 - r_m^{\widehat{a}})}, \tag{4.11}$$

where $\widehat{a}$ is the solution of the non-linear equation:

$$G(a) = \frac{m}{a} + \frac{m r_m^a \ln r_m}{(1 - r_m^a)\ln(1 - r_m^a)} + \sum_{i=1}^{m} \frac{\ln r_i}{1 - r_i^a} = 0. \tag{4.12}$$

Proof 4.1: $G(a)$ can be rewritten as

$$G(a) = \frac{m}{a}\left[1 + G_1(a) + \frac{G_2(a)}{G_3(a)}\right], \tag{4.13}$$

where $G_1(a) = (1/m)\sum_{i=1}^{m} \ln s_i/(1 - s_i)$, $G_2(a) = s_m \ln s_m/m(1 - s_m)$, $G_3(a) = \ln(1 - s_m)/m$ and $s_i = r_i^a$, $i = 1, ..., m$.

The limit of $G(a)$ is considered as $a \to 0^+$ and $a \to \infty$. It is easily obtained that $\lim_{a\to 0^+} G(a) = \infty$ and $\lim_{a\to\infty} G(a) < 0$. By the intermediate value theorem $G(a)$ has at least one root in $(0, \infty)$. If it can be show that $G^{'}(a) < 0$, then the proof will be completed. Since $r_i < r_m$, $1/(1 - r_i^a) < 1/(1 - r_m^a)$, $i = 1, ..., m - 1$ for $a > 0$ and

$$\begin{aligned}
G^{'}(a) &= \frac{-m}{a^2} + \frac{m r_m^a \left(\ln r_m^a\right)^2}{a^2\left(1 - r_m^a\right)^2}\frac{\left(r_m^a + \ln(1 - r_m^a)\right)}{\left(\ln(1 - r_m^a)\right)^2} + \sum_{i=1}^{m}\frac{r_i^a\left(\ln r_i^a\right)^2}{a^2\left(1 - r_i^a\right)^2} \\
&< \frac{m}{a^2}\left\{-1 + r_m^a\frac{\left(\ln r_m^a\right)^2}{\left(1 - r_m^a\right)^2}\frac{\left(r_m^a + \ln(1 - r_m^a)\right)}{\left(\ln(1 - r_m^a)\right)^2} + \frac{r_m^a\left(\ln r_m^a\right)^2}{\left(1 - r_m^a\right)^2}\right\} \qquad (4.14) \\
&= \frac{-m}{a^2} + \frac{m r_m^a \left(\ln r_m^a\right)^2}{a^2(1 - r_m^a)^2}\left\{1 + \frac{r_m^a}{\left(\ln(1 - r_m^a)\right)^2} + \frac{r_m^a}{\ln(1 - r_m^a)}\right\}.
\end{aligned}$$

According to the order of convergence of the corresponding terms in the last expression, it can be easily show that $G^{'}(a) < 0$.

Finally, we will show that the ML estimates of a and b maximizes the log-likelihood function $l(a, b; \mathbf{r})$. Let $H(a, b)$ be the Hessian matrix of $l(a, b; \mathbf{r})$ at (a, b). It is clear that $H_{11}(\widehat{a}, \widehat{b}) < 0$ and the determinant of the Hessian matrix

$$\begin{aligned} D(\widehat{a}, \widehat{b}) &= H_{11}(\widehat{a}, \widehat{b})H_{22}(\widehat{a}, \widehat{b}) - (H_{12}(\widehat{a}, \widehat{b}))^2 \\ &= G^{'}(\widehat{a}) \left\{ \frac{\left(\ln(1 - r_m^{\widehat{a}})\right)^2}{m} \right\}, \end{aligned} \tag{4.15}$$

and $D(\widehat{a}, \widehat{b}) > 0$. Hence, $(\widehat{a}, \widehat{b})$ is the local maximum point of $l(a, b; \mathbf{r})$. Since there is no singular point of $l(a, b; \mathbf{r})$ and it has a single critical point then, it is enough to show that the absolute maximum of the function is indeed the local maximum. Assume that there exist a $\widehat{a}_0$ in the domain in which $l^*(\widehat{a}_0) > l^*(\widehat{a})$, where $l^*(\widehat{a}) = l(\widehat{a}, \widehat{b}; \mathbf{r},)$. Since $\widehat{a}$ is the local maximum there should be some point a_1 in the neighborhood of $\widehat{a}$ such that $l^*(\widehat{a}) > l^*(a_1)$. Let $k(a) = l^*(a) - l^*(\widehat{a})$ then $k(\widehat{a}_0) > 0$, $k(a_1) < 0$ and $k(\widehat{a}) = 0$. This implies that a_1 is a local minimum of the $l^*(a)$, but $\widehat{a}$ is the only critical point so it is a contradiction. Therefore, $(\widehat{a}, \widehat{b})$ is the absolute maximum of $l(a, b; \mathbf{r})$. ■

4.2.2. Bayesian Estimation

Assume that the parameters a and b have a joint bivariate prior density function that was first suggested by [Al-Hussaini and Jaheen, 1995] as,

$$\pi(a, b) = \pi_1(b\,|a)\pi_2(a), \tag{4.16}$$

where

$$\pi_1(b\,|a) = \frac{a^{\alpha+1}}{\Gamma(\alpha+1)\gamma^{\alpha+1}} b^{\alpha} e^{-ba/\gamma},\ \alpha > -1,\ \gamma > 0, \tag{4.17}$$

is the gamma conjugate prior, was first introduced by [Papadopoulos, 1978] and was also used later on by [Al-Hussaini and Jaheen, 1992], and a has gamma prior with parameters (δ, β),

$$\pi_2(a) = \frac{a^{\delta-1}}{\Gamma(\delta)\beta^{\delta}} e^{-a/\beta}, \ \beta > 0, \ \delta > 0. \tag{4.18}$$

From equations (4.6) and (4.16), the joint posterior density function of a and b is

$$\pi\left(a, b\left|\mathbf{r}\right.\right) = \frac{L(a, b; \mathbf{r})\pi(a, b)}{\int_0^\infty \int_0^\infty L(a, b; \mathbf{r})\pi(a, b) da db}$$

$$= \frac{a^{\delta+m+\alpha} b^{\alpha+m} e^{-a\left(\frac{1}{\beta} - \sum_{i=1}^m \ln r_i\right)} e^{-b\left(\frac{a}{\gamma} - \ln(1-r_m^a)\right)}}{\Gamma(m+\alpha+1)\psi(0,1,0,0)\prod_{i=1}^m (1-r_i^a)}, \tag{4.19}$$

where

$$\psi(c, d, h, f) = \int_0^\infty \frac{t^{\delta+m+\alpha+c} e^{-t\left(h + \frac{1}{\beta} - \sum_{i=1}^m \ln r_i\right)}}{\prod_{i=1}^m (1-r_i^t)\left[f + \frac{t}{\gamma} - \ln(1-r_m^t)\right]^{m+\alpha+d}} dt. \tag{4.20}$$

If the loss function is the SE loss function, then the Bayes estimates of a and b are the given by their marginal posterior expectations as

$$\widehat{a}_{BS} = E(a\left|\mathbf{r}\right.) = \frac{\psi(1,1,0,0)}{\psi(0,1,0,0)}, \tag{4.21}$$

and

$$\widehat{b}_{BS} = E(b\left|\mathbf{r}\right.) = (m+\alpha+1)\frac{\psi(0,2,0,0)}{\psi(0,1,0,0)}, \tag{4.22}$$

respectively. If we use the LINEX loss function, the Bayes estimates of a and b are given by

$$\widehat{a}_{BL} = -\frac{1}{v}\ln E(e^{-av}\left|\mathbf{r}\right.) = -\frac{1}{v}\ln\left(\frac{\psi(0,1,v,0)}{\psi(0,1,0,0)}\right), \tag{4.23}$$

and

$$\widehat{b}_{BL} = -\frac{1}{v}\ln E(e^{-bv}\left|\mathbf{r}\right.) = -\frac{1}{v}\ln\left(\frac{\psi(1,1,0,v)}{\psi(0,1,0,0)}\right), \tag{4.24}$$

respectively. It should be point out that equations (4.21)-(4.24) are not in explicit form, but the practitioner should not be discouraged, there are several numerical methods that can be used to evaluate those expressions.

4.2.3. Prediction of Future Record Values

In this section, the problem of prediction of the sth $(s > m)$ upper record value are considered by using non-Bayesian and Bayesian approaches.

4.2.3.1. Non-Bayesian Prediction Approach

When the first m upper record values are observed from a population with pdf $f(x;\theta)$, the predictive likelihood function of $Y = R_s$, $s > m$ and the parameters θ is given by [Basak and Balakrishnan, 2003] as

$$L(y,\theta;\mathbf{r}) = \prod_{i=1}^{m} \frac{f(r_i;\theta)}{1-F(r_i;\theta)} \frac{[H(y;\theta)-H(r_m;\theta)]^{s-m-1}}{\Gamma(s-m)} f(y;\theta), \tag{4.25}$$

where $\theta = (a,b)$, $\mathbf{r} = (r_1, ..., r_m)$ and $H(y;\theta) = -\ln(1-F(y;\theta))$. From equations (4.1), (4.2) and (4.25), we have

$$L(y,a,b;\mathbf{r}) = \frac{a^{m+1}b^s y^{a-1}}{(1-y^a)^{1-b}} \frac{[\ln(1-r_m^a)-\ln(1-y^a)]^{s-m-1}}{\Gamma(s-m)} \prod_{i=1}^{m} \frac{r_i^{a-1}}{1-r_i^a}, \tag{4.26}$$

where $y > r_m > r_{m-1} > ... > r_1 > 0$. Then, the PMLE of a and b and the MLP of $Y = R_s$ are obtained by minimizing the logarithm of the predictive likelihood function in equation (4.26) with respect to the above mentioned parameters. After some simplifications these equations are

$$\frac{m+1}{a} + (s-m-1)\frac{(y^a \ln y/(1-y^a)) - (r_m^a \ln r_m)/(1-r_m^a)}{\ln(1-r_m^a)-\ln(1-y^a)}$$

$$+\ln y - (b-1)\frac{y^a \ln y}{1-y^a} + \sum_{i=1}^{m} \frac{\ln r_i}{1-r_i^a} = 0, \tag{4.27}$$

$$\frac{s}{b} + \ln(1-y^a) = 0, \tag{4.28}$$

$$(s-m-1)\frac{ay^{a-1}/(1-y^a)}{\ln(1-r_m^a)-\ln(1-y^a)} + \frac{a-1}{y} - (b-1)\frac{ay^{a-1}}{1-y^a} = 0. \tag{4.29}$$

The above system of three equations can be reduced to a system of two equations by replacing $b = -s/(\ln(1-y^a))$ into the equations (4.27) and (4.29) and obtained as follows

$$\frac{m+1}{a} + (s-m-1)\frac{(y^a \ln y/(1-y^a)) - (r_m^a \ln r_m)/(1-r_m^a)}{\ln(1-r_m^a)-\ln(1-y^a)} + \ln y$$
$$+ \left(\frac{s}{\ln(1-y^a)} + 1\right)\frac{y^a \ln y}{1-y^a} + \sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a} = 0, \tag{4.30}$$

$$\frac{(s-m-1)ay^{a-1}/(1-y^a)}{\ln(1-r_m^a)-\ln(1-y^a)} + \frac{a-1}{y} + \left(\frac{s}{\ln(1-y^a)} + 1\right)\frac{ay^{a-1}}{1-y^a} = 0. \tag{4.31}$$

The above non-linear equation system can be easily solved numerically.

4.2.3.2. Bayesian Prediction Approach

The prediction of future records based on a Bayesian approach is considered under the SE and the LINEX loss functions. The conditional density of $Y = R_s$, $s > m$ given the past m records is

$$f(y\,|\mathbf{r}\,,\theta) = \frac{[H(y;\theta) - H(r_m;\theta)]^{s-m-1}}{\Gamma(s-m)}\frac{f(y\,|\theta)}{1-F(r_m\,|\theta)}$$

$$= \frac{ab^{s-m}}{\Gamma(s-m)}\left[\ln\left(\frac{1-r_m^a}{1-y^a}\right)\right]^{s-m-1}\left(\frac{1-y^a}{1-r_m^a}\right)^b\frac{y^{a-1}}{1-y^a}, \tag{4.32}$$

where $r_m < y < 1$. The Bayes predictive density function of Y given $\mathbf{r}$, see [Arnold et al., 1998], is given by

$$h(y\,|\mathbf{r}) = \int_0^\infty \int_0^\infty f(y\,|\mathbf{r}\,, a, b)\pi(a, b\,|\mathbf{r})dadb. \tag{4.33}$$

Using equations (4.19) and (4.32), the Bayesian predictive density function of $Y = R_s$ is obtained as follows

$$h(y\,|\mathbf{r}\,) = \frac{\sum_{i=0}^{s-m-1} \begin{pmatrix} s-m-1 \\ i \end{pmatrix} \xi_i(\mathbf{r}, y)}{B(s-m, m+\alpha+1)\psi(0,1,0,0)y}, \tag{4.34}$$

where

$$\xi_i(\mathbf{r}, u) = \int_0^\infty \frac{t^{k_1}\left[\ln(1-r_m^t)\right]^i \left[-\ln(1-u^t)\right]^{k_2} e^{-t\left(\frac{1}{\beta}-\ln u-\sum_{i=1}^m \ln r_i\right)}}{\prod_{i=1}^m(1-r_i^t)\left[\frac{t}{\gamma}-\ln(1-u^t)\right]^{s+\alpha+1}(1-u^t)} dt, \tag{4.35}$$

$k_1 = \delta+\alpha+m+1$, $k_2 = s-m-1-i$ and $B(x,y)$ is the Beta function. Then, the Bayes point predictor of $Y = R_s$, $s > m$ under the SE and the LINEX loss functions are given by

$$\widehat{Y}_{SEL} = \int_{r_m}^1 yh(y\,|\mathbf{r})dy = \frac{\sum_{i=0}^{s-m-1} \begin{pmatrix} s-m-1 \\ i \end{pmatrix} \int_{r_m}^1 \xi_i(\mathbf{r}, y)dy}{B(s-m, m+\alpha+1)\psi(0,1,0,0)}, \tag{4.36}$$

and $\widehat{Y}_{Linex} = -\left(\ln E(e^{-vY}\,|\mathbf{r})\right)/v$ where

$$E_{Y|\mathbf{r}}(e^{-vY}\,|\mathbf{r}\,) = \int_{r_m}^1 e^{-vY} h(y\,|\mathbf{r}\,)dy$$

$$= \frac{\sum_{i=0}^{s-m-1} \begin{pmatrix} s-m-1 \\ i \end{pmatrix} \int_{r_m}^1 e^{-vy-\ln y}\xi_i(\mathbf{r}, y)dy}{B(s-m, m+\alpha+1)\psi(0,1,0,0)}. \tag{4.37}$$

4.2.4. Simulation Study

The two examples are given to illustrate the findings of Section 4.2. The former is a real data set is obtained from the Shasta reservoir in California while latter example uses simulated data set. In both examples, the mathematical package MATLAB 7.7.0 is used to obtain the estimates of the parameters a and b and the prediction of future record value(s).

The first example deals with the monthly water capacity data from the Shasta reservoir in California, USA and are taken for the month of February from 1991 to 2010 (see http://cdec.water.ca.gov/reservoir_map.html). The maximum capacity of the reservoir is 4552000 AF and the data are transformed to the interval $[0, 1]$. The actual and transformed data are given in Table 4.1. The 20 values are used to verify that the transformed data follow Kumaraswamy's distribution. The Kolmogorov–Smirnov test shows that indeed the observations follow the double bounded Kumaraswamy's distribution (p-value > 0.2). The prediction of the 7th record value based on the first 5 records is computed by using the ML and Bayes prediction approaches.

The parameters of the priors given by equations (4.17) and (4.18) are estimated by using the method of moments. From equations (4.17) and (4.18), we have $a \sim Gamma(\delta, \beta)$ and $b\,|\widehat{a} \sim Gamma(\alpha+1, \gamma/\widehat{a})$ where $\widehat{a} = 2.4446$ is obtained from the data. Using the method of moments, we have $\overline{X} = \delta\beta$ and $\sum_{i=1}^{n} x_i^2/n = \delta\beta^2 + (\delta\beta)^2$ by equating the sample moments with the population moments. β and δ are obtained by solving these two equations and are given by

$$\beta = \frac{\sum_{i=1}^{n} x_i^2/n - \overline{X}^2}{\overline{X}}, \ \delta = \frac{\overline{X}^2}{\sum_{i=1}^{n} x_i^2/n - \overline{X}^2}. \tag{4.38}$$

Then, these two quantities are used as estimates of the population parameters δ and β. Similarly, using the third and fourth moments of $b\,|\widehat{a}$, we have

$$\frac{1}{n}\sum_{i=1}^{n} x_i^3 = \left(\frac{\gamma}{\widehat{a}}\right)^3 (\alpha+1)(\alpha+2)(\alpha+3), \tag{4.39}$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i^4 = \left(\frac{\gamma}{\widehat{a}}\right)^4 (\alpha+1)(\alpha+2)(\alpha+3)(\alpha+4). \tag{4.40}$$

Solving these two equations for γ and α, the following non-linear equations are obtained

$$\left(\frac{\gamma}{\widehat{a}}\right) = \frac{1}{\alpha+4}\frac{\sum_{i=1}^{n} x_i^4/n}{\sum_{i=1}^{n} x_i^3/n}, \ \frac{(\alpha+1)(\alpha+2)(\alpha+3)}{(\alpha+4)^3} = \frac{\left(\sum_{i=1}^{n} x_i^3/n\right)^4}{\left(\sum_{i=1}^{n} x_i^4/n\right)^3}. \tag{4.41}$$

Then again one can use these two quantities as estimates of the population parameters γ and α. For this data set, the estimated parameters are $\alpha = 15.903$, $\beta = 0.0659$,

Table 4.1: Monthly capacity for February and proportion of total capacity.

Year	Capacity	Proportion of total capacity	Year	Capacity	Proportion of total capacity
1991	1542838	0.338936	2001	3495969	0.768007
1992	1966077	0.431915	2002	3839544	0.843485
1993	3459209	0.759932	2003	3584283	0.787408
1994	3298496	0.724626	2004	3868600	0.849868
1995	3448519	0.757583	2005	3168056	0.695970
1996	3694201	0.811556	2006	3834224	0.842316
1997	3574861	0.785339	2007	3772193	0.828689
1998	3567220	0.783660	2008	2641041	0.580194
1999	3712733	0.815627	2009	1960458	0.430681
2000	3857423	0.847413	2010	3380147	0.742563

Table 4.2: Estimates of a and b and predictors of R_7.

			Bayes	
Parameter	MLE	PMLE	SEL	LINEX($v = 0.2$)
a	1.7846	3.1127	3.1877	3.0523
b	5.2496	2.7521	2.0343	1.9752
R_7	-	0.8425	0.9768	0.9771

$\gamma = 0.0970$ and $\delta = 9.5779$. When the actual 7th record value is 0.849868, the ML prediction for the 7th record value is 0.6663 and the Bayesian point predictor under the SE loss function is 0.8518.

Several authors such as [Raqab, 2002], [Ahmadi and Doostparast, 2006], [Hendi et al., 2007] have used simulated data to illustrate their findings. In our second example, a simulated data is used for demonstration purposes. $a = 5.8191$ and $b = 1.0515$ are generated from the priors by using the values of $\alpha = 1$, $\beta = 3$, $\gamma = 2$ and $\delta = 3$. A random sample of 7 record values from the Kumaraswamy distribution are generated, which are $0.5454, 0.6417, 0.8723, 0.9242, 0.9446, 0.9549, 0.9571$. The first 5 will be used to estimate the parameters a and b, and also to predict the 7th record value. Based on this sample, the ML, PML and Bayes estimates of the parameters (a, b), and MLP and Bayes point predictor of $Y = R_7$ are obtained under the SE and the LINEX loss functions and results are listed in Table 4.2.

4.2.5. Conclusion

In this section, the ML and Bayes estimates of two shape parameters of the Kumaraswamy distribution are obtained. The Bayes estimates are derived under the SE and LINEX loss function for bivariate prior density function. Non-Bayesian and Bayesian point predictors of the future record values are obtained. The real life and computer generated data sets are considered to illustrate the use of the methods proposed in this section. The examples reveals out that the performance of the Bayes estimates of the parameters and Bayesian point predictor of the future record values are better than that of ML case.

4.3. Estimation of The Parameters Based on Records and Inter-Record Times

In this section, the parameter estimates of Kumaraswamy distribution based on lower record values and their corresponding inter-record times are obtained under the classical and Bayesian frameworks. The Lindley approximation and MCMC technique are used to obtain the Bayes estimates under different loss functions. Finally, a Monte Carlo simulation is performed to compare the estimates of the parameters. The non-Bayesian and Bayesian point predictors and the Bayesian prediction interval for future lower record values are obtained based on the observed lower record values with their corresponding inter-record times. To see the effect of the inter-record times in parameter estimates, the estimates based on lower record values with inter-record times and upper records which are obtained from the same random sample of the Kumaraswamy distribution are constructed. Finally, the two approach are compared by using Monte Carlo simulations.

4.3.1. ML Estimation

Let $X_1, X_2, \ldots$ be i.i.d. random variables, coming from a population with the cdf and the pdf $F(.)$ and $f(.)$, respectively. Then the likelihood function associated with the sequence $\{R_1, K_1, \ldots, R_m, K_m\}$ is given by [Samaniego and Whitaker, 1986] as

$$L(\mathbf{r},\mathbf{k})=\prod_{i=1}^{m} f(r_i)\left\{1-F(r_i)\right\}^{k_i-1} I_{(-\infty,r_{i-1})}(r_i), \tag{4.42}$$

where $r_0 \equiv \infty$, $k_m \equiv 1$ and $I_A(x)$ is the indicator function of the set A. From equations (4.1), (4.2) and (4.42), we have

$$L(a,b;\mathbf{r},\mathbf{k})=a^m b^m \exp\left\{(a-1)\sum_{i=1}^{m}\ln r_i+\sum_{i=1}^{m}(bk_i-1)\ln(1-r_i^a)\right\}, \tag{4.43}$$

where $r_1 > \ldots > r_m$ and so the log-likelihood function is

$$l(a,b;\mathbf{r},\mathbf{k})=m\ln a+m\ln b+(a-1)\sum_{i=1}^{m}\ln r_i+\sum_{i=1}^{m}(bk_i-1)\ln(1-r_i^a). \tag{4.44}$$

The MLEs of a and b are given by

$$\widehat{b}=\frac{m}{T_a}, \tag{4.45}$$

where $T_a = -\sum_{i=1}^{m} k_i \ln(1-r_i^a)$ and $\widehat{a}$ is the solution of the following non-linear equation

$$\frac{m}{a}+\sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a}-\frac{m}{T_a}\sum_{i=1}^{m}\frac{k_i r_i^a \ln r_i}{1-r_i^a}=0. \tag{4.46}$$

Therefore, $\widehat{a}$ can be obtained as the solution of the non-linear equation of the form $h(a)=a$ where

$$h(a)=m\left[\frac{m}{T_a}\sum_{i=1}^{m}\frac{k_i r_i^a \ln r_i}{1-r_i^\alpha}-\sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a}\right]^{-1}. \tag{4.47}$$

Since, $\widehat{a}$ is a fixed point solution of the non-linear equation (4.47), its value can be obtained using an iterative scheme as like: $a_{(j+1)} = h(a_{(j)})$, where $a_{(j)}$ is the jth iterate of $\widehat{a}$. After $\widehat{a}$ is obtained, $\widehat{b}$ can be obtained from equation (4.45). The iteration procedure should stopped when $\left|a_{(j)}-a_{(j+1)}\right|$ is sufficiently small.

Next, the existence and uniqueness of the ML estimates of the parameters a and b of the Kumaraswamy distribution based on lower record values and their corresponding inter-record times are proved.

Theorem 4.2: The MLEs of the parameters a and b are unique and given by $\widehat{b} = m/T_{\widehat{a}}$ where $\widehat{a}$ is the solution of the non-linear equation:

$$G(a) = \frac{m}{a} - \frac{m}{T_a}\sum_{i=1}^{m}\frac{k_i r_i^a \ln r_i}{1-r_i^a} + \sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a} = 0. \tag{4.48}$$

Proof 4.2: It is clear that

$$\begin{aligned} G(a) &= \frac{m}{a} + \frac{m\sum_{i=1}^{m}\left(k_i r_i^a \ln r_i/\left(1-r_i^a\right)\right)}{\sum_{i=1}^{m} k_i \ln(1-r_i^a)} + \sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a} \\ &> \frac{m}{a} + \frac{m r_1^a \sum_{i=1}^{m} k_i \ln r_i}{\sum_{i=1}^{m} k_i \ln(1-r_i^a)} + \sum_{i=1}^{m}\ln r_i. \end{aligned} \tag{4.49}$$

Then, we have

$$G(0) \equiv \lim_{a\to 0} G(a) > \lim_{a\to 0}\left(\frac{m}{a} + \frac{\sum_{i=1}^{m} k_i \ln r_i}{\sum_{i=1}^{m} k_i \ln(1-r_i^a)}\right) + \sum_{i=1}^{m}\ln r_i = \infty, \tag{4.50}$$

and $G(0) = \infty$. Moreover,

$$\begin{aligned} G(\infty) \equiv \lim_{a\to\infty} G(a) &= \lim_{a\to\infty}\left(\frac{m}{a} + \frac{m\sum_{i=1}^{m}\left(k_i r_i^a \ln r_i/\left(1-r_i^a\right)\right)}{\sum_{i=1}^{m} k_i \ln(1-r_i^a)} + \sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a}\right) \\ &= \lim_{a\to\infty}\frac{m\sum_{i=1}^{m}\left(k_i \ln r_i/\left(1-r_i^a\right)\right)(r_i/r_1)^a}{\sum_{i=1}^{m} k_i(\ln(1-r_i^a)/r_1^a)} + \sum_{i=1}^{m}\ln r_i \\ &= \sum_{i=1}^{m}(\ln r_i - \ln r_1) < 0. \end{aligned} \tag{4.51}$$

Hence, we obtain that $\lim_{a\to 0} G(a) = \infty$ and $\lim_{a\to\infty} G(a) < 0$. By the intermediate value theorem $G(a)$ has at least one root in $(0,\infty)$. If it can be shown that $\partial G(a)/\partial a <$

0 then the proof will be completed. It is easily obtained that

$$\frac{\partial G(a)}{\partial a} = -\frac{1}{a^2}\left[G_1(a) - \frac{mG_2(a)}{\left(\sum_{i=1}^{m} k_i \ln(1-r_i^a)\right)^2}\right], \tag{4.52}$$

where

$$G_1(a) = m - \sum_{i=1}^{m} r_i^a \left(\frac{\ln r_i^a}{1-r_i^a}\right)^2, \tag{4.53}$$

$$G_2(a) = \sum_{i=1}^{m} k_i r_i^a \left(\frac{\ln r_i^a}{1-r_i^a}\right)^2 \left(\sum_{i=1}^{m} k_i \ln(1-r_i^a)\right) + \left(\sum_{i=1}^{m} \frac{k_i r_i^a \ln r_i^a}{1-r_i^a}\right)^2. \tag{4.54}$$

It is easily obtained that $G_1(a) > 0$*, because* $f(x) = x(\ln x)^2/(1-x)^2$, $x \in (0,1)$ *is an increasing function on* $(0,1)$ *and* $\lim_{a\to 0} f(x) = 0$, $\lim_{a\to 1} f(x) = 1$. *Therefore,* $f(x) < 1$ *for* $x \in (0,1)$. *Moreover,* $G_2(a) < 0$ *is obtained by using the Cauchy-Schwarz inequality and* $x < -\ln(1-x)$, $x \in (0,1)$. *Notice that* $g(x) = x + \ln(1-x)$, $x \in (0,1)$ *then* $g(x)$ *is a decreasing function on* $(0,1)$ *and* $g(x) < 0$ *for* $x \in (0,1)$. *Since* $G_1(a) > 0$ *and* $G_2(a) < 0$*, we have* $\partial G(a)/\partial a < 0$ *from equation (4.52).*

Finally, we will show that the MLEs of (a,b) *maximizes the log-likelihood function* $l(a,b;\mathbf{r},\mathbf{k})$. *Let* $H(a,b)$ *be the Hessian matrix of* $l(a,b;\mathbf{r},\mathbf{k})$ *at* (a,b). *It is clear that*

$$H_{11}(\widehat{a},\widehat{b}) = -G_1(\widehat{a}) - \widehat{b}\sum_{i=1}^{m} k_i r_i^{\widehat{a}} \left(\frac{\ln r_i}{1-r_i^{\widehat{a}}}\right)^2 < 0, \tag{4.55}$$

and the determinant of the Hessian matrix

$$D(\widehat{a},\widehat{b}) = H_{11}(\widehat{a},\widehat{b})H_{22}(\widehat{a},\widehat{b}) - \left(H_{12}(\widehat{a},\widehat{b})\right)^2$$

$$= \frac{mG_1(\widehat{a})}{\widehat{a}^2\widehat{b}^2} - G_2(\widehat{a}) > 0. \tag{4.56}$$

Hence, $(\widehat{a},\widehat{b})$ *is the local maximum of* $l(a,b;\mathbf{r},\mathbf{k})$. *Since there is no singular point of* $l(a,b;\mathbf{r},\mathbf{k})$ *and it has a single critical point then, it is enough to show that the* *absolute maximum of the function is indeed the local maximum. Assume that there exist a $\widehat{a}_0$ in the domain in which $l^*(\widehat{a}_0) > l^*(\widehat{a})$, where $l^*(\widehat{a}) = l(\widehat{a}, \widehat{b}; \mathbf{r}, \mathbf{k})$. Since $\widehat{a}$ is the local maximum there should be some point a_1 in the neighborhood of $\widehat{a}$ such that $l^*(\widehat{a}) > l^*(a_1)$. Let $k(a) = l^*(a) - l^*(\widehat{a})$ then $k(\widehat{a}_0) > 0$, $k(a_1) < 0$ and $k(\widehat{a}) = 0$. This implies that a_1 is a local minimum of the $l^*(a)$, but $\widehat{a}$ is the only critical point so it is a contradiction. Therefore, $(\widehat{a}, \widehat{b})$ is the absolute maximum of $l(a, b; \mathbf{r}, \mathbf{k})$.* ■

4.3.2. ML Estimation When a Is Known

Without loss of generality, we assume that $a = a_0$. Then, by (4.43)

$$L(\alpha_0, \beta; \mathbf{r}, \mathbf{k}) = a_0^m b^m \exp\left\{(a_0 - 1)\sum_{i=1}^{m} \ln r_i + \sum_{i=1}^{m}(bk_i - 1)\ln(1 - r_i^{a_0})\right\}, \quad (4.57)$$

where $r_1 > \ldots > r_m$. In this case, $T_{a_0} = -\sum_{i=1}^{m} k_i \ln(1 - r_i^{a_0})$ is a sufficient statistic for b and the MLE of b is $\widehat{b}_{ML} = m/T_{a_0}$. The moment generating function of T_{a_0} is $M(t) = 1/(1 - t/b)^m, b > t$. By the uniqueness of the moment generating function, T_{a_0} is distributed as $Gamma(m, 1/b)$ and its mean and variance are m/b and m/b^2, respectively. Therefore, $E(\widehat{b}_{ML}) = (m/(m-1))b$ and an unbiased estimator for b is $\widehat{b}_U = (m-1)/T_{a_0}$. Notice that, $MSE(\widehat{b}_U) < MSE(\widehat{b}_{ML})$ and $MSE(\widehat{b}_{ML}) \to 0$ as $m \to \infty$ then $\widehat{b}_{ML}$ and $\widehat{b}_U$ converge to b in mean square. Moreover, we have $2bT_{a_0} \sim \chi^2_{2m}$ and the exact $100(1-\eta)\%$ confidence interval of b is

$$\left(\frac{\chi^2_{2m,\eta/2}}{2T_{a_0}}, \frac{\chi^2_{2m,1-\eta/2}}{2T_{a_0}}\right). \quad (4.58)$$

4.3.3. Asymptotic Confidence Interval

In practice, the observed information matrix is used as a consistent estimator of the Fisher information matrix. An asymptotic confidence intervals for the parameters a and b based on lower record values and their corresponding inter-record times are obtained by using the observed information matrix. The observed information matrix $\mathbf{J}_m(a, b)$ is given by

$$\mathbf{J}_m(a,b) = -\begin{bmatrix} \frac{\partial^2 l}{\partial a^2} & \frac{\partial^2 l}{\partial a \partial b} \\ \frac{\partial^2 l}{\partial b \partial a} & \frac{\partial^2 l}{\partial b^2} \end{bmatrix} = \begin{bmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{bmatrix}, \tag{4.59}$$

where

$$J_{11} = \frac{m}{a^2} + \sum_{i=1}^{m}(bk_i - 1)r_i^a \left(\frac{\ln r_i}{1-r_i^a}\right)^2, J_{12} = J_{21} = \sum_{i=1}^{m} \frac{k_i r_i^a \ln r_i}{1-r_i^\alpha}, J_{22} = \frac{m}{b^2}. \tag{4.60}$$

By the asymptotic normality of the MLE, we have $\left[\sqrt{m}(\widehat{a}_{ML} - a), \sqrt{m}(\widehat{b}_{ML} - b)\right] \overset{a}{\sim} N_2(\mathbf{0}, \mathbf{I}^{-1})$ for large m, where $\overset{a}{\sim}$ means approximately distributed and $\mathbf{I}^{-1}$ is the inverse of the Fisher information matrix. If the likelihood equations have a unique solution $\widehat{\theta}_n$, then $\widehat{\theta}_n$ is consistent, asymptotically normal and efficient, see [Lehmann and Casella, 1998]. Since our likelihood equations have a unique solution, these results are satisfied for our estimates. The observed information matrix $\mathbf{J}_m(\widehat{a}_{ML}, \widehat{b}_{ML})/m$ is a consistent estimator for $\mathbf{I}_m(a,b)/m$ under the regularity conditions, see [Lawless, 2003]. Therefore, we use the observed information matrix in the asymptotic normality of the MLE. For large m (the number of record values) under inverse sampling, we can construct the approximate $100(1-\eta)\%$ equi-tail confidence intervals for a and b as

$$a \in \left(\widehat{a}_{ML} \pm c\sqrt{\frac{J_{22}}{J_{11}J_{22} - J_{12}^2}}\right) \text{ and } b \in \left(\widehat{b}_{ML} \pm c\sqrt{\frac{J_{11}}{J_{11}J_{22} - J_{12}^2}}\right), \tag{4.61}$$

where $c = z_{1-\eta/2}$ and z_η is the upper ηth quantile of the standard normal distribution.

4.3.4. Bayes Estimation

In this section, we consider the Bayes estimates of the Kumaraswamy distribution parameters under different loss functions for the inverse sampling scheme.

4.3.4.1. Bayesian Estimation When a Is Known

When the parameter a is assumed to be known, say $a = a_0$, we use the gamma conjugate prior density for the parameter b, was used in section 4.2, i.e.

$$\pi(b) = \frac{a_0^{\alpha+1}}{\Gamma(\alpha+1)\gamma^{\alpha+1}} b^{\alpha} e^{-ba_0/\gamma},\ b > 0\ (\alpha > -1, \gamma > 0). \tag{4.62}$$

Then, the posterior density function of b is readily obtained from equations (4.57) and (4.62) as

$$b|\left(\mathbf{r}, \mathbf{k}\right) \sim Gamma\left(m + \alpha + 1, (a_0/\gamma + T_{a_0})^{-1}\right). \tag{4.63}$$

The Bayes estimate of a under the SE loss function, $\widehat{b}_{BS,1}$, is the mean of the $b|\left(\mathbf{r}, \mathbf{k}\right)$. Therefore

$$\widehat{b}_{BS,1} = \frac{m + \alpha + 1}{a_0/\gamma + T_{a_0}}, \tag{4.64}$$

and the Bayes estimate of b under the LINEX loss function, $\widehat{b}_{BL,1}$, is

$$\widehat{b}_{BL,1} = -\frac{1}{v} \ln E_{b|(\mathbf{r},\mathbf{k})}(e^{-vb}) = \frac{m + \alpha + 1}{v} \ln\left(1 + \frac{v}{a_0/\gamma + T_{a_0}}\right). \tag{4.65}$$

If we use the Jeffrey's non-informative prior, that is $\pi(b) = 1/b$, then we have $b|\left(r, k\right) \sim Gamma\left(m, 1/T_{a_0}\right)$. Therefore, the Bayes estimates of b under the SE and the LINEX loss functions are obtained as

$$\widehat{b}_{BS,0} = \frac{m}{T_{a_0}} \text{ and } \widehat{b}_{BL,0} = \frac{m}{v} \ln\left(1 + \frac{v}{T_{a_0}}\right), \tag{4.66}$$

respectively.

*T*he $100(1-\eta)\%$ Bayesian credible interval can be easily constructed by using the posterior density function of b. It is clear that $2(a_0/\gamma + T_{a_0})b|\left(\mathbf{r}, \mathbf{k}\right) \sim \chi^2_{2(m+\alpha+1)}$. Therefore, a Bayesian credible interval for b is given by

$$\left(\frac{\chi^2_{2(m+\alpha+1)}(\eta/2)}{2(a_0/\gamma + T_{a_0})}, \frac{\chi^2_{2(m+\alpha+1)}(1-\eta/2)}{2(a_0/\gamma + T_{a_0})}\right). \tag{4.67}$$

In the following proposition the comparison of Bayes estimates are given under the SE and the LINEX loss functions.

Proposition 4.1:

i) $\widehat{b}_{BL,1} \leq \widehat{b}_{BS,1}$ *for* $v > 0$.

ii) $\widehat{b}_{BL,1} \geq \widehat{b}_{BS,1}$ *for* $-(a_0/\gamma + T_{a_0}) < v < 0$.

Proof 4.1It is known that

$$\ln(1+x) \leq x \text{ for every } x > -1. \tag{4.68}$$

i) Suppose $v > 0$. $v/(a_0/\gamma + T_{a_0}) > 0$, *when* $a_0/\gamma > 0$ *and* $T_{a_0} > 0$. *We have* $\ln(1 + v/(a_0/\gamma + T_{a_0})) \leq v/(a_0/\gamma + T_{a_0})$ *by the inequality (4.68). Therefore,* $\widehat{b}_{BL,1} \leq \widehat{b}_{BS,1}$

ii) Suppose $v < 0$ *and* $-(a_0/\gamma + T_{a_0}) < v$, *then* $v/(a_0/\gamma + T_{a_0}) > -1$. *We have* $\ln(1 + v/(a_0/\gamma + T_{a_0})) \leq v/(a_0/\gamma + T_{a_0})$ *by the inequality (4.68). Therefore,* $\widehat{b}_{BL,1} \geq \widehat{b}_{BS,1}$. ■

4.3.4.2. Bayesian Estimation When a and b Are Unknown

We consider the Bayes estimates of a and b when the parameters a and b are both unknown and random variables. We assume that a and b have a joint bivariate density function, say $\pi(a, b)$, which is used in section 4.2

$$\pi(a, b) = \pi_1(b\,|a)\pi_2(a), \tag{4.69}$$

where

$$\pi_1(b\,|a) = \frac{a^{\alpha+1}}{\Gamma(\alpha+1)\gamma^{\alpha+1}} b^{\alpha} e^{-ba/\gamma},\ b > 0\ (\alpha > -1, \gamma > 0), \tag{4.70}$$

and a has gamma distribution with parameters (δ, β). From equations (4.43) and (4.69), the joint posterior density function of a and b can be rewritten as

$$\pi(a, b|\,\mathbf{r}, \mathbf{k}) = I(\mathbf{r}, \mathbf{k}) a^{m+\alpha+\delta} b^{m+\alpha}$$

$$\exp\left\{-b\left(\frac{a}{\gamma} + T_a\right) - a\left(\frac{1}{\beta} - \sum_{i=1}^{m} \ln r_i\right) - \sum_{i=1}^{m} \ln(1 - r_i^a)\right\}, \tag{4.71}$$

where

$$\frac{[I(\mathbf{r},\mathbf{k})]^{-1}}{\Gamma(m+\alpha+1)} =_0^\infty \alpha^{m+\alpha+\delta}\left(\frac{a}{\gamma}+T_a\right)^{-m-\alpha-1}$$

$$\exp\left\{-a\left(\frac{1}{\beta}-\sum_{i=1}^{m}\ln r_i\right)-\sum_{i=1}^{m}\ln(1-r_i^a)\right\}da. \quad (4.72)$$

The Bayes estimate of a given measurable function of a and b, say $g(a,b)$, under the SE loss function is

$$\widehat{g}_{BS} = E_{a,b|\mathbf{r},\mathbf{k}}(g(a,b)) = \frac{\int_0^\infty\int_0^\infty g(a,b)L(a,b;\mathbf{r},\mathbf{k})\pi(a,b)dadb}{\int_0^\infty\int_0^\infty L(a,b;\mathbf{r},\mathbf{k})\pi(a,b)dadb}. \quad (4.73)$$

It is not possible to compute equation (4.73) analytically. Two approaches are suggested here to approximate equation (4.73), namely Lindley's approximation and MCMC method.

- Lindley's approximation

  For the two parameter case (a,b), we have from equation (4.71)

$$Q = \ln I(\mathbf{r},\mathbf{k}) + (m+\alpha+\delta)\ln a + (m+\alpha)\ln b - b\left(\frac{a}{\gamma}+T_a\right)$$

$$-\,a\left(\frac{1}{\beta}-\sum_{i=1}^{m}\ln r_i\right)-\sum_{i=1}^{m}\ln(1-r_i^a). \quad (4.74)$$

The joint posterior mode is the obtained from the equations $\partial Q/\partial a = 0$ and $\partial Q/\partial b = 0$ as

$$\widetilde{b} = \frac{m+\alpha}{\widetilde{a}/\gamma + T_{\widetilde{a}}}, \quad (4.75)$$

and $\widetilde{a}$ is the solution of the nonlinear equation

$$\frac{m+\alpha+\delta}{a} - \frac{m+\alpha}{a/\gamma+T_a}\left(\frac{1}{\gamma}+\sum_{i=1}^{m}\frac{k_i r_i^a\ln r_i}{1-r_i^a}\right) - \frac{1}{\beta} + \sum_{i=1}^{m}\frac{\ln r_i}{1-r_i^a} = 0. \quad (4.76)$$

It can be solved by using the same procedure in equations (4.45) and (4.47). The elements of the Q^* are

$$Q_{11}^* = \frac{m+\alpha+\delta}{a^2} + \sum_{i=1}^{m}(bk_i - 1)r_i^a \left(\frac{\ln r_i}{1-r_i^a}\right)^2, \tag{4.77}$$

$$Q_{12}^* = Q_{21}^* = \frac{1}{\gamma} + \sum_{i=1}^{m} \frac{k_i r_i^a \ln r_i}{1-r_i^a},\ Q_{22}^* = \frac{m+\alpha}{b^2}, \tag{4.78}$$

and $\tau_{ij},\ i,j = 1,2$ are obtained by using equations (4.77) and (4.78). Moreover, we have

$$Q_{12} = 0,\ Q_{21} = -\sum_{i=1}^{m} k_i r_i^a \left(\frac{\ln r_i}{1-r_i^a}\right)^2,\ Q_{03} = \frac{2\,(m+\alpha)}{b^3}, \tag{4.79}$$

$$Q_{30} = \frac{2\,(m+\alpha+\delta)}{a^3} - \sum_{i=1}^{m}(bk_i - 1)r_i^a(1+r_i^a)\left(\frac{\ln r_i}{1-r_i^a}\right)^3. \tag{4.80}$$

Therefore, the approximate Bayes estimates of a and b under the SE and the LINEX loss functions are

$$\widehat{a}_{BS,Lind} = \widetilde{a} + \frac{1}{2}\left[Q_{30}\tau_{11}^2 + 3Q_{21}\tau_{11}\tau_{12} + Q_{03}\tau_{21}\tau_{22}\right], \tag{4.81}$$

$$\widehat{a}_{BL,Lind} = \widetilde{a} - \frac{1}{v}\ln\left[1 + \frac{v}{2}\left(v\tau_{11} - Q_{30}\tau_{11}^2 - 3Q_{21}\tau_{11}\tau_{12} - Q_{03}\tau_{21}\tau_{22}\right)\right], \tag{4.82}$$

$$\widehat{b}_{BS,Lind} = \widetilde{b} + \frac{1}{2}\left[Q_{30}\tau_{12}\tau_{11} + Q_{21}(\tau_{11}\tau_{22} + 2\tau_{12}^2) + Q_{03}\tau_{22}^2\right], \tag{4.83}$$

$$\widehat{b}_{BL,Lind} = \widetilde{b} - \frac{1}{v}\ln\left[1 + \frac{v^2}{2}\tau_{22} - \frac{v}{2}Q_{21}(\tau_{11}\tau_{22} + 2\tau_{12}^2) - \frac{v}{2}Q_{30}\tau_{12}\tau_{11} - \frac{v}{2}Q_{03}\tau_{22}^2\right]. \tag{4.84}$$

Notice that all approximate Bayes estimates are evaluated at $(\widetilde{a}, \widetilde{b})$.

- MCMC method

In the previous subsection, the Bayes estimates of a and b are obtained under the SE and the LINEX loss functions by using the Lindley's approximation. Since the exact probability distributions of these estimates are not known, it is difficult to evaluate HPD credible intervals of parameters. For this reason, the MCMC method are used to compute the Bayes estimates of a and b under the SE and the LINEX loss functions as well the HPD credible intervals.

The MCMC method are considered to generate samples from the posterior distributions and then compute the Bayes estimates of a and b under the SE and the LINEX loss functions. The joint posterior density of a and b is given by equation (4.71). It is easy to see that

$$b|\, a, \mathbf{r}, \mathbf{k} \sim Gamma\left(m + \alpha + 1, a/\gamma + T_a\right) \tag{4.85}$$

and

$$\pi(\, a|\, b, \mathbf{r}, \mathbf{k}) \propto a^{m+\alpha+\delta} \exp\left\{ -b\left(\frac{a}{\gamma} + T_a\right) \right.$$
$$\left. -a\left(\frac{1}{\beta} - \sum_{i=1}^{m} \ln r_i\right) - \sum_{i=1}^{m} \ln(1 - r_i^a) \right\}. \tag{4.86}$$

Therefore, samples of b can be generated by using the gamma distribution. However, the posterior distribution of a cannot be reduced analytically to well known distribution and therefore it is not possible to sample directly by standard methods. If the posterior density of a is unimodal and roughly symmetric then it is often convenient to approximate it by a normal distribution centered at the mode (see, [Gelman et al., 2003]). Since the posterior density of a is log-concave density (so unimodal) and the posterior density of a is roughly symmetric with respect to mode (by experimentation), we use the Metropolis-Hasting algorithm with the normal proposal distribution to generate a random sample from the posterior density of a. The hybrid Metropolis-Hastings and Gibbs sampling algorithm, which will be used to solve our problem, is suggested by [Tierney, 1994]. This algorithm combines the Metropolis-Hastings with Gibbs sampling scheme under the Gaussian proposal distribution.

- Step 1: Take some initial guess of a and b, say $a^{(0)}$ and $b^{(0)}$.

- Step 2: Set $t = 1$.
- Step 3: Generate $a^{(t)}$ from $\pi(a|\, b, \mathbf{r}, \mathbf{k})$ using the Metropolis-Hastings algorithm with the proposed value distribution $q(a) \equiv N(\widetilde{a}, V_{\widetilde{a}})$ where $\widetilde{a}$ is a mode of $\pi(a|\, b^{(t-1)}, \mathbf{r}, \mathbf{k})$ and $V_a = \left(-d^2(\ln \pi(a|\, b^{(t-1)}, \mathbf{r}, \mathbf{k}))/da^2\right)^{-1}$:

  -Step 3.1: Let $v = a^{(t-1)}$.

  -Step 3.2: Generate w from the proposal distribution q.

  -Step 3.3: Let $p(v, w) = \min \left\{ 1, \dfrac{\pi(w|\, b^{(t-1)}, \mathbf{r}, \mathbf{k})\; q(v)}{\pi(v|\, b^{(t-1)}, \mathbf{r}, \mathbf{k})\; q(w)} \right\}$.

  -Step 3.4: Generate u from $Uniform(0, 1)$. If $u \leq p(v, w)$ then accept the proposal and set $a^{(t)} = w$; otherwise, set $a^{(t)} = v$.
- Step 4: Generate $b^{(t)}$ from $Gamma\left(m + \alpha + 1, a/\gamma - \sum_{i=1}^{m} k_i \ln(1 - r_i^{a(t)})\right)$.
- Step 5: Set $t = t + 1$.
- Step 6: Repeat Steps 3-5, N times, and obtain the posterior samples $(a^{(i)}, b^{(i)})$, $i = 1, \ldots, N$.

The samples obtained from the algorithm are used to compute the Bayes estimates and to construct the HPD credible intervals. The Bayes estimator of $g \equiv g(a, b)$ based on the SE and the LINEX loss function are given, respectively, by

$$\widehat{g}_{BS,MH} = E(g|\, \mathbf{r}, \mathbf{k}) = \frac{1}{N - M} \sum_{i=M+1}^{N-M} g(a^{(i)}, b^{(i)}), \tag{4.87}$$

and

$$\widehat{g}_{BL,MH} = -\frac{1}{v} \ln \left[\frac{1}{N - M} \sum_{i=M+1}^{N-M} \exp\left(-v\, g(a^{(i)}, b^{(i)})\right) \right], \tag{4.88}$$

where M is the burn-in period.

The HPD $100(1 - \gamma)\%$ credible intervals of a and b can be obtained by the method of [Chen and Shao, 1999]. In particular for a:

From MCMC, the sequence $a_1, \ldots, a_N$, are obtained, and ordered as $a_{(1)} < \ldots < a_{(N)}$. The credible intervals are constructed as $\left(a_{(j)}, a_{(j+[N(1-\gamma)])}\right)$ for $j = 1, ..., N - [N(1 - \gamma)]$ where $[x]$ denotes the largest integer less than or equal to x. Then, the HPD credible interval of a is that interval which has the shortest length. Similarly, the HPD credible interval of b can also be constructed.

4.3.5. Prediction of Future Record Values

In this section, we consider the problem of prediction and prediction interval for the s th ($s > m$) record value using non-Bayesian and Bayesian approaches.

4.3.5.1. Non-Bayesian Prediction

When the first m lower record values are observed, the predictive likelihood function of $Y = R_s$, $s > m$ and the parameters θ is given by [Basak and Balakrishnan, 2003] as

$$L(y, \theta; \mathbf{r}) =_{i=1}^{m} \frac{f(r_i; \theta)}{F(r_i; \theta)} \frac{[H(y; \theta) - H(r_m; \theta)]^{s-m-1}}{\Gamma(s-m)} f(y; \theta), \tag{4.89}$$

where $\theta = (\alpha, \beta)$, $r = (r_1, .., r_m)$ and $H(y; \theta) = -\ln F(y; \theta)$. Moreover, the likelihood function associated with the sequence $\{R_1, K_1, \ldots, R_m, K_m\}$ is given by [Samaniego and Whitaker, 1986] in equation (4.42). Similarly, the predictive likelihood function for the future record R_s based on the sequence $\{R_1, K_1, \ldots, R_m, K_m, R_s\}$ is derived in section 2.3 as

$$L(y, \theta; \mathbf{r}, \mathbf{k}) =_{i=1}^{m} f(r_i; \theta) \{1 - F(r_i; \theta)\}^{k_i - 1} I_{(-\infty, r_{i-1})}(r_i)$$
$$\frac{[H(y; \theta) - H(r_m; \theta)]^{s-m-1}}{\Gamma(s-m)} f(y; \theta). \tag{4.90}$$

Notice that, $K_m \equiv 1$ is defined for convenience, when the inverse sampling is employed (see [Samaniego and Whitaker, 1986]). The PMLE of (a, b) and the MLP of $Y = R_s$ are obtained by maximizing the logarithm of the predictive likelihood function in equation (4.90) with respect to mentioned parameters. For a special case, when $s = m + 1$, the MLP of $Y = R_{m+1}$, say $\widehat{Y}_{MLP}$, is obtained as

$$\widehat{Y}_{MLP} = \left(\frac{\widehat{a}_{PML} - 1}{\widehat{a}_{PML} \widehat{b}_{PML} - 1} \right)^{1/\widehat{a}_{PML}}. \tag{4.91}$$

Notice that $\widehat{Y}_{MLP}$ is an increasing function with respect to $\widehat{a}_{PML}$ when $\widehat{a}_{PML} > 1$ and $\widehat{Y}_{MLP}$ is a decreasing function with respect to $\widehat{b}_{PML}$ when $\widehat{b}_{PML} > 1$. It can be seen that $\widehat{Y}_{MLP}$ is very big compared to the last record values which violates $\widehat{Y}_{MLP} < r_m$. Hence, we will not list the $\widehat{Y}_{MLP}$ in Tables 4.5 and 4.7.

4.3.5.2. Bayesian Prediction

In this subsection, we consider the problem of prediction and prediction interval of future records based on a Bayesian approach using the SE and the LINEX loss functions by using the bivariate prior in equation (4.69). The conditional density function of $Y = R_s, s > m$ given the past m records is

$$f(y|\,\mathbf{r}, a, b) = \frac{[H(y) - H(r_m)]^{s-m-1}}{\Gamma(s-m)} \frac{f(y)}{F(r_m)}$$

$$= \frac{f(y)}{\Gamma(s-m)F(r_m)} \sum_{j=0}^{s-m-1} \binom{s-m-1}{j} (-1)^j \left[\ln F(y)\right]^j \left[\ln F(r_m)\right]^{s-m-1-j}, \quad (4.92)$$

where $0 < y < r_m$. The Bayes predictive density function Y is given by

$$h(y|\,\mathbf{r}, \mathbf{k}) = \int_0^\infty \int_0^\infty f(y|\,\mathbf{r}, a, b)\pi(a, b|\,\mathbf{r}, \mathbf{k})dadb. \quad (4.93)$$

It is clear that $h(y|\,\mathbf{r}, \mathbf{k})$ cannot be expressed in closed form and hence it cannot be computed analytically.

The consistent estimator of $h(y|\,\mathbf{r}, \mathbf{k})$ is constructed by using the hybrid Metropolis-Hastings and Gibbs sampling procedure as described in MCMC case. Suppose that $\{(a_i, b_i),\ i = 1, ..., N\}$ are MCMC samples obtained from $\pi(a, b|\,\mathbf{r}, \mathbf{k})$ using the hybrid Metropolis -Hastings and Gibbs sampling technique. The consistent estimator of $h(y|\,\mathbf{r}, \mathbf{k})$ based on the simulation can be obtained as

$$\widehat{h}(y|\,\mathbf{r}, \mathbf{k}) = \frac{1}{N} \sum_{i=1}^{N} f(y|\,\mathbf{r}, a_i, b_i), \quad (4.94)$$

and a consistent estimator of the predictive distribution of $Y = R_s$ based on the simulation, say $H(y|\,\mathbf{r},\mathbf{k})$, can be obtained as

$$\widehat{H}(y|\,\mathbf{r},\mathbf{k}) = \frac{1}{N}\sum_{i=1}^{N} F^{*}(y|\,\mathbf{r},a_i,b_i), \tag{4.95}$$

and $F^{*}(y|\,\mathbf{r},a,b)$ denotes the distribution function corresponding to the density function $f(y|\,\mathbf{r},a,b)$ and

$$F^{*}(y|\,\mathbf{r},a,b) = \int_0^y f(t|\,\mathbf{r},a,b)dt$$

$$= \sum_{j=0}^{s-m-1}\binom{s-m-1}{j}\frac{[\ln F(r_m)]^{s-m-1-j}\,\Gamma(j+1,-\ln F(y))}{\Gamma(s-m)F(r_m)}, \tag{4.96}$$

where $\Gamma(x,y)$ is the incomplete Gamma function, i.e. $\Gamma(x,y) = \int_y^{\infty} t^{x-1}e^{-t}dt$. It should be noted that the same MCMC samples $\{(a_i,b_i),\ i=1,...,N\}$ can be used to compute $\widehat{h}(y|\,\mathbf{r},\mathbf{k})$ or $\widehat{H}(y|\,\mathbf{r},\mathbf{k})$ for all y. Then, the point predictor of $Y = R_s$ under the SE loss function is

$$\widehat{Y}_S = \int_0^{r_m} y\widehat{h}(y|\,\mathbf{r},\mathbf{k})dy = \frac{1}{N}\sum_{i=1}^{N}\int_0^{r_m} yf(y|\,\mathbf{r},a_i,b_i)dy$$

$$= \frac{1}{N}\sum_{i=1}^{N}\int_0^{r_m}\frac{yf(y;\mathbf{r},a_i,b_i)}{\Gamma(s-m)F(r_{m;}a_i,b_i)}\left[\ln\left(\frac{F(r_m;a_i,b_i)}{F(y;a_i,b_i)}\right)\right]^{s-m-1}dy. \tag{4.97}$$

The point predictor of $Y = R_s$ under the LINEX loss function is

$$\widehat{Y}_L = -\frac{1}{v}\ln\left\{\frac{1}{N}\sum_{i=1}^{N}\int_0^{r_m}\frac{e^{-vy}f(y;\mathbf{r},a_i,b_i)}{\Gamma(s-m)F(r_{m;}a_i,b_i)}\right.$$

$$\left.\left(\ln\left(\frac{F(r_m;a_i,b_i)}{F(y;a_i,b_i)}\right)\right)^{s-m-1}dy\right\}. \tag{4.98}$$

For a special case, when $s = m+1$, the conditional density function of $Y = R_s$, $s > m$ given $\mathbf{r}$ is $f(y|\,\mathbf{r},a,b) = f(y)/F(r_m)$. Hence, the distribution function of $f(y|\,\mathbf{r},a,b)$ is given by

$$F^{*}(y|\,\mathbf{r},a,b)=\frac{1-(1-y^{a})^{b}}{1-(1-r_{m}^{a})^{b}}. \tag{4.99}$$

Therefore, $\widehat{h}(y|\,\mathbf{r},\mathbf{k})$, $\widehat{H}(y|\,\mathbf{r},\mathbf{k})$, $\widehat{Y}_S$ and $\widehat{Y}_L$ are obtained from equations (4.94), (4.95), (4.97) and (4.98), respectively by using $f(y|\,\mathbf{r},a,b)$ and $F^{*}(y|\,\mathbf{r},a,b)$.

Moreover, a symmetric $100\eta\%$ prediction interval for Y, can be obtained by solving the following non-linear equations, for the lower bound L and upper bound U,

$$\frac{1+\eta}{2}=P(Y>\,L|\,\mathbf{r},\mathbf{k})=1-H(L|\,\mathbf{r},\mathbf{k})\Rightarrow H(L|\,\mathbf{r},\mathbf{k})=\frac{1-\eta}{2}, \tag{4.100}$$

$$\frac{1-\eta}{2}=P(Y>\,U|\,\mathbf{r},\mathbf{k})=1-H(U|\,\mathbf{r},\mathbf{k})\Rightarrow H(U|\,\mathbf{r},\mathbf{k})=\frac{1+\eta}{2}. \tag{4.101}$$

These equations can be easily solved by using the Newton-Raphson method.

4.3.6. Simulation Study

In this section, we present some numerical results to compare the performance of the different methods for different sample sizes and different priors. The performances of the point estimators and predictors are compared by using ERs and MSPEs, respectively. The performances of the confidence, credible and prediction intervals are compared by using average confidence lengths and cps. The ER of θ, when θ is estimated by $\widehat{\theta}$, is given by

$$ER(\theta)=\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{\theta}_i-\theta_i\right)^2, \tag{4.102}$$

under the SE loss function. Moreover, the estimated risk of θ under the LINEX loss function is given by

$$ER(\theta)=\frac{1}{N}\sum_{i=1}^{N}\left(e^{v\left(\widehat{\theta}_i-\theta_i\right)}-v\left(\widehat{\theta}_i-\theta_i\right)-1\right), \tag{4.103}$$

Table 4.3: Results for the true value of $b = 2.8545$, $(\alpha, \gamma) = (1, 4)$ and a is known.

MLE	SEL	Bayes estimates LINEX $v=-2$	$v=-1$	$v=1$	$v=2$	Exact MLE confidence interval	Bayesian credible interval
			$m=5$				
3.8549	3.0277	4.6575	3.5362	2.6931	2.4472	(1.3242,8.3535)	(1.2343,5.7272)
6.1387	1.5483	1.4888	0.2401	0.1582	0.5332	7.0293/0.9347	4.4932/0.9527
			$m=7$				
3.6323	3.0306	4.1218	3.4216	2.7537	2.5407	(1.4604,6.7765)	(1.3858,5.3080)
4.1676	1.2032	0.9735	0.1784	0.1295	0.4574	5.3162/0.9483	3.9222/0.9560
			$m=10$				
3.3350	2.9634	3.6384	3.2380	2.7527	2.5822	(1.5992,5.6977)	(1.5312,4.8605)
2.2451	0.9010	0.6113	0.1281	0.1001	0.3660	4.0985/0.9513	3.3292/0.9553
			$m=12$				
3.1984	2.9161	3.4460	3.1414	2.7366	2.5876	(1.6527,5.2459)	(1.5943,4.6304)
1.5528	0.7788	0.4823	0.1090	0.0883	0.3246	3.5933/0.9533	3.0362/0.9533
			$m=15$				
3.1576	2.9234	3.3502	3.1104	2.7696	2.6390	(1.7673,4.9448)	(1.7030,4.4682)
1.2799	0.6526	0.4120	0.0896	0.0753	0.2813	3.1775/0.9440	2.7652/0.9503

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their lengths and cp's.

where N is the number of replication. Similarly, the MSPEs can be computed with respect to the SE and the LINEX loss functions. All of the computations are performed by using Matlab R2010a. All the results are based on 3000 replications.

In Table 4.3, the ML and Bayes estimates under the SE and the LINEX ($v = -2, -1, 1$ and 2) loss functions with their corresponding ERs are listed when a is known $(a = 3)$, the true value of $b = 2.8545$ and the prior parameters of b are chosen to be $(\alpha, \gamma) = (1, 4)$. Since the exact distribution of the MLE of b is known, the 95% exact confidence intervals are easily constructed. Moreover, the 95% Bayesian credible interval for b which is obtained by using the posterior distribution of b are listed. From Table 4.3, the average ERs of b decrease as the sample size increases in all cases, as expected. The Bayes estimates under the SE and the LINEX loss functions have smaller ER than that of MLEs. The average lengths of the intervals decrease as the sample size increases. The lengths of the Bayesian credible intervals are smaller than that of the exact confidence intervals. Also, the coverage probabilities are quite close to the nominal level 95%.

In Tables 4.4 and 4.5, the ML and Bayes estimates under the SE and the LINEX loss functions with their corresponding ERs are listed for a and b when $a = 3$, $b = 5$

and the prior parameters are Prior 1: $(\delta, \beta) = (0.5, 4)$ and $(\alpha, \gamma) = (3, 2)$. The point predictors and the 95% prediction interval for $Y = R_{m+1}$ by using Bayesian (with respect to the SE and the LINEX loss functions) method are also listed. Moreover, the 95% asymptotic and HPD credible intervals with their coverage probabilities are listed. The Bayes point estimates are computed by using Lindley's approximation and MCMC method under the SE and the LINEX ($v = -2, -1, 1$ and 2) loss functions.

From Table 4.4, the average ERs of a and b generally decrease as the sample size increases. The ERs of Bayes estimates under the SE and the LINEX loss functions are smaller than that of MLEs. On the other hand, the ERs of the Bayes estimates for a based on the Lindley's approximation and MCMC methods are close to each other under the SE and the LINEX loss functions. However, the ERs of the Bayes estimates for b based on the Lindley's approximation and MCMC methods are close to each other under both the SE and the LINEX loss functions when $v > 0$. The ERs of the Bayes estimates for b based on the Lindley's approximation and MCMC methods close each other under the LINEX loss function as the sample size increases when $v < 0$. Furthermore, the average lengths of the intervals decrease as the sample size increases. The average lengths of the HPD credible intervals are smaller than that of the asymptotic confidence intervals but the cp values of asymptotic confidence intervals are more close to the nominal value.

From Table 4.5, the average MSPEs of the point predictors decrease as the sample size increases in all cases. Also, the average lengths of the prediction intervals decrease as the sample size increases and their coverage probabilities are quite close to the nominal value.

In Tables 4.6 and 4.7, the ML and Bayes estimates under the SE and the LINEX loss functions with their corresponding ERs are listed for a and b when $a = 10, b = 4.5$ and the prior parameters are Prior 2: $(\delta, \beta) = (5, 2)$ and $(\alpha, \gamma) = (6, 5)$. The point predictors and the 95% prediction interval for $Y = R_{m+1}$ by using Bayesian (with respect to the SE and the LINEX loss functions) method are also listed. Moreover, the 95% asymptotic and HPD credible intervals with their coverage probabilities are listed. The Bayes point estimates are computed by using Lindley's approximation and MCMC method under the SE and the LINEX ($v = -2, -1, 1$ and 2) loss functions.

From Table 4.6, the average ERs of a and b generally decrease as the sample size

Table 4.4: Results for the true values of $(a, b) = (3, 5)$ using Prior 1.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic	HPD
			LINEX					LINEX				confidence interval	credible interval
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp	length/cp
							$m=5$						
a	3.4171	2.5827	2.6797	2.6413	2.5091	2.4365	2.5819	2.7430	2.6596	2.5088	2.4392	(1.4731,5.3613)	(1.8603,3.3305)
	1.3539	0.2476	0.5810	0.1287	0.1249	0.4987	0.2489	0.4648	0.1212	0.1252	0.4952	3.8882/0.9543	1.4703/0.8520
b	430.3868	3.2481	3.5730	3.5596	2.6050	2.3103	3.2417	5.7582	4.3524	2.6591	2.2910	(0,3606.8)	(1.1221,5.6552)
	$6.75\text{x}10^7$	3.2667	23.1728	2.3254	1.4906	4.3850	3.2885	2.0629	0.8544	1.4422	4.4230	3606.8/0.9410	4.5331/0.8213
							$m=7$						
a	3.1461	2.6725	2.7323	2.7067	2.6320	2.5900	2.6710	2.7544	2.7122	2.6305	2.5907	(1.9651,4.3271)	(2.1359,3.2166)
	0.4827	0.1504	0.3305	0.0771	0.0752	0.2999	0.1521	0.2979	0.0756	0.0757	0.2994	2.3620/0.9387	1.0806/0.8367
b	65.0268	3.2386	3.5418	3.5306	2.6154	2.3115	3.2324	5.6603	4.2932	2.6653	2.3034	(0,414.0633)	(1.1422,5.6069)
	$2.96\text{x}10^6$	3.2720	24.2915	2.3826	1.4806	4.3824	3.2956	2.1911	0.8698	1.4360	4.3982	414.0633/0.9300	4.4647/0.8153
							$m=10$						
a	3.0802	2.7582	2.7973	2.7797	2.7337	2.7080	2.7570	2.8063	2.7816	2.7326	2.7082	(2.2888,3.8717)	(0.23416,3.1760)
	0.1836	0.0838	0.1745	0.0422	0.0422	0.1703	0.0847	0.1656	0.0420	0.0426	0.1705	1.5829/0.9490	0.8344/0.8597
b	11.7617	3.3278	3.6282	3.6149	2.7198	2.4177	3.3241	5.6591	4.3419	2.7665	2.4062	(0.0010,34.8776)	(1.2519,5.6680)
	1416.4	2.9878	21.3242	2.1336	1.3876	4.1716	3.0030	2.5387	0.8252	1.3466	4.1940	34.8766/0.9460	4.4161/0.8317
							$m=12$						
a	3.0891	2.7929	2.8276	2.8118	2.7717	2.7494	2.7918	2.8344	2.8131	2.7707	2.7496	(2.3939,3.7844)	(2.4058,3.1800)
	0.1517	0.0651	0.1328	0.0325	0.0330	0.1340	0.0656	0.1268	0.0323	0.0332	0.1340	1.3905/0.9500	0.7743/0.8900
b	9.9987	3.4032	3.7088	3.6911	2.8101	2.5142	3.4000	5.6307	4.3689	2.8543	2.4978	(0.0071,26.5346)	(1.3251,5.7096)
	525.894	2.7525	18.4583	1.9068	1.3080	3.9802	2.7653	2.6523	0.7783	1.2698	4.0123	26.5275/0.9540	4.3574/0.8383
							$m=15$						
a	3.0912	2.8193	2.8493	2.8355	2.8015	2.7828	2.8183	2.8538	2.8361	2.8007	2.7831	(2.4889,3.6936)	(2.4669,3.1690)
	0.1096	0.0495	0.0979	0.0244	0.0253	0.1038	0.0501	0.0947	0.0244	0.0256	0.1041	1.2047/0.9513	0.7021/0.9010
b	8.3803	3.5405	3.8555	3.8339	2.9539	2.6653	3.5382	5.7044	4.4808	2.9976	2.6412	(0.0630,19.5804)	(1.4998,5.8237)
	187.155	2.3791	14.5525	1.5848	1.1846	3.6817	2.3897	2.3789	0.7085	1.1481	3.7287	19.5173/0.9563	4.3239/0.8530

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 4.5: Predictions based on lower records with inter-record times using Prior 1.

r_m / r_{m+1}	Bayes point predictors					Prediction interval length/cp
	SEL	LINEX				
		$v=-2$	$v=-1$	$v=1$	$v=2$	
			$m=5$			
0.1170	0.0838	0.0845	0.0842	0.0835	0.0831	(0.0264,0.1159)
0.0707	0.0013	0.0026	0.0006	0.0007	0.0027	0.0895/0.9480
			$m=7$			
0.0610	0.0443	0.0445	0.0444	0.0442	0.0441	(0.0149,0.0605)
0.0366	0.0004	0.0008	0.0002	0.0002	0.0008	0.0455/0.9403
			$m=10$			
0.0361	0.0264	0.0265	0.0265	0.0264	0.0264	(0.0093,0.0357)
0.0230	0.0001	0.0002	0.0001	0.0001	0.0002	0.0264/0.9633
			$m=12$			
0.0338	0.0249	0.0249	0.0249	0.0249	0.0248	(0.0089,0.0335)
0.0222	0.0001	0.0002	0	0	0.0002	0.0246/0.9603
			$m=15$			
0.0310	0.0229	0.0229	0.0228	0.0228	0.0228	(0.0083,0.0307)
0.0215	0.0001	0.0001	0	0	0.0001	0.0224/0.9717

Notes: First column: The first row represents the average of the r_m th record values and second row represents the average of the true values (r_{m+1}) which we want to predict. Last column, the first row represents a 95% PI and second row represents their lengths and cp's. For the others, the first row represents the average predictors and second row represents corresponding MSPEs for each choice of m.

increases. The ERs of Bayes estimates under the SE and the LINEX loss functions are smaller than that of MLEs. On the other hand, the ERs of the Bayes estimates for a and b based on the Lindley's approximation and MCMC methods are generally close to each other under both the SE and the LINEX loss functions except for some cases. Furthermore, the average lengths of the intervals decrease as the sample size increases. The average lengths of the HPD credible intervals are smaller than that of the asymptotic confidence intervals. The HPD credible interval is preferable to the asymptotic confidence interval with respect to length and the cp value.

From Table 4.7, the average MSPEs of the point predictors decrease as the sample size increases in all cases. Also, the average lengths of the prediction intervals decrease and their coverage probabilities close to the nominal value as the sample size increases.

In Tables 4.8-4.11, to observe the effect of the inter-record times in parameter estimates, we generate lower and upper records by using the following procedure.

Table 4.6: Results for the true values of $(a,b)=(10,4.5)$ using Prior 2.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic confidence interval	HPD credible interval
			LINEX					LINEX					
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp	length/cp
							$m=5$						
a	11.1515	9.3657	9.8095	9.7561	8.7593	8.5317	9.3638	11.0250	10.1665	8.7278	8.1952	(4.7492,17.5541)	(7.1711,11.6745)
	11.2941	1.1322	6.1442	0.6723	0.6374	2.1100	1.1379	2.2856	0.5266	0.6554	2.6953	12.8049/0.9667	4.5033/0.9460
b	41.2933	3.8319	4.2420	4.1922	3.2468	3.0038	3.8285	5.8857	4.7297	3.2757	2.8945	(0,243.6592)	(1.7532,6.1200)
	185040	0.6776	1.2415	0.2534	0.5595	2.0609	0.6830	2.0568	0.2337	0.5393	2.2595	243.6592/0.9540	4.3668/0.9907
							$m=7$						
a	10.4953	9.4765	9.8009	9.7269	9.1163	8.9119	9.4733	10.2903	9.8703	9.1115	8.7785	(6.5285,14.4622)	(7.8665,11.1342)
	5.0228	0.7117	2.5000	0.3850	0.3903	1.4173	0.7260	1.3478	0.3337	0.3941	1.6309	7.9337/0.9473	3.2677/0.9290
b	32.5050	3.7841	4.1658	4.1198	3.2175	2.9638	3.7798	5.7329	4.6299	3.2483	2.8781	(0,155.8264)	(1.7439,6.0205)
	85021	0.6821	1.3588	0.2525	0.5738	2.1289	0.6904	1.8582	0.1909	0.5524	2.2885	155.8264/0.9313	4.2766/0.9857
							$m=10$						
a	10.3130	9.6236	9.8646	9.7917	9.3979	9.2355	9.6188	10.0819	9.8480	9.3977	9.1849	(7.6552,12.9709)	(8.3716,10.8901)
	2.2960	0.3834	1.0174	0.1911	0.2173	0.8681	0.3928	0.8115	0.1834	0.2179	0.9388	5.3157/0.9467	2.5185/0.9417
b	15.3741	3.7958	4.1604	4.1155	3.2473	2.9897	3.7910	5.6484	4.5963	3.2771	2.9156	(0.0010,59.9217)	(1.7949,5.9857)
	42705	0.6668	1.3921	0.2570	0.5531	2.0797	0.6793	1.7657	0.1955	0.5329	2.2173	59.9207/0.9393	4.1909/0.9777
							$m=12$						
a	10.2938	9.6735	9.8969	9.8253	9.4746	9.3248	9.6684	10.0688	9.8676	9.4739	9.2851	(7.9692,12.6184)	(8.5035,10.8589)
	1.6954	0.3147	0.6909	0.1502	0.1803	0.7367	0.3224	0.5785	0.1457	0.1810	0.7890	4.6492/0.9533	2.3554/0.9507
b	9.2548	3.8223	4.1821	4.1350	3.2905	3.0353	3.8168	5.5897	4.5869	3.3180	2.9636	(0.0105,25.2429)	(1.8528,5.9728)
	901.303	0.6476	1.4251	0.2611	0.5251	1.9957	0.6614	1.7470	0.2089	0.5072	2.1272	25.2323/0.9467	4.1200/0.9773
							$m=15$						
a	10.3292	9.7397	9.9429	9.8736	9.5692	9.4345	9.7374	10.0752	9.9064	9.5710	9.4094	(8.3089,12.3496)	(8.6608,10.8290)
	1.2728	0.2367	0.4965	0.1121	0.1366	0.5760	0.2389	0.4256	0.1078	0.1361	0.6064	4.0407/0.9407	2.1681/0.9613
b	7.5184	3.8870	4.2443	4.1950	3.3693	3.1171	3.8843	5.5945	4.6256	3.3971	3.0487	(0.0619,17.2718)	(1.9445,6.0066)
	112.329	0.5752	1.2522	0.2359	0.4737	1.8454	0.5810	1.6956	0.2013	0.4565	1.9684	17.2099/0.9580	4.0621/0.9813

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 4.7: Predictions based on lower records with inter-record times using Prior 2.

r_m / r_{m+1}	Bayes point predictors					Prediction interval
		LINEX				
	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp
			$m=5$			
0.5245	0.4728	0.4750	0.4739	0.4716	0.4704	(0.3485,0.5231)
0.4485	0.0041	0.0078	0.0020	0.0021	0.0087	0.1746/0.9297
			$m=7$			
0.4307	0.3891	0.3905	0.3898	0.3884	0.3876	(0.2896,0.4295)
0.3677	0.0029	0.0056	0.0014	0.0015	0.0061	0.1399/0.9160
			$m=10$			
0.3676	0.3328	0.3338	0.3333	0.3323	0.3317	(0.2495,0.3666)
0.3172	0.0019	0.0036	0.0009	0.0010	0.0039	0.1171/0.9273
			$m=12$			
0.3600	0.3261	0.3270	0.3265	0.3256	0.3251	(0.2450,0.3590)
0.3131	0.0017	0.0032	0.0008	0.0008	0.0034	0.1140/0.9280
			$m=15$			
0.3509	0.3181	0.3190	0.3186	0.3177	0.3172	(0.2396,0.3500)
0.3089	0.0013	0.0026	0.0006	0.0007	0.0027	0.1104/0.9417

Notes: First column: The first row represents the average of the r_m th record values and second row represents the average of the true values (r_{m+1}) which we want to predict. Last column, the first row represents a 95% PI and second row represents their lengths and cp's. For the others, the first row represents the average predictors and second row represents corresponding MSPEs for each choice of m.

- Step 1: A random sample are generated from the Kumaraswamy distribution with parameters (a, b) and sample size n.
- Step 2: The lower record values with their corresponding inter-record times and the upper record values are saved. Notice that the sample sizes of the lower and the upper record values can be different. Moreover, the number of recod values in a random sample with size n is approximately $\ln(n)$.
- Step 3: The estimates of a and b are computed based on lower record values with their corresponding inter-record times.
- Step 4: The estimates of a and b are also computed based on only upper record values.
- Step 5: Repeat Steps 1-4, 3000 times and obtain the samples (a_i, b_i).

In Tables 4.8 and 4.9, the ML and Bayes estimates under the SE and the LINEX loss functions with their corresponding ERs for a and b are listed when $a = 4$, $b = 10$

and the prior parameters are Prior 3: $(\delta, \beta) = (8, 0.5)$ and $(\alpha, \gamma) = (4, 7)$. Moreover, the 95% asymptotic and HPD credible intervals with their coverage probabilities are listed. The Bayes point estimates are computed by using Lindley's approximation and MCMC method under the SE and the LINEX ($v = -2, -1, 1$ and 2) loss functions.

From Tables 4.8 and 4.9, the average ERs of a and b generally decrease as the sample size increases except for some cases. Moreover, the average lengths of the intervals decrease as the sample size increases. The ERs of a which are obtained by using lower records with their corresponding inter-record times are smaller than the one based on only upper record values. The average lengths of the intervals for a which are obtained by using lower records with their corresponding inter-record times are shorter than the one based on only upper record values. However, the cp values for upper record case are more close to the nominal value than that of lower record case. The ERs of b which are obtained by using lower records with their corresponding inter-record times are smaller than the one based on only upper record values except for some LINEX cases. The results for the asymptotic intervals of b is similar to the interval results of a. The HPD credible intervals of b which are obtained by using upper record values have a good results with respect to lower record case.

In Tables 4.10 and 4.11, the ML and Bayes estimates under the SE and the LINEX loss functions with their corresponding ERs for a and b are listed when $a = 2$, $b = 3$ and the prior parameters are Prior 4: $(\delta, \beta) = (2, 1)$ and $(\alpha, \gamma) = (2, 1.5)$. Moreover, the 95% asymptotic and HPD credible intervals with their coverage probabilities are listed. The Bayes point estimates are computed by using Lindley's approximation and MCMC method under the SE and the LINEX ($v = -2, -1, 1$ and 2) loss functions.

From Tables 4.10 and 4.11, the average ERs of a and b generally decrease as the sample size increases except for some cases. Moreover, the average lengths of the intervals generally decrease as the sample size increases. The comparison of the lower records with their corresponding inter-record times and upper records are almost same in Tables 6 and 7. Moreover, it is observed that the cp values of HPD credible intervals of a and b are around the nominal value.

In the all MCMC case, we run three MCMC chains with fairly different initial values and generated 10000 iterations for each chain. To diminish the effect of the starting distribution, we generally discard the first half of each sequence and focus on the second half. To provide relatively independent samples for improvement of prediction accuracy, we calculate the Bayesian MCMC estimates by the means of every 5^{th} sampled values after discarding the first half of the chains (see[Gelman et al., 2003]). The scale reduction factor estimate $\sqrt{\widehat{R}} = \sqrt{\frac{Var(\psi)}{W}}$ is used to monitor convergence of MCMC simulations where ψ is the estimand of interest, $Var(\psi) = \frac{n-1}{n}W + \frac{1}{n}B$ with the iteration number n for each chain, the between- and within- sequence variances B and W (see [Gelman et al., 2003]). It is observed that the scale factor value of the MCMC estimators are found below 1.1 which is an acceptable value for their convergency for all cases in Tables 4.4, 4.6 and 4.8-4.11.

4.3.7. Conclusion

In this paper, firstly we consider the non-Bayesian and Bayesian point estimates as well as confidence intervals for the unknown parameters of the Kumaraswamy distribution based on lower record values with their corresponding inter-record times. The ML estimates of the unknown parameters are derived under the inverse sampling scheme. The Lindley's approximation and MCMC methods are used to get the Bayes estimates under the SE and the LINEX loss function for the bivariate prior. Monte Carlo simulation reveals out that the ERs of the Bayes estimates are smaller than that of MLEs under the SE and the LINEX loss functions. The average length of the HPD credible intervals are smaller than that of the asymptotic intervals. Moreover, the Bayesian point predictors as well as prediction intervals for the future lower record values are considered. The point predictors and prediction intervals of the future lower record values are computed based on the lower record values with their corresponding inter-record times. The result of the point predictors and prediction intervals are satisfactory when it is compared to the real values.

Secondly, a random sample generating from the Kumaraswamy distribution is allocated lower records with correspondig inter-record times and upper records, non-Bayesian and Bayesian parameter estimates are considered by using these record values. Therefore, we can see the effect of considering the inter-record times in the parameter estimates. We obtain that using the inter-record times generally decreases

Table 4.8: Results based on lower records with inter-record times using Prior 3.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic	HPD
			LINEX					LINEX				confidence interval	credible interval
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp	length/cp
							$n=1000$						
a	4.0419	3.7969	3.8770	3.8438	3.7402	3.6833	3.7746	3.8704	3.8231	3.7252	3.6752	(2.7899,5.3252)	(3.2119,4.3149)
	1.2785	0.1160	0.3065	0.0621	0.0585	0.2399	0.1373	0.3532	0.0764	0.0648	0.2530	2.5354/0.8680	1.1030/0.8230
b	9.4442	9.1931	9.6058	9.9769	6.5682	6.6498	9.0888	19.6408	16.7792	6.0969	4.8328	(0.5056,19.8549)	(3.5204,15.3559)
	134.613	2.4399	40.5245	1.4691	2.6752	320.192	3.0135	44.3458	6.5657	2.9291	9.3345	19.3493/0.6713	11.8355/0.9740
							$n=5000$						
a	3.9782	3.8365	3.8951	3.8698	3.7974	3.7571	3.8154	3.8797	3.8480	3.7819	3.7476	(3.0719,4.8881)	(3.3499,4.2584)
	0.5562	0.0758	0.1666	0.0385	0.0389	0.1611	0.0948	0.2394	0.0519	0.0449	0.1747	1.8161/0.8507	0.9085/0.8180
b	9.5895	9.1210	9.5083	9.8717	6.5239	6.5866	9.0037	19.1799	16.3547	6.1002	4.8578	(0.6752,19.6067)	(3.5493,15.1206)
	91.7067	2.4170	138.481	2.0592	2.5297	5.9913	3.1168	174.226	7.1909	2.9266	9.2845	18.9315/0.7097	11.5713/0.9647
							$n=10000$						
a	3.9713	3.8484	3.9002	3.8775	3.8146	3.7793	3.8228	3.8767	3.8502	3.7945	3.7654	(3.1605,4.7827)	(3.3962,4.2191)
	0.4311	0.0637	0.1344	0.0319	0.0329	0.1367	0.0874	0.2347	0.0491	0.0406	0.1560	1.6222/0.8470	0.8228/0.7907
b	8.8843	9.1033	9.4827	9.8427	6.5196	6.5750	8.9493	18.8550	16.0598	6.0999	4.8691	(0.8750,17.6104)	(3.5652,14.9682)
	69.0105	2.5247	90.3264	2.1482	2.5315	5.8615	3.4372	922.744	8.4144	2.9281	9.2620	16.7354/0.6683	11.4030/0.9543

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 4.9: Results based on only upper records using Prior 3.

	MLE	Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic confidence interval	HPD credible interval
			LINEX					LINEX					
		SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp	length/cp
							$n=1000$						
a	6.1204	3.9586	4.2724	4.1965	3.6227	3.4268	3.9549	4.8117	4.3568	3.6342	3.3719	(0.6795,11.8871)	(2.4070,5.6048)
	78.6163	0.2497	0.5449	0.1465	0.1208	0.5732	0.2492	1.1492	0.1991	0.1187	0.6124	11.2075/0.9667	3.1978/0.9920
b	6.93×10^{11}	9.3019	9.9866	10.2166	7.3608	7.3649	9.2971	17.1038	14.4537	6.9235	5.7158	(0.1665,6.57×10^{12})	(4.5503,14.5171)
	1.44×10^{27}	4.0220	174.592	3.4530	2.0026	9.2925	4.0352	13.2527	3.7573	2.2073	7.5729	6.57×10^{12}/0.9530	9.9668/0.9743
							$n=5000$						
a	5.4688	3.9575	4.2697	4.1937	3.6245	3.4287	3.9550	4.8039	4.3522	3.6364	3.3748	(0.7061,10.4100)	(2.4121,5.5979)
	29.1548	0.2384	0.5289	0.1411	0.1177	0.5640	0.2394	1.1262	0.1922	0.1164	0.6104	9.7039/0.9610	3.1858/0.9933
b	1.54×10^{7}	9.3480	10.0461	10.2452	7.5361	7.5147	9.3436	16.4863	13.9217	7.1359	5.9620	(0.5582,5.24×10^{8})	(4.8109,14.2990)
	7.08×10^{17}	3.9976	144.865	3.2914	1.9173	9.5429	4.0061	12.0605	3.3157	2.0408	7.0846	5.24×10^{8}/0.9503	9.4881/0.9673
							$n=10000$						
a	5.2535	3.9434	4.2544	4.1784	3.6140	3.4190	3.9412	4.7824	4.3346	3.6253	3.3659	(0.8225,10.0112)	(2.4062,5.5763)
	14.0463	0.2445	0.5276	0.1412	0.1230	0.5832	0.2446	1.1018	0.1901	0.1208	0.6210	9.1887/0.9533	3.1701/0.9937
b	17.9920	9.3707	10.0728	10.2596	7.6095	7.5784	9.3675	16.2145	13.6981	7.2235	6.0611	(0.8318,59.4862)	(4.9193,14.2159)
	17991	3.8621	265.925	3.6365	1.8020	7.3480	3.8699	11.7712	3.2411	1.9557	6.8868	58.6541/0.9473	9.2966/0.9667

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 4.10: Results based on lower records with inter-record times using Prior 4.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic	HPD
			LINEX					LINEX				confidence interval	credible interval
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp	length/cp
							$n=1000$						
a	2.1120	1.8584	1.8952	1.8789	1.8344	1.8087	1.8548	1.9018	1.8783	1.8312	1.8077	(1.3899,2.8453)	(1.4604,2.2562)
	0.4366	0.0489	0.1049	0.0248	0.0249	0.1032	0.0512	0.1045	0.0257	0.0258	0.1045	1.4553/0.9009	0.7958/0.9159
b	2.9683	2.5961	2.8175	2.8040	2.1243	1.8232	2.5880	4.5357	3.4077	2.1509	1.8703	(0.1919,5.9990)	(0.8085,4.6520)
	6.6099	0.3715	0.9396	0.1836	0.3138	1.4703	0.3814	2.4488	0.2831	0.3024	1.3813	5.8072/0.7718	3.8435/0.9803
							$n=5000$						
a	2.1470	1.9421	1.9680	1.9562	1.9260	1.9088	1.9399	1.9718	1.9559	1.9238	1.9078	(1.6085,2.6875)	(1.6127,2.2715)
	0.1795	0.0218	0.0450	0.0109	0.0111	0.0462	0.0226	0.0455	0.0113	0.0115	0.0467	1.0789/0.8740	0.6588/0.9533
b	3.5778	2.6517	2.8758	2.8604	2.1846	1.8859	2.6461	4.5494	3.4434	2.2110	1.9287	(0.2854,7.1125)	(0.8658,4.6988)
	15.2483	0.3392	0.8821	0.1753	0.2840	1.3620	0.3464	2.4896	0.2866	0.2739	1.2840	6.8372/0.8350	3.8331/0.9813
							$n=10000$						
a	2.2291	1.9886	2.0109	2.0006	1.9751	1.9607	1.9866	2.0130	1.9999	1.9732	1.9598	(1.7200,2.7390)	(1.6898,2.2856)
	0.2354	0.0161	0.0344	0.0083	0.0080	0.0321	0.0169	0.0362	0.0087	0.0083	0.0327	1.0190/0.8270	0.5958/0.9677
b	3.2761	2.7254	2.9575	2.9417	2.2457	1.9465	2.7198	4.6204	3.5224	2.2753	1.9860	(0.4814,6.1369)	(0.9147,4.7926)
	6.7265	0.2997	0.7515	0.1589	0.2542	1.2608	0.3072	2.5805	0.3052	0.2434	1.1885	5.6555/0.8073	3.8779/0.9853

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

Table 4.11: Results based on only upper records using Prior 4.

		Bayes estimates using Lindley					Bayes estimates using MCMC					Asymptotic	HPD
			LINEX					LINEX				confidence interval	credible interval
	MLE	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	SEL	$v=-2$	$v=-1$	$v=1$	$v=2$	length/cp	length/cp
							$n=1000$						
a	3.9429	1.9190	2.0980	2.0541	1.6918	1.5104	1.9092	2.4710	2.1821	1.6997	1.5426	(0.0448,8.7381)	(0.7509,3.2652)
	58.8860	0.1780	0.3815	0.1004	0.0764	0.4004	0.1759	0.7083	0.1332	0.0805	0.3815	8.6933/0.9643	2.5143/0.9746
b	16.5897	2.7509	3.0504	2.9799	2.4191	2.2184	2.7498	3.8272	3.1995	2.4381	2.2058	(0.4493,133.828)	(1.2440,4.4202)
	181621	0.5727	1.5667	0.3192	0.2753	1.1622	0.5736	1.6309	0.3307	0.2695	1.0341	133.379/0.9666	3.1762/0.9393
							$n=5000$						
a	3.7150	1.9185	2.0966	2.0522	1.6948	1.5142	1.9099	2.4737	2.1824	1.7017	1.5454	(0.0457,8.1304)	(0.7497,3.2725)
	37.3458	0.1668	0.3614	0.0944	0.0740	0.3934	0.1674	0.6983	0.1283	0.0789	0.3839	8.0846/0.9647	2.5228/0.9790
b	4.3137	2.7885	3.0795	3.0042	2.4892	2.3012	2.7881	3.7102	3.1777	2.5052	2.2876	(0.7569,12.8132)	(1.3513,4.3681)
	911.213	0.5228	1.3702	0.2889	0.2473	0.9893	0.5234	1.4097	0.2963	0.2433	0.9203	12.0563/0.9657	3.0168/0.9417
							$n=10000$						
a	3.5896	1.9083	2.0849	2.0405	1.6878	1.5091	1.9023	2.4642	2.1741	1.6947	1.5392	(0.0432,7.8318)	(0.7470,3.2624)
	22.8608	0.1712	0.3634	0.0953	0.0765	0.4008	0.1702	0.6934	0.1295	0.0782	0.3724	7.7886/0.9610	2.5154/0.9793
b	3.6209	2.8088	3.0962	3.0194	2.5206	2.3371	2.8074	3.6804	3.1782	2.5343	2.3222	(0.8745,6.4874)	(1.3951,4.3575)
	3.6558	0.4979	1.3410	0.2791	0.2306	0.8959	0.4983	1.3760	0.2863	0.2274	0.8529	5.6129/0.9690	2.9624/0.9420

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their cp's.

the ERs of the parameter estimates. As a result, we suggest using lower record values with their corresponding inter-record times instead of just using upper record values to get the parameter estimates when these records are obtained from the common sample.

4.4. Estimation of The Reliability Based on Record Values

The problem of estimating of R on random samples has been extensively studied under various distributional assumptions on X and Y. A comprehensive account of this topic is presented by [Kotz et al., 2003]. It is provided an excellent review of the development of the stress-strength reliability under classical and Bayesian point of views up to the year 2003. For most recent results on the topic see [Kundu and Gupta, 2005], [Mokhlis, 2005], [Baklizi, 2008], [Rezaei et al., 2010], [Nadar et al., 2014] and the references therein.

The main purpose of this section is to improve inference procedures for the stress-strength model based on upper record values when the measurements follow the Kumaraswamy distribution with the first shape parameters are same. Different estimators of R are obtained, namely, ML, UMVU and Bayesian and empirical Bayesian estimates under the SE and the LINEX loss functions corresponding to conjugate and non informative priors. Moreover, exact, asymptotic and Bayesian credible intervals of R are also obtained.

4.4.1. Estimation of R When a Is Common and Unknown

The ML estimates, its existence and uniqueness, asymptotic confidence intervals, as well as Bayes estimates and Bayesian credible interval for R are obtained when the first shape parameter is common for the distributions of X and Y.

4.4.1.1. ML Estimation of R

Let $X \sim Kum(a, b_1)$ and $Y \sim Kum(a, b_2)$ are independent random variables. Then, the reliability R is

$$R = P(X < Y) = \int_0^1 f_Y(y) P(X < Y \mid Y = y) dy$$

$$= \frac{b_1}{b_1 + b_2}. \tag{4.104}$$

The estimate of R are considered based on upper record data on both variables. Let $R_1, \ldots, R_n$ be a set of upper record values from $Kum(a, b_1)$ and $S_1, \ldots, S_m$ be a set of upper records from $Kum(a, b_2)$ independently from the first sample. The joint likelihood function of (b_1, b_2, a) given $(\underline{r}, \underline{s})$ based on records is given by, see [Arnold et al., 1998]

$$L(b_1, b_2, a \,|\underline{r}, \underline{s}) = h_1(\underline{r}; a) h_2(\underline{s}; a) a^{n+m} b_1^n b_2^m e^{-b_1 T_1(r_n; a)} e^{-b_2 T_2(s_m; a)} \tag{4.105}$$

where $\underline{r} = (r_1, \ldots, r_n)$, $\underline{s} = (s_1, \ldots, s_m)$, $T_1(r_n; a) = -\ln(1 - r_n^a)$, $T_2(s_m; a) = -\ln(1 - s_m^a)$ and

$$h_1(\underline{r}; a) = \prod_{i=1}^{n} \frac{r_i^{a-1}}{1 - r_i^a}, \; h_2(\underline{s}; a) = \prod_{j=1}^{m} \frac{s_j^{a-1}}{1 - s_j^a}. \tag{4.106}$$

Then, the joint log-likelihood function is

$$l(b_1, b_2, a \,|\underline{r}, \underline{s}) = \ln h_1(\underline{r}; a) + \ln h_2(\underline{s}; a) + (n + m) \ln a$$

$$+ n \ln b_1 + m \ln b_2 + -b_1 T_1(r_n; a) - b_2 T_2(s_m; a). \tag{4.107}$$

The ML estimates of b_1, b_2 and a, say $\widehat{b}_1$, $\widehat{b}_2$ and $\widehat{a}$ respectively, are given by

$$\widehat{b}_1 = -\frac{n}{\ln(1 - r_n^{\widehat{a}})}, \tag{4.108}$$

$$\widehat{b}_2 = -\frac{m}{\ln(1 - s_m^{\widehat{a}})}, \tag{4.109}$$

and $\widehat{a}$ is the solution of the following non-linear equation

$$\frac{n+m}{a}+\sum_{i=1}^{n}\frac{\ln r_i}{1-r_i^a}+\frac{nr_n^a\ln r_n}{(1-r_n^a)\ln(1-r_n^a)}$$

$$+\sum_{j=1}^{m}\frac{\ln s_j}{1-s_j^a}+\frac{ms_m^a\ln s_m}{(1-s_m^a)\ln(1-s_m^a)}=0. \quad (4.110)$$

Therefore, $\widehat{a}$ can be obtained as a solution of the non-linear equation of the form $h(a)=a$ where

$$h(a)=-(n+m)\left[\sum_{i=1}^{n}\frac{\ln r_i}{1-r_i^a}+\frac{nr_n^a\ln r_n}{(1-r_n^a)\ln(1-r_n^a)}\right.$$

$$\left.+\sum_{j=1}^{m}\frac{\ln s_j}{1-s_j^a}+\frac{ms_m^a\ln s_m}{(1-s_m^a)\ln(1-s_m^a)}\right]^{-1}. \quad (4.111)$$

Since, $\widehat{a}$ is a fixed point solution of the non-linear equation (4.110), its value can be obtained using an iterative scheme as: $a_{(j+1)}=h(a_{(j)})$, where $a_{(j)}$ is the jth iterate of $\widehat{a}$. The iteration procedure should be stopped when $\left|a_{(j+1)}-a_{(j)}\right|$ is sufficiently small. After $\widehat{a}$ is obtained, $\widehat{b}_1$ and $\widehat{b}_2$ can be obtained from equations (4.108) and (4.109), respectively. Therefore, the MLE of R, say $\widehat{R}$, is

$$\widehat{R}=\frac{\widehat{b}_1}{\widehat{b}_1+\widehat{b}_2}. \quad (4.112)$$

Next, the existence and uniqueness of the ML estimates of the parameters b_1, b_2 and a are proved.

Theorem 4.3: The ML estimates of the parameters b_1, b_2 and a are unique and given by $\widehat{b}_1=-n/\ln(1-r_n^{\widehat{a}})$, $\widehat{b}_2=-m/\ln(1-s_m^{\widehat{a}})$ where $\widehat{a}$ is the solution of the non-linear equation

$$G(a)\equiv\frac{n+m}{a}+\sum_{i=1}^{n}\frac{\ln r_i}{1-r_i^a}+\frac{nr_n^a\ln r_n}{(1-r_n^a)\ln(1-r_n^a)}$$

$$+\sum_{j=1}^{m}\frac{\ln s_j}{1-s_j^a}+\frac{ms_m^a\ln s_m}{(1-s_m^a)\ln(1-s_m^a)}=0. \quad (4.113)$$

Proof 4.3: $G(a)$ can be rewritten as

$$G(a)=\frac{n}{a}\left[1+G_1(a)+\frac{G_2(a)}{G_3(a)}\right]+\frac{m}{a}\left[1+H_1(a)+\frac{H_2(a)}{H_3(a)}\right], \quad (4.114)$$

where

$$G_1(a)=\frac{1}{n}\sum_{i=1}^{n}\frac{\ln v_i}{1-v_i}, G_2(a)=\frac{1}{n}\frac{v_n\ln v_n}{1-v_n}, G_3(a)=\frac{1}{n}\ln(1-v_n), \quad (4.115)$$

$$H_1(a)=\frac{1}{m}\sum_{j=1}^{m}\frac{\ln w_j}{1-w_j}, H_2(a)=\frac{1}{m}\frac{w_m\ln w_m}{1-w_m}, H_3(a)=\frac{1}{m}\ln(1-w_m), \quad (4.116)$$

$v_i=r_i^a$, $i=1,\dots,n$ and $w_j=s_j^a$, $j=1,\dots,m$. The limit of $G(a)$ is considered as $a\to 0^+$ and $a\to\infty$. It is obtained that $\lim_{a\to 0^+}G(a)=\infty$ and $\lim_{a\to\infty}G(a)<0$. By the intermediate value theorem $G(a)$ has at least one root in $(0,\infty)$. If it can be shown that $G^{'}(a)<0$, then the proof will be completed. Since $r_i<r_n$, $1/(1-r_n^a)>1/(1-r_i^a)$, $i=1,\dots,n-1$ and $s_j<s_m$, $1/(1-s_m^a)>1/(1-s_j^a)$, $j=1,\dots,m-1$ for $a>0$,

$$G^{'}(a)<\frac{-(n+m)}{a^2}+\frac{nr_n^a}{a^2}\left(\frac{\ln r_n^a}{1-r_n^a}\right)^2\left[1+\frac{r_n^a+\ln(1-r_n^a)}{(\ln(1-r_n^a))^2}\right]$$

$$+\frac{ms_m^a}{a^2}\left(\frac{\ln s_m^a}{1-s_m^a}\right)^2\left[1+\frac{s_m^a+\ln(1-s_m^a)}{(\ln(1-s_m^a))^2}\right]$$

$$=\frac{-(n+m)}{a^2}+\frac{nv_n}{a^2}\left(\frac{\ln v_n}{1-v_n}\right)^2\left[1+\frac{v_n+\ln(1-v_n)}{(\ln(1-v_n))^2}\right] \quad (4.117)$$

$$+\frac{mw_m}{a^2}\left(\frac{\ln w_m}{1-w_m}\right)^2\left[1+\frac{w_m+\ln(1-w_m)}{(\ln(1-w_m))^2}\right]$$

$$=\frac{n}{a^2}h(v_n)+\frac{m}{a^2}h(w_m),$$

where

$$h(x) = -1 + x\left(\frac{\ln x}{1-x}\right)^2\left(1 + \frac{x + \ln(1-x)}{(\ln(1-x))^2}\right),\ 0 < x < 1. \tag{4.118}$$

It can be easily shown that $h(x)$ *is a monotone increasing function and* $h(x) < 0$ *for all* $0 < x < 1$. *Hence,* $G^{'}(a) < 0$ *is obtained.*

Finally, we will show that the MLEs of (b_1, b_2, a) *maximizes the log-likelihood function* $l(b_1, b_2, a\,|\underline{r}, \underline{s})$. *Let* $H(b_1, b_2, a)$ *be the Hessian matrix of* $l(b_1, b_2, a\,|\underline{r}, \underline{s})$ *at* (b_1, b_2, a). *It is clear that if* $\det(H) \neq 0$ *for the critical point* (b_1, b_2, a) *and* $\det(H_1) < 0$, $\det(H_2) > 0$, $\det(H_3) < 0$ *at* (b_1, b_2, a) *then it is a local maximum of* $l(b_1, b_2, a\,|\underline{r}, \underline{s})$, *where*

$$H_1 = \frac{\partial^2 l}{\partial b_1^2},\ H_2 = \begin{pmatrix} \frac{\partial^2 l}{\partial b_1^2} & \frac{\partial^2 l}{\partial b_1 \partial b_2} \\ \frac{\partial^2 l}{\partial b_2 \partial b_1} & \frac{\partial^2 l}{\partial b_2^2} \end{pmatrix},\ H_3 = H \text{ and } l = l(b_1, b_2, a\,|\underline{r}, \underline{s}). \tag{4.119}$$

It can be easily seen that

$$\det(H_1(\widehat{b}_1, \widehat{b}_2, \widehat{a})) = \frac{-\left(\ln(1 - r_n^{\widehat{a}})\right)^2}{n} < 0, \tag{4.120}$$

$$\det(H_2(\widehat{b}_1, \widehat{b}_2, \widehat{a})) = \frac{\left(\ln(1 - r_n^{\widehat{a}})\right)^2}{n}\frac{\left(\ln(1 - s_m^{\widehat{a}})\right)^2}{m} > 0, \tag{4.121}$$

and

$$\det(H(\widehat{b}_1, \widehat{b}_2, \widehat{a})) = G^{'}(\widehat{a})\frac{\left(\ln(1 - r_n^{\widehat{a}})\right)^2}{n}\frac{\left(\ln(1 - s_m^{\widehat{a}})\right)^2}{m} < 0. \tag{4.122}$$

Hence, $(\widehat{b}_1, \widehat{b}_2, \widehat{a})$ *is the local maximum of* $l(b_1, b_2, a\,|\underline{r}, \underline{s})$. *Since there is no singular point of* $l(b_1, b_2, a\,|\underline{r}, \underline{s})$ *and it has a single critical point then, it is enough to show that the absolute maximum of the function is indeed the local maximum. Assume that there exist an* $\widehat{a}_0$ *in the domain in which* $l^*(\widehat{a}_0) > l^*(\widehat{a})$, *where* $l^*(\widehat{a}) = l(\widehat{b}_1, \widehat{b}_2, \widehat{a}\,|\underline{r}, \underline{s})$. *Since* $\widehat{a}$ *is the local maximum there should be some point* a_1 *in the neighborhood of* $\widehat{a}$ *such that* $l^*(a_1) < l^*(\widehat{a})$. *Let* $K(a) = l^*(a) - l^*(\widehat{a})$ *then* $K(\widehat{a}_0) > 0$, $K(a_1) < 0$ *and* $K(\widehat{a}) = 0$. *This implies that* a_1 *is a local minimum of the* $l^*(a)$, *but* $\widehat{a}$ *is the only critical point so it is a contradiction. Therefore,* $(\widehat{b}_1, \widehat{b}_2, \widehat{a})$ *is the absolute maximum of* $l(b_1, b_2, a\,|\underline{r}, \underline{s})$. ■

4.4.1.2. Asymptotic Distribution and Confidence Intervals For R

The Fisher information matrix of (b_1, b_2, a) as $I \equiv I(b_1, b_2, a)$ is given by

$$I = -\begin{pmatrix} E\left(\frac{\partial^2 l}{\partial b_1^2}\right) & E\left(\frac{\partial^2 l}{\partial b_1 \partial b_2}\right) & E\left(\frac{\partial^2 l}{\partial b_1 \partial a}\right) \\ E\left(\frac{\partial^2 l}{\partial b_2 \partial b_1}\right) & E\left(\frac{\partial^2 l}{\partial b_2^2}\right) & E\left(\frac{\partial^2 l}{\partial b_2 \partial a}\right) \\ E\left(\frac{\partial^2 l}{\partial a \partial b_1}\right) & E\left(\frac{\partial^2 l}{\partial a \partial b_2}\right) & E\left(\frac{\partial^2 l}{\partial a^2}\right) \end{pmatrix} = \begin{pmatrix} I_{11} & I_{12} & I_{13} \\ I_{21} & I_{22} & I_{23} \\ I_{31} & I_{32} & I_{33} \end{pmatrix}, \tag{4.123}$$

where $I_{11} = n/b_1^2$, $I_{22} = m/b_2^2$, $I_{12} = I_{21} = 0$, $I_{13} = I_{31}$, $I_{23} = I_{32}$

$$I_{13} = \int_0^1 \frac{r_n^a \ln r_n}{1 - r_n^a} f_{R_n}(r_n) dr_n, \; I_{23} = \int_0^1 \frac{s_m^a \ln s_m}{1 - s_m^a} g_{S_m}(s_m) ds_m, \tag{4.124}$$

where $f_{R_n}(r_n)$ is a pdf of nth upper record value from $Kum(a, b_1)$ and $g_{S_m}(s_m)$ is a pdf of mth upper record value from $Kum(a, b_2)$,

$$I_{33} = -\sum_{i=1}^{n} \int_0^1 r_i^a (\frac{\ln r_i}{1 - r_i^a})^2 f_{R_i}(r_i) dr_i - \sum_{j=1}^{m} \int_0^1 s_j^a (\frac{\ln s_j}{1 - s_j^a})^2 g_{S_j}(s_j) ds_j + \frac{n+m}{a^2}$$

$$+ b_1 \int_0^1 r_n^a (\frac{\ln r_n}{1 - r_n^a})^2 f_{R_n}(r_n) dr_n + b_2 \int_0^1 s_m^a (\frac{\ln s_m}{1 - s_m^a})^2 g_{S_m}(s_m) ds_m, \tag{4.125}$$

where $f_{R_i}(r_i)$ is a pdf of ith upper record value from $Kum(a, b_1)$ and $g_{S_j}(s_j)$ is a pdf of jth upper record value from $Kum(a, b_2)$. After making suitable transformations, it is obtained that

$$I_{13} = \frac{b_1^n}{a} \sum_{i=1}^{\infty} \frac{1}{i} \left[\frac{1}{(b_1 + i)^n} - \frac{1}{(b_1 + i - 1)^n} \right], \tag{4.126}$$

$$I_{32} = \frac{b_2^m}{a} \sum_{j=1}^{\infty} \frac{1}{j} \left[\frac{1}{(b_2 + j)^m} - \frac{1}{(b_2 + j - 1)^m} \right], \tag{4.127}$$

and

$$I_{33} = \frac{n+m}{a^2} - \frac{2}{a^2} \left[\sum_{i=1}^{n} b_1^i A_i(b_1) - b_1^{n+1} A_n(b_1) \right.$$

$$+\sum_{j=1}^{m} b_2^j B_j(b_2) - b_2^{m+1} B_m(b_2)\Bigg], \tag{4.128}$$

where

$$A_i(b_1) = \left[\sum_{k=1}^{\infty} \frac{1}{k+1}\left(\frac{1}{(b_1+k-1)^i} - \frac{1}{(b_1+k)^i}\right)\left(\sum_{q=1}^{k} \frac{1}{q}\right)\right], \tag{4.129}$$

and

$$B_j(b_2) = \left[\sum_{k=1}^{\infty} \frac{1}{k+1}\left(\frac{1}{(b_2+k-1)^j} - \frac{1}{(b_2+k)^j}\right)\left(\sum_{q=1}^{k} \frac{1}{q}\right)\right], \tag{4.130}$$

see [Gradshteyn and Ryzhik, 1994] (formula 1.516(1), 4.272(6)).

Theorem 4.4: As $n \to \infty$ and $m \to \infty$ and $n/m \to p$ then

$$\left[\sqrt{n}(\widehat{b}_1 - b_1), \sqrt{m}(\widehat{b}_2 - b_2), \sqrt{n}(\widehat{a} - a)\right] \to N_3(0, U^{-1}(b_1, b_2, a)), \tag{4.131}$$

where

$$U(b_1, b_2, a) = \begin{pmatrix} u_{11} & 0 & u_{13} \\ 0 & u_{22} & u_{23} \\ u_{31} & u_{32} & u_{33} \end{pmatrix}, \tag{4.132}$$

and $u_{11} = \lim_{n,m\to\infty}(I_{11}/n)$, $u_{13} = u_{31} = \lim_{n,m\to\infty}(I_{13}/n)$, $u_{22} = \lim_{n,m\to\infty}(I_{22}/m)$, $u_{23} = u_{32} = \lim_{n,m\to\infty}(\sqrt{p}I_{23}/n)$, $u_{33} = \lim_{n,m\to\infty}(I_{33}/n)$.

Proof 4.4: The proof follows from the asymptotic normality of MLE. ■

Theorem 4.5: As $n \to \infty$ and $m \to \infty$ and $n/m \to p$ then

$$\sqrt{n}(\widehat{R} - R) \to N(0, \sigma^2), \tag{4.133}$$

where

$$\sigma^2 = \frac{b_1^2 p(u_{11}u_{33} - u_{13}^2) - 2b_1 b_2 \sqrt{p} u_{13} u_{23} + b_2^2(u_{22}u_{33} - u_{23}^2)}{k(b_1+b_2)^4}, \tag{4.134}$$

and $k = u_{11}u_{22}u_{33} - u_{11}u_{23}u_{32} - u_{13}u_{22}u_{31}$.

Proof 4.5: $\sqrt{n}\widehat{R}$ *is asymptotically normal with mean* $\sqrt{n}R$ *and variance*

$$\sigma^2 = \lim_{n,m\to\infty} n \sum_{j=1}^{3}\sum_{i=1}^{3} \frac{\partial R}{\partial b_i}\frac{\partial R}{\partial b_j} I_{ij}^{-1}, \tag{4.135}$$

where I_{ij}^{-1} *is the* (i,j) *th element of the inverse of the I, see [Rao, 1965]. Since* $\partial R/\partial b_3 = \partial R/\partial a = 0$,

$$\sigma^2 = \lim_{n,m\to\infty} n \left[\frac{\partial R}{\partial b_1}\frac{\partial R}{\partial b_1} I_{11}^{-1} + \frac{\partial R}{\partial b_2}\frac{\partial R}{\partial b_1}\left(I_{21}^{-1} + I_{12}^{-1}\right) + \frac{\partial R}{\partial b_2}\frac{\partial R}{\partial b_2} I_{22}^{-1}\right]$$

$$= \lim_{n,m\to\infty} n \left[\frac{b_1^2(I_{11}I_{33} - I_{13}^2) - 2b_1b_2I_{13}I_{23} + b_2^2(I_{22}I_{33} - I_{23}^2)}{(b_1+b_2)^4(I_{11}I_{22}I_{33} - I_{11}I_{23}^2 - I_{22}I_{13}^2)}\right]. \tag{4.136}$$

When this expression is multiplied by $\frac{1}{n^2m}n^2m$ *a suitable form is obtained, considering* $n/m \to p$ *as* $n \to \infty$ *and* $m \to \infty$, *the desired result is obtained.* ■

Remark 4.1: Theorem 4.5 can be used to construct the asymptotic confidence interval of R. *The variance* σ^2 *needs to be estimated to compute the confidence interval of* R. *The empirical Fisher information matrix and the MLEs of* b_1, b_2 *and* a *are used to estimate* σ^2 *as follows* $\widehat{u}_{11} = 1/\widehat{b}_1^2$, $\widehat{u}_{22} = 1/\widehat{b}_2^2$

$$\widehat{u}_{13} = \frac{\widehat{b}_1^n}{n\widehat{a}} \sum_{i=1}^{\infty} \frac{1}{i}\left[\frac{1}{(\widehat{b}_1 + i)^n} - \frac{1}{(\widehat{b}_1 + i - 1)^n}\right], \tag{4.137}$$

$$\widehat{u}_{23} = \frac{\sqrt{p}}{n}\frac{\widehat{b}_2^m}{\widehat{a}} \sum_{j=1}^{\infty} \frac{1}{j}\left[\frac{1}{(\widehat{b}_2 + j)^m} - \frac{1}{(\widehat{b}_2 + j - 1)^m}\right], \tag{4.138}$$

$$\widehat{u}_{33} = \frac{n+m}{n\widehat{a}^2} - \frac{2}{n\widehat{a}^2}\left[\sum_{i=1}^{n} \widehat{b}_1^i A_i(\widehat{b}_1) - \widehat{b}_1^{n+1}A_n(\widehat{b}_1)\right.$$

$$\left. + \sum_{j=1}^{m} \widehat{b}_2^j B_j(\widehat{b}_2) - \widehat{b}_2^{m+1}B_m(\widehat{b}_2)\right]. \tag{4.139}$$

4.4.1.3. Bayes Estimation of R

Assume that all parameters $b_1,\ b_2$ and a are unknown and have independent gamma priors with parameters $(\alpha_i, \beta_i),\ i = 1, 2, 3$, respectively. The density function of a gamma random variable X with the shape and scale parameters α and β, respectively, is

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\beta},\ x > 0,\ \alpha, \beta > 0. \tag{4.140}$$

The joint prior density function of b_1, b_2 and a is $\pi(b_1, b_2, a) = \pi(b_1)\pi(b_2)\pi(a)$, and the joint posterior density function of b_1, b_2 and a given $(\underline{r}, \underline{s})$ is

$$\pi\left(b_1, b_2, a \left| \underline{r}, \underline{s}\right.\right) = \frac{h_1(\underline{r}; a) h_2(\underline{s}; a) b_1^{n+\alpha_1-1} b_2^{m+\alpha_2-1} a^{n+m+\alpha_3-1}}{\Gamma(n+\alpha_1)\Gamma(m+\alpha_2) I_0(\underline{r}, \underline{s})}$$

$$\exp\left\{-b_1(\beta_1 + T_1(r_n; a)) - b_2(\beta_2 + T_2(s_m; a)) - a\beta_3\right\}, \tag{4.141}$$

where

$$I_0(\underline{r}, \underline{s}) = \int_0^\infty \frac{a^{n+m+\alpha_3-1} h_1(\underline{r}; a) h_2(\underline{s}; a) e^{-a\beta_3}}{(\beta_1 + T_1(r_n; a))^{n+\alpha_1} (\beta_2 + T_2(s_m; a))^{m+\alpha_2}} da. \tag{4.142}$$

Then, the Bayes estimate of a given measurable function of b_1, b_2 and a, say $u(b_1, b_2, a)$ under the SE loss function is

$$\widehat{u}_B = \int_0^\infty \int_0^\infty \int_0^\infty u(b_1, b_2, a) \pi(b_1, b_2, a \left| \underline{r}, \underline{s}\right.) db_1 db_2 da$$

$$= \frac{\int_0^\infty \int_0^\infty \int_0^\infty u(b_1, b_2, a) L(b_1, b_2, a \left| \underline{r}, \underline{s}\right.) \pi(b_1, b_2, a) db_1 db_2 da}{\int_0^\infty \int_0^\infty \int_0^\infty L(b_1, b_2, a \left| \underline{r}, \underline{s}\right.) \pi(b_1, b_2, a) db_1 db_2 da}. \tag{4.143}$$

The ratio of two integrals equation (4.143) cannot be solved analytically. We may use a numerical integration method to calculate the integrals or use approximate methods such as the approximate form due to Lindley [Lindley, 1980] or that of Tierney and Kadane [Tierney and Kadane, 1986]. Lindley has proposed approximations for moments that capture the first-order error terms of the normal approximation. This is generally accurate enough, but, as Lindley points out, the required evaluation of third derivatives of the posterior can be rather tedious, especially, in problems with several parameters. Moreover, the error of Tierney and Kadane's approximate is of the order $O(n^{-2})$ while the error in using Lindley's approximate form is of the order $O(n^{-1})$. Therefore, the Tierney-Kadane approximation is preferred for our case. The regularity condition required for using Tierney-Kadane's form is that the posterior density function should be unimodal.

To show that the posterior density function is unimodal, it suffices to show that the function $Q(b_1, b_2, a) \equiv \ln \pi\left(b_1, b_2, a\,|\underline{r}, \underline{s}\right)$ has the unique mode. The extremum points of $Q(b_1, b_2, a)$ are given by

$$\widetilde{b}_1 = \frac{q_1}{\beta_1 + T_1(r_n; \widetilde{a})},\ \widetilde{b}_2 = \frac{q_2}{\beta_2 + T_2(s_m; \widetilde{a})}, \tag{4.144}$$

and $\widetilde{a}$ is the solution of the non-linear equation

$$P(a) = \frac{q_3}{a} - \frac{q_1 r_n^a \ln r_n/(1 - r_n^a)}{\beta_1 + T_1(r_n; a)} - \frac{q_2 s_m^a \ln s_m/(1 - s_m^a)}{\beta_2 + T_2(s_m; a)} - \beta_3 = 0. \tag{4.145}$$

$P(a)$ can be rewritten as

$$P(a) = \frac{1}{a}\left[q_3 - \frac{q_1 v_n \ln v_n/(1 - v_n)}{\beta_1 - \ln(1 - v_n)} - \frac{q_2 w_m \ln w_m/(1 - w_m)}{\beta_2 - \ln(1 - w_m)}\right] - \beta_{3,} \tag{4.146}$$

where $v_n = r_n^a$, $w_m = s_m^a$, $q_1 = n + \alpha_1 - 1$, $q_2 = m + \alpha_2 - 1$ and $q_3 = n + m + \alpha_3 - 1$. It is easily seen that $\lim_{a \to 0^+} P(a) = \infty$ and $\lim_{a \to \infty} P(a) < 0$. If it can be shown that $P(a)$ is monotone decreasing for all a, then the equation $P(a) = 0$ has a unique solution in $(0, \infty)$.

$$P^{'}(a) = -\frac{1}{a^2}\left[q_1 v_n \left(\frac{\ln v_n}{1 - v_n}\right)^2 \left\{\frac{1}{\beta_1 - \ln(1 - v_n)} - \frac{v_n}{(\beta_1 - \ln(1 - v_n))^2}\right\}\right.$$

$$\left. + q_3 + q_2 w_m \left(\frac{\ln w_m}{1 - w_m}\right)^2 \left\{\frac{1}{\beta_2 - \ln(1 - w_m)} - \frac{w_m}{(\beta_2 - \ln(1 - w_m))^2}\right\}\right] \tag{4.147}$$

$$= -\frac{1}{a^2}\left[q_3 + q_1 h_1(v_n) + q_2 h_1(w_m)\right],$$

where

$$h_1(x) = x\left(\frac{\ln x}{1-x}\right)^2 \left\{\frac{1}{\beta_1 - \ln(1-x)} - \frac{x}{(\beta_1 - \ln(1-x))^2}\right\}, \tag{4.148}$$

where $0 < x < 1$. Let $f_1(x) = \beta_1 - \ln(1-x) - x$, then $f_1(0) > 0$ and $f_1(x)$ is a monotone increasing function for all $0 < x < 1$. It can be easily shown that $h_1(x) > 0$ for all $0 < x < 1$, by noticing $h_1(x) = x\,(\ln x/(1-x))^2\,(f_1(x)/(\beta_1 - \ln(1-x))^2)$. Hence, $P^{'}(a) < 0$ is obtained. Now, we want to show that the function $Q(b_1, b_2, a)$ is the maximum at the point $(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})$. Let $H^*(b_1, b_2, a)$ be the Hessian matrix of $Q(b_1, b_2, a)$. We obtain that

$$\det(H_1^*(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})) = -\frac{(\beta_1 - \ln(1 - r_n^{\widetilde{a}}))^2}{n + \alpha_1 - 1} < 0, \tag{4.149}$$

$$\det(H_2^*(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})) = \frac{(\beta_1 - \ln(1 - r_n^{\widetilde{a}}))^2}{n + \alpha_1 - 1}\frac{(\beta_2 - \ln(1 - s_m^{\widetilde{a}}))^2}{m + \alpha_2 - 1} > 0, \tag{4.150}$$

and

$$\det(H^*(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})) = P^{'}(\widetilde{a})\frac{(\beta_1 - \ln(1 - r_n^{\widetilde{a}}))^2}{n + \alpha_1 - 1}\frac{(\beta_2 - \ln(1 - s_m^{\widetilde{a}}))^2}{m + \alpha_2 - 1} < 0. \tag{4.151}$$

Therefore, $Q(b_1, b_2, a)$ has unique mode and so the posterior density function is unimodal. Consequently, Tierney and Kadane's approximation can be applied to our case.

The posterior mean of the $u(b_1, b_2, a)$, equation (4.143), can be rewritten as

$$E\left[u(b_1, b_2, a)\,|\underline{r}, \underline{s}\right] = \frac{\int_0^\infty \int_0^\infty \int_0^\infty e^{n\Lambda^*(b_1,b_2,a)} db_1 db_2 da}{\int_0^\infty \int_0^\infty \int_0^\infty e^{n\Lambda(b_1,b_2,a)} db_1 db_2 da}, \tag{4.152}$$

where

$$\Lambda(b_1, b_2, a) = \frac{[\ln(L(b_1, b_2, a\,|\underline{r}, \underline{s})) + \ln(\pi(b_1, b_2, a))]}{n}, \tag{4.153}$$

and

$$\Lambda^*(b_1, b_2, a) = \Lambda(b_1, b_2, a) + \frac{1}{n}\ln(u(b_1, b_2, a)). \tag{4.154}$$

Following the [Tierney and Kadane, 1986], equation (4.152) can be approximated in the form

$$\widehat{u}_{BT}(b_1, b_2, a) = \left[\frac{\det \Sigma^*}{\det \Sigma}\right]^{1/2} \exp\left(n\left[\Lambda^*(\widetilde{b}_1^*, \widetilde{b}_2^*, \widetilde{a}^*) - \Lambda(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})\right]\right), \tag{4.155}$$

where $(\widetilde{b}_1^*, \widetilde{b}_2^*, \widetilde{a}^*)$ and $(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})$ maximize $\Lambda^*(b_1, b_2, a)$ and $\Lambda(b_1, b_2, a)$, respectively, and Σ^* and Σ are the negatives of the inverse Hessians of $\Lambda^*(b_1, b_2, a)$ and $\Lambda(b_1, b_2, a)$ at $(\widetilde{b}_1^*, \widetilde{b}_2^*, \widetilde{a}^*)$ and $(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})$, respectively.

In our case, we have

$$\Lambda(b_1, b_2, a) = \frac{1}{n}\left[l(b_1, b_2, a\,|\underline{r}, \underline{s}) + \ln C + (\alpha_1 - 1)\ln b_1 + (\alpha_2 - 1)\ln b_2\right.$$

$$\left. + (\alpha_3 - 1)\ln a - b_1\beta_1 - b_2\beta_2 - a\beta_3\right], \tag{4.156}$$

where $C = \beta_1^{\alpha_1}\beta_2^{\alpha_2}\beta_3^{\alpha_3}/(\Gamma(\alpha_1)\Gamma(\alpha_2)\Gamma(\alpha_3))$. $(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})$ can be obtained by solving the following equations

$$\Lambda_1 = \frac{\partial\Lambda(b_1, b_2, a)}{\partial b_1} = 0,\ \Lambda_2 = \frac{\partial\Lambda(b_1, b_2, a)}{\partial b_2} = 0,\ \Lambda_3 = \frac{\partial\Lambda(b_1, b_2, a)}{\partial a} = 0, \tag{4.157}$$

and are given by

$$\widetilde{b}_1 = \frac{n + \alpha_1 - 1}{\beta_1 + T_1(r_n; \widetilde{a})}, \widetilde{b}_2 = \frac{m + \alpha_2 - 1}{\beta_2 + T_2(s_m; \widetilde{a})}, \tag{4.158}$$

and $\widetilde{a}$ is the solution of the non-linear equation

$$\frac{q_3}{a} + \sum_{i=1}^{n}\frac{\ln r_i}{1 - r_i^a} - \frac{q_1 r_n^a \ln r_n/(1 - r_n^a)}{\beta_1 + T_1(r_n; a)}$$

$$+ \sum_{j=1}^{m}\frac{\ln s_j}{1 - s_j^a} - \frac{q_2 s_m^a \ln s_m/(1 - s_m^a)}{\beta_2 + T_2(s_m; a)} - \beta_3 = 0. \tag{4.159}$$

The fixed point method is applied as in the MLE of a. The units of the Hessian matrix of $\Lambda(b_1, b_2, a)$ are obtained as

$$\Lambda_{11} = \frac{\partial^2 \Lambda(b_1, b_2, a)}{\partial b_1^2} = \frac{-q_1}{n b_1^2},\ \Lambda_{22} = \frac{\partial^2 \Lambda(b_1, b_2, a)}{\partial b_2^2} = \frac{-q_2}{n b_2^2}, \tag{4.160}$$

$$\Lambda_{12} = \Lambda_{21} = 0,\ \Lambda_{13} = \Lambda_{31} = \frac{\partial^2 \Lambda(b_1, b_2, a)}{\partial b_1 \partial a} = \frac{-r_n^a \ln r_n}{n(1 - r_n^a)}, \tag{4.161}$$

$$\Lambda_{23} = \Lambda_{32} = \frac{\partial^2 \Lambda(b_1, b_2, a)}{\partial b_2 \partial a} = \frac{-s_m^a \ln s_m}{n(1 - s_m^a)}, \tag{4.162}$$

$$\Lambda_{33} = \frac{\partial^2 \Lambda(b_1, b_2, a)}{\partial a^2} = \frac{1}{n}\left[\frac{-q_3}{a^2} + \sum_{i=1}^{n} r_i^a \left(\frac{\ln r_i}{1 - r_i^a}\right)^2 - b_1 r_n^a \left(\frac{\ln r_n}{1 - r_n^a}\right)^2 \right.$$
$$\left. + \sum_{j=1}^{m} s_j^a \left(\frac{\ln s_j}{1 - s_j^a}\right)^2 - b_2 s_m^a \left(\frac{\ln s_m}{1 - s_m^a}\right)^2\right]. \tag{4.163}$$

Hence,

$$\Sigma = -\begin{pmatrix} \Lambda_{11} & 0 & \Lambda_{13} \\ 0 & \Lambda_{22} & \Lambda_{23} \\ \Lambda_{13} & \Lambda_{23} & \Lambda_{33} \end{pmatrix}^{-1}, \tag{4.164}$$

and the determinant of Σ is evaluated at $(\widetilde{b}_1, \widetilde{b}_2, \widetilde{a})$.

The Bayes estimate of R under the SE loss function is obtained by using $u(b_1, b_2, a) = R$. Equation (4.154) takes the form

$$_{BS}\Lambda^*(b_1, b_2, a) = \Lambda(b_1, b_2, a) + \frac{1}{n} \ln R. \tag{4.165}$$

The maximum value of the function $_{BS}\Lambda^*(b_1, b_2, a)$, say at $(_{BS}\widetilde{b}_1^*, _{BS}\widetilde{b}_2^*, _{BS}\widetilde{a}^*)$, is a solution of the non-linear equation system

$$\frac{n + \alpha_1 - 1}{b_1} - \beta_1 - T_1(r_n; a) + \frac{b_2}{b_1(b_1 + b_2)} = 0, \tag{4.166}$$

$$\frac{m + \alpha_2 - 1}{b_2} - \beta_2 - T_2(s_m; a) + \frac{1}{b_1 + b_2} = 0, \tag{4.167}$$

and

$$\frac{q_3}{a}+\sum_{i=1}^{n}\frac{\ln r_i}{1-r_i^a}+\sum_{j=1}^{m}\frac{\ln s_j}{1-s_j^a}-\frac{b_1 r_n^a\ln r_n}{1-r_n^a}-\frac{b_2 s_m^a\ln s_m}{1-s_m^a}-\beta_3=0. \tag{4.168}$$

The solution of the system can be obtained by using the fixed point method. The Hessian matrix of ${}_{BS}\Lambda^*(b_1,b_2,a)$ can be computed following the same arguments as in the first case. Therefore, the value of $\det({}_{BS}\Sigma^*)$ at $({}_{BS}\widetilde{b}_1^*,{}_{BS}\widetilde{b}_2^*,{}_{BS}\widetilde{a}^*)$ is obtained. The Bayes estimate of R under the SE loss function is obtained by using equation (4.155) and is given by

$$\widehat{R}_{BS}=\left[\frac{\det_{BS}\Sigma^*}{\det\Sigma}\right]^{1/2}\exp\left(n\left[{}_{BS}\Lambda^*({}_{BS}\widetilde{b}_1^*,{}_{BS}\widetilde{b}_2^*,{}_{BS}\widetilde{a}^*)-\Lambda(\widetilde{b}_1,\widetilde{b}_2,\widetilde{a})\right]\right). \tag{4.169}$$

If we choose $u(b_1,b_2,a)=e^{-vR}$, the Bayes estimate of R is obtained under the LINEX loss function. Similar to the SE loss function case, we get

$${}_{BL}\Lambda^*(b_1,b_2,a)=\Lambda(b_1,b_2,a)-\frac{vR}{n}, \tag{4.170}$$

from equation (4.154). The maximum value of the function ${}_{BL}\Lambda^*(b_1,b_2,a)$, say at $({}_{BL}\widetilde{b}_1^*,{}_{BL}\widetilde{b}_2^*,{}_{BL}\widetilde{a}^*)$, is a solution of the non-linear equation system

$$\frac{n+\alpha_1-1}{b_1}-\beta_1-T_1(r_n;a)-\frac{vb_2}{(b_1+b_2)^2} \;=\; 0, \tag{4.171}$$

$$\frac{m+\alpha_2-1}{b_2}-\beta_2-T_2(s_m;a)+\frac{vb_1}{(b_1+b_2)^2} \;=\; 0, \tag{4.172}$$

and

$$\frac{q_3}{a}+\sum_{i=1}^{n}\frac{\ln r_i}{1-r_i^a}+\sum_{j=1}^{m}\frac{\ln s_j}{1-s_j^a}-\frac{b_1 r_n^a\ln r_n}{1-r_n^a}-\frac{b_2 s_m^a\ln s_m}{1-s_m^a}-\beta_3=0. \tag{4.173}$$

The Bayes estimate of R under the LINEX loss function is obtained by using equation (4.155) and is given by

$$\widehat{R}_{BL}=\left[\frac{\det_{BL}\Sigma^*}{\det\Sigma}\right]^{1/2}\exp\left(n\left[{}_{BL}\Lambda^*({}_{BL}\widetilde{b}_1^*,{}_{BL}\widetilde{b}_2^*,{}_{BL}\widetilde{a}^*)-\Lambda(\widetilde{b}_1,\widetilde{b}_2,\widetilde{a})\right]\right). \tag{4.174}$$

4.4.2. Estimation of R When a Is Common and Known

The estimation of R is considered when the parameter a is assumed to be known, say $a = 1$. Let $R_1, \ldots, R_n$ be a set of upper records from $Kum(1, b_1)$ and $S_1, \ldots, S_m$ be an independent set of upper records from $Kum(1, b_2)$.

4.4.2.1. ML Estimation and Confidence Intervals of R

Based on the above samples, the MLE of R, say $\widehat{R}_{MLE}$, is

$$\widehat{R}_{MLE} = \frac{\widehat{b}_1}{\widehat{b}_1 + \widehat{b}_2} = \frac{n\ln(1-s_m)}{n\ln(1-s_m) + m\ln(1-r_n)}. \tag{4.175}$$

It is easy to see that $-2b_1\ln(1-r_n) \sim \chi^2(2n)$ and $-2b_2\ln(1-s_m) \sim \chi^2(2m)$. Therefore,

$$F^* = \left(\frac{R}{1-R}\right)\left(\frac{1-\widehat{R}_{MLE}}{\widehat{R}_{MLE}}\right), \tag{4.176}$$

is an F distributed random variable with $(2n, 2m)$ degrees of freedom. The pdf of $\widehat{R}_{MLE}$ is as follows;

$$f_{\widehat{R}_{MLE}}(r) = \frac{1}{r^2B(m,n)}\left(\frac{nb_1}{mb_2}\right)^n \frac{\left(\frac{1-r}{r}\right)^{n-1}}{\left(1+\frac{nb_1(1-r)}{mb_2 r}\right)^{n+m}}, \tag{4.177}$$

where $0 < r < 1$. The $100(1-\alpha)\%$ confidence interval for R can be obtained as

$$\left(\frac{1}{1+F_{2m,2n;\frac{\alpha}{2}}\left(\frac{1-\widehat{R}_{MLE}}{\widehat{R}_{MLE}}\right)}, \frac{1}{1+F_{2m,2n;1-\frac{\alpha}{2}}\left(\frac{1-\widehat{R}_{MLE}}{\widehat{R}_{MLE}}\right)}\right), \tag{4.178}$$

where $F_{2m,2n;\frac{\alpha}{2}}$ and $F_{2m,2n;1-\frac{\alpha}{2}}$ are the lower and upper $\frac{\alpha}{2}$th percentile points of a F distribution with $(2m, 2n)$ degrees of freedom.

On the other hand, the approximate confidence interval of R can be easily obtained by using the Fisher information matrix. The Fisher information matrix of

(b_1, b_2) is

$$I = -\begin{pmatrix} E\left(\frac{\partial^2 l}{\partial b_1^2}\right) & E\left(\frac{\partial^2 l}{\partial b_1 \partial b_2}\right) \\ E\left(\frac{\partial^2 l}{\partial b_2 \partial b_1}\right) & E\left(\frac{\partial^2 l}{\partial b_2^2}\right) \end{pmatrix} = \begin{pmatrix} n/b_1^2 & 0 \\ 0 & m/b_2^2 \end{pmatrix}. \tag{4.179}$$

By the asymptotic properties of the MLE, $\widehat{R}_{MLE}$ is approximately distributed as normal with mean R and variance

$$\sigma^2 = \sum_{j=1}^{2} \sum_{i=1}^{2} \frac{\partial R}{\partial b_i} \frac{\partial R}{\partial b_j} I_{ij}^{-1}, \tag{4.180}$$

where I_{ij}^{-1} is the (i, j) th element of the inverse of the I, see [Rao, 1965]. Then, an approximate $100(1-\alpha)\%$ confidence interval for R is

$$\left(\widehat{R}_{MLE} - cz_{\alpha/2}\widehat{R}_{MLE}(1-\widehat{R}_{MLE}), \widehat{R}_{MLE} + cz_{\alpha/2}\widehat{R}_{MLE}(1-\widehat{R}_{MLE})\right), \tag{4.181}$$

where $z_{\alpha/2}$ is the upper $\frac{\alpha}{2}$th percentile points of a standard normal distribution and $c = \sqrt{(1/n) + (1/m)}$.

4.4.2.2. UMVUE of R

When the first shape parameter $a = 1$, the joint likelihood function is

$$L(b_1, b_2, a\,|\underline{r}, \underline{s}) = h_1(\underline{r})h_2(\underline{s})b_1^n b_2^m e^{-b_1 T_1(r_n)} e^{-b_2 T_2(s_m)}, \tag{4.182}$$

where $h_1(\underline{r}) =_{i=1}^{n} 1/(1-r_i)$, $h_2(\underline{s}) =_{j=1}^{m} 1/(1-s_j)$, $T_1(r_n) = -\ln(1-r_n)$ and $T_2(s_m) = -\ln(1-s_m)$. It is clear that $(T_1(r_n), T_2(s_m))$ is a sufficient statistic for (b_1, b_2). It can be shown that it is also a complete sufficient statistic by using Theorem 10-9 in [Arnold, 1990]. Let us define

$$\phi(R_1, S_1) = \begin{cases} 1 & \text{if } R_1 < S_1 \\ 0 & \text{if } R_1 \geq S_1 \end{cases}. \tag{4.183}$$

Then, $E\left(\phi(R_1,S_1)\right)=R$ so it is an unbiased estimator of R. Let $P_1=-\ln(1-R_1)$ and $P_2=-\ln(1-S_1)$. The UMVUE of R, say $\widehat{R}_U$, can be obtained by using the Rao-Blackwell and the Lehmann-Scheffe's Theorems, (see [Arnold, 1990])

$$\widehat{R}_U=E\left(\phi(P_1,P_2)\left|(T_1,T_2)\right.\right)$$

$$=\int_{P_2}\int_{P_1}\phi(P_1,P_2)\,f(p_1,p_2\left|T_1,T_2\right.)dp_1dp_2 \tag{4.184}$$

$$=\int_{P_2}\int_{P_1}\phi(P_1,P_2)f_{P_1|T_1}(p_1\left|T_1\right.)f_{P_2|T_2}(p_2\left|T_2\right.)dp_1dp_2,$$

where $(T_1,T_2)=(T_1(r_n),T_2(s_m))$, $f(p_1,p_2\left|T_1,T_2\right.)$ is the conditional pdf of (P_1,P_2) given (T_1,T_2). Using the joint pdf of (R_1,R_n) and (S_1,S_m) and after making a simple transformation, we obtain the $f_{P_1|T_1}(p_1\left|T_1\right.)$ and $f_{P_2|T_2}(p_2\left|T_2\right.)$, and are given by

$$f_{P_1|T_1}(p_1\left|T_1\right.)=(n-1)\frac{(t_1-p_1)^{n-2}}{t_1^{n-1}},\ 0<p_1<t_1, \tag{4.185}$$

$$f_{P_2|T_2}(p_2\left|T_2\right.)=(m-1)\frac{(t_2-p_2)^{m-2}}{t_2^{m-1}},\ 0<p_2<t_2. \tag{4.186}$$

Therefore,

$$\widehat{R}_U=\iint\limits_{P_1<P_2}f_{P_1|T_1}(p_1\left|T_1\right.)f_{P_2|T_2}(p_2\left|T_2\right.)dp_1dp_2$$

$$=\begin{cases}\int_0^{t_1}\int_{p_1}^{t_2}\dfrac{(n-1)(m-1)(t_1-p_1)^{n-2}(t_2-p_2)^{m-2}}{t_1^{n-1}t_2^{m-1}}dp_2dp_1 & \text{if } t_2\geq t_1\\ \int_0^{t_2}\int_0^{p_2}\dfrac{(n-1)(m-1)(t_1-p_1)^{n-2}(t_2-p_2)^{m-2}}{t_1^{n-1}t_2^{m-1}}dp_2dp_1 & \text{if } t_2<t_1\end{cases} \tag{4.187}$$

$$=\begin{cases}{}_2F_1(1,1-m;n;t_1/t_2) & \text{if } t_2\geq t_1\\ 1-{}_2F_1(1,1-n;m;t_2/t_1) & \text{if } t_2<t_1\end{cases},$$

where ${}_2F_1(.,.;.;.)$ is Gauss hypergeometric function, (see formula 3.196(1) in [Gradshteyn and Ryzhik, 1994]).

4.4.2.3. Bayesian Estimation of R

Assume that the parameter b_1 and b_2 have independent gamma priors with the parameters (α_i, β_i), $i = 1, 2$. Then, the joint posterior density function of b_1 and b_2 given $(\underline{r}, \underline{s})$ is

$$\pi\left(b_1, b_2 \left|\underline{r}, \underline{s}\right.\right) = \frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{\Gamma(\delta_1)\Gamma(\delta_2)} b_1^{\delta_1 - 1} b_2^{\delta_2 - 1} e^{-b_1\lambda_1} e^{-b_2\lambda_2}, \tag{4.188}$$

where $\lambda_1 = \beta_1 + T_1(r_n)$, $\lambda_2 = \beta_2 + T_2(s_m)$, $\delta_1 = n + \alpha_1$, $\delta_2 = m + \alpha_2$. The posterior pdf of R can be obtained by using the joint posterior density function and is given by

$$f_R(r) = \frac{\lambda_1^{\delta_1}\lambda_2^{\delta_2}}{B(\delta_1, \delta_2)} \frac{r^{\delta_1 - 1}(1-r)^{\delta_2 - 1}}{(r\lambda_1 + (1-r)\lambda_2)^{\delta_1 + \delta_2}},\ 0 < r < 1. \tag{4.189}$$

After making a suitable transformations and simplifications by using formula 3.197(3) in [Gradshteyn and Ryzhik, 1994], the Bayes estimate of R, say $\widehat{R}_{BS}$, under the SE loss function is

$$\widehat{R}_{BS} = \begin{cases} c_1(\frac{\lambda_1}{\lambda_2})^{\delta_1}\ {}_2F_1(c_1^*, \delta_1 + 1; c_1^* + 1; 1 - \frac{\lambda_1}{\lambda_2}) & \text{if } \lambda_1 < \lambda_2 \\ c_1(\frac{\lambda_2}{\lambda_1})^{\delta_2}\ {}_2F_1(c_1^*, \delta_2; c_1^* + 1; 1 - \frac{\lambda_2}{\lambda_1}) & \text{if } \lambda_2 \le \lambda_1 \end{cases}, \tag{4.190}$$

where $c_1 = \delta_1 / c_1^*$, $c_1^* = \delta_1 + \delta_2$.

The Bayes estimate of R under the LINEX loss function, say $\widehat{R}_{BL}$, is $\widehat{R}_{BL} = \left\{-\ln E_R(e^{-vR})\right\}/v$, where $E_R(.)$ denotes posterior expectation with respect to the posterior density of R. It can be easily obtained that

$$E(e^{-vR}) = \begin{cases} (\frac{\lambda_1}{\lambda_2})^{\delta_1}\ \Phi_1(\delta_1, c_1^*, c_1^*, 1 - \frac{\lambda_1}{\lambda_2}, -v) & \text{if } \lambda_1 < \lambda_2 \\ (\frac{\lambda_2}{\lambda_1})^{\delta_2}\ e^{-v}\ \Phi_1(\delta_2, c_1^*, c_1^*, 1 - \frac{\lambda_2}{\lambda_1}, v) & \text{if } \lambda_2 \le \lambda_1 \end{cases}, \tag{4.191}$$

where $\Phi_1(.,.,.,.,.)$ is confluent hypergeometric series of two variables, (see formulas 3.385 and 9.261(1) in [Gradshteyn and Ryzhik, 1994]). Therefore,

$$\widehat{R}_{BL} = \begin{cases} -\frac{1}{v}\left(c_2 + \ln\left[\Phi_1(\delta_1, c_1^*, c_1^*, 1 - \frac{\lambda_1}{\lambda_2}, -v)\right]\right) & \text{if } \lambda_1 < \lambda_2 \\ -\frac{1}{v}\left(c_3 + \ln\left[\Phi_1(\delta_2, c_1^*, c_1^*, 1 - \frac{\lambda_2}{\lambda_1}, v)\right]\right) & \text{if } \lambda_2 \leq \lambda_1 \end{cases}, \tag{4.192}$$

where $c_2 = \delta_1 \ln(\lambda_1/\lambda_2)$ and $c_3 = \delta_2 \ln(\lambda_2/\lambda_1) - v$.

If we use the Jeffrey's non informative prior which is given by $\sqrt{\det I(b_1, b_2)}$, then the joint prior density function is $\pi(b_1, b_2) \propto 1/b_1 b_2$. Therefore, the joint posterior density function of b_1 and b_2 given $(\underline{r}, \underline{s})$ is

$$\pi(b_1, b_2 \,|\underline{r}, \underline{s}) = \frac{(T_1(r_n))^n \,(T_2(s_m))^m}{\Gamma(n)\Gamma(m)} b_1^{n-1} b_2^{m-1} e^{-b_1 T_1(r_n)} e^{-b_2 T_2(s_m)}, \tag{4.193}$$

and the posterior pdf of R is given by

$$f_R(r) = \frac{(T_1(r_n))^n \,(T_2(s_m))^m}{B(n,m)} \frac{r^{n-1}(1-r)^{m-1}}{(rT_1(r_n) + (1-r)T_2(s_m))^{n+m}}, \; 0 < r < 1. \tag{4.194}$$

The Bayes estimate of R under the SE and the LINEX loss function, say $\widehat{R}^*_{BS}$ and $\widehat{R}^*_{BL}$, respectively, are

$$\widehat{R}^*_{BS} = \begin{cases} c(\frac{T_1}{T_2})^n \; {}_2F_1(c^*, n+1; c^*+1; 1 - \frac{T_1}{T_2}) & \text{if } T_1 < T_2 \\ c(\frac{T_2}{T_1})^m \; {}_2F_1(c^*, m; c^*+1; 1 - \frac{T_2}{T_1}) & \text{if } T_2 \leq T_1 \end{cases}, \tag{4.195}$$

and

$$\widehat{R}^*_{BL} = \begin{cases} -\frac{1}{v}\left(c_4 + \ln\left[\Phi_1(n, c^*, c^*, 1 - \frac{T_1}{T_2}, -v)\right]\right) & \text{if } T_1 < T_2 \\ -\frac{1}{v}\left(c_5 + \ln\left[\Phi_1(m, c^*, c^*, 1 - \frac{T_2}{T_1}, v)\right]\right) & \text{if } T_2 \leq T_1 \end{cases}, \tag{4.196}$$

where $c = n/c^*$, $c^* = n + m$, $c_4 = n \ln(T_1/T_2)$, $c_5 = m \ln(T_2/T_1) - v$, $T_1 = T_1(r_n)$ and $T_2 = T_2(s_m)$.

Alternatively, the Bayes estimate of R under the SE and the LINEX loss functions can be obtained approxiametly by using the Lindley's approximation. The approximate Bayes estimate of R under the SE and LINEX loss functions, say $\widehat{R}_{BS,Lindley}$ and $\widehat{R}_{BL,Lindley}$, respectively, are

$$\widehat{R}_{BS,Lindley} = \widetilde{R}\left(1 + \frac{(1-\widetilde{R})^2}{n+\alpha_1-1} - \frac{\widetilde{R}(1-\widetilde{R})}{m+\alpha_2-1}\right), \tag{4.197}$$

and

$$\widehat{R}_{BL,Lindley} = \widetilde{R} - \frac{1}{v}\ln\left(1 + \frac{\widetilde{R}_1(1-\widetilde{R})(v\widetilde{R}-2)}{2(n+\alpha_1-1)} + \frac{\widetilde{R}\widetilde{R}_1(v-v\widetilde{R}+2)}{2(m+\alpha_2-1)}\right), \tag{4.198}$$

where $\widetilde{R} = \widetilde{b}_1/(\widetilde{b}_1+\widetilde{b}_2)$, $\widetilde{R}_1 = v\widetilde{R}(1-\widetilde{R})$, $\widetilde{b}_1 = (n+\alpha_1-1)/(\beta_1+T_1(r_n))$ and $\widetilde{b}_2 = (m+\alpha_2-1)/(\beta_2+T_2(s_m))$.

If we use the Jeffrey's non informative prior, then the approximate Bayes estimate of R under the SE and the LINEX loss functions, say $\widehat{R}^*_{BS,Lindley}$ and $\widehat{R}^*_{BL,Lindley}$, respectively, are

$$\widehat{R}^*_{BS,Lindley} = \widetilde{R}\left(1 + \frac{(1-\widetilde{R})^2}{n-1} - \frac{\widetilde{R}(1-\widetilde{R})}{m-1}\right), \tag{4.199}$$

and

$$\widehat{R}^*_{BL,Lindley} = \widetilde{R} - \frac{1}{v}\ln 1\left(+\frac{\widetilde{R}_1(1-\widetilde{R})(v\widetilde{R}-2)}{2(n-1)} + \frac{\widetilde{R}\widetilde{R}_1(v-v\widetilde{R}+2)}{2(m-1)}\right), \tag{4.200}$$

where $\widetilde{R} = \widetilde{b}_1/(\widetilde{b}_1+\widetilde{b}_2)$, $\widetilde{R}_1 = v\widetilde{R}(1-\widetilde{R})$, $\widetilde{b}_1 = (n-1)/T_1(r_n)$ and $\widetilde{b}_2 = (m-1)/T_2(s_m)$.

4.4.2.4. Empirical Bayes Estimation of R

The Bayes estimates of R are obtained by using two different ways. It is clear that these estimators depend on the parameters of the prior distributions of b_1 and

b_2. However, the Bayes estimators can be also obtained independently of the prior parameters.

These prior parameters could be estimated by means of an empirical Bayes procedure, see [Lindley, 1969], [Awad and Gharraf, 1986]. Let $R_1, \ldots, R_n$ and $S_1, \ldots, S_m$ be two independent random samples from $Kum(1, b_1)$ and $Kum(1, b_2)$, respectively. For fixed $\underline{r}$, the function $L(b_1, 1\,|\underline{r})$ of b_1 can be considered as a gamma density with parameters $(n+1, T_1(r_n))$. Therefore, it is proposed to estimate the prior parameters α_1 and β_1 from the samples as $n+1$ and $T_1(r_n)$, respectively. Similarly, α_2 and β_2 could be estimated from the samples as $m+1$ and $T_2(s_m)$, respectively. Hence, the empirical Bayes estimate of R with respect to the SE and the LINEX loss functions, say $\widehat{R}_{EBS}$ and $\widehat{R}_{EBL}$, respectively, could be given as

$$\widehat{R}_{EBS} = \begin{cases} c_6 c_7 \, {}_2F_1(c_{13}, 2n+2; c_{13}+1; c_9) & \text{if } T_1 < T_2 \\ c_6 c_8 \, {}_2F_1(c_{13}, 2m+1; c_{13}+1; c_{10}) & \text{if } T_2 \leq T_1 \end{cases} \tag{4.201}$$

and

$$\widehat{R}_{EBL} = \begin{cases} -\dfrac{1}{v}\left((2n+1)\ln(T_1/T_2) + \ln c_{11}\right) & \text{if } T_1 < T_2 \\ -\dfrac{1}{v}\left((2m+1)\ln(T_2/T_1) - v + \ln c_{12}\right) & \text{if } T_2 \leq T_1 \end{cases} . \tag{4.202}$$

where $c_6 = (2n+1)/(2n+2m+2)$, $c_7 = (T_1/T_2)^{2n+1}$, $c_8 = (T_2/T_1)^{2m+1}$, $c_9 = 1-(T_1/T_2)$, $c_{10} = 1-(T_2/T_1)$, $c_{11} = \Phi_1(2n+1, c_{13}, c_{13}, c_9, -v)$, $c_{12} = \Phi_1(2m+1, c_{13}, c_{13}, c_{10}, v)$ and $c_{13} = 2n+2m+2$.

4.4.2.5. Bayesian Credible Intervals for R

It is known that $b_1\,|\underline{r} \sim Gamma(\delta_1, \lambda_1)$ and $b_2\,|\underline{s} \sim Gamma(\delta_2, \lambda_2)$. Then, $2(\beta_1 + T_1(r_n))b_1|\,\underline{r} \sim \chi^2(2(n+\alpha_1))$ and $2(\beta_2 + T_2(s_m))b_2|\,\underline{s} \sim \chi^2(2(m+\alpha_2))$. Therefore,

$$W = \frac{2(\beta_2 + T_2(s_m))b_2 \mid \underline{s}/2(m+\alpha_2)}{2(\beta_1 + T_1(r_n))b_1 \mid \underline{r}/2(n+\alpha_1)} \tag{4.203}$$

is an F distributed random variable with $(2(m+\alpha_2), 2(n+\alpha_1))$ degrees of freedom and the $100(1-\alpha)\%$ Bayesian credible interval for R can be obtained as

$$\left(\frac{1}{1+C_1\left(F_{2(m+\alpha_2),2(n+\alpha_1);\frac{\alpha}{2}}\right)}, \frac{1}{1+C_1\left(F_{2(m+\alpha_2),2(n+\alpha_1);1-\frac{\alpha}{2}}\right)}\right) \tag{4.204}$$

where $C_1 = \delta_2\lambda_1/\delta_1\lambda_2$, $F_{2(m+\alpha_2),2(n+\alpha_1);\frac{\alpha}{2}}$ and $F_{2(m+\alpha_2),2(n+\alpha_1);1-\frac{\alpha}{2}}$ are the lower and upper $\frac{\alpha}{2}$th percentile points of a F distribution with $(2(m+\alpha_2), 2(n+\alpha_1))$ degrees of freedom.

Moreover, this interval can be obtained independently of these parameters by using the empirical method. In this case, the posterior distributions of b_1 and b_2 have gamma distributions with parameters $(2n+1, 2T_1(r_n))$ and $(2m+1, 2T_2(s_m))$, respectively and the $100(1-\alpha)\%$ Bayesian credible interval for R can be obtained as

$$\left(\frac{1}{1+C_2\left(F_{(4m+2),(4n+2);\frac{\alpha}{2}}\right)}, \frac{1}{1+C_2\left(F_{(4m+2),(4n+2);1-\frac{\alpha}{2}}\right)}\right) \tag{4.205}$$

where $C_2 = ((4m+2)T_1(r_n)) \,/\, ((4n+2)T_2(s_m))$, $F_{(4m+2),(4n+2);\frac{\alpha}{2}}$ and $F_{(4m+2),(4n+2);1-\frac{\alpha}{2}}$ are the lower and upper $\frac{\alpha}{2}$th percentile points of a F distribution with $(4m+2, 4n+2)$ degrees of freedom.

4.4.3. Simulation Study

In this section, the results of simulation study are presented for comparing the risk of different estimators based on Monte Carlo simulation. All of the computations are performed by using MATLAB R2007a. All the results are based on 2500 replications.

We consider two cases separately to draw inference on R, namely when the common first shape parameter a is unknown and known. Without loss of generality, we take $a = 1$ when a is known. In both cases, we generate the record values with the sample sizes; $(n, m) = (5, 5)$, $(8, 8)$, $(10, 10)$, $(12, 12)$ from Kumaraswamy distribution.

In Table 4.12, the estimate of a is computed by using the iterative algorithm. The initial estimate of a is taken 1 and the iterative process stops when the difference between the two consecutive iterates are less than 10^{-6}. Once the estimate of a is obtained, the estimate of b_1 and b_2 are obtained by using equations (4.108) and (4.109), respectively. Finally, the MLE of R is obtained by using equation (4.112). The Bayes estimates under the SE and the LINEX $(v = 1)$ loss functions are obtained by using the Tierney and Kadane approximation. Prior 1: $(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) = (8, 10, 5, 4, 5, 5)$ and Prior 2: $(\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3) = (9, 5, 7, 1, 6, 5)$ are used for the true values of R are 0.501731 and 0.908896 and their results are tabulated in Table 4.12. Moreover, the average length of approximate confidence intervals and their cps are computed based on the asymptotic distribution of $\widehat{R}$ and is denoted by $\overline{L}_{AMLE}$. The nominal α value is 0.05.

From Table 4.12, it is observed that as the sample size increases in all the cases the average ERs of the estimators decrease, as expected. It verifies the consistency properties of all the estimates. The average length of the approximate confidence intervals also decrease as the sample size increases while the coverage probability is around 0.95. It is observed that the ER of Bayes estimate is smaller than that of ML estimate. Heuristically, in the Bayes approach we have extra information or data based on accumulated knowledge about the parameters as opposed to the MLE approach, therefore the Bayes estimate to be better than the MLE, in the sense that it has smaller ER.

In Table 4.13, the ML, UMVU and Bayesian estimates of R and their corresponding ERs are listed when a is known $(a = 1)$. The Bayes estimates are computed under the SE and the LINEX $(v = 1)$ loss functions for different prior parameters. The first two Bayes estimates are based on series expansion and the other two based on Lindley's approximation for the conjugate prior distributions. In addition, the empirical Bayes estimates are also obtained. Prior 3: $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (6, 8, 3, 5)$ and Prior 4: $(\alpha_1, \alpha_2, \beta_1, \beta_2) = (10, 6, 1, 8)$ are used for the true values of R are 0.548264 and 0.925025, respectively. Furthermore, the approximate and the exact confidence intervals for R are obtained by using equations (4.181) and (4.178). Finally, the Bayesian credible intervals are also obtained by using equations (4.205). The average length of the interval, denoted by $\overline{L}_{Bayes}$, and average length of exact confidence interval, denoted by $\overline{L}_{MLE}$, along with their cp's are reported in Table 4.13

From Table 4.13, the average ERs decrease as the sample size increases in all the cases. The Bayes estimate of R has the smallest ER. The Bayes estimates for series expansion and Lindley's methods are very close to each other. From this, we can infer that when the Bayes estimation can not be obtained in the closed form, the Lindley approximation is a good alternative. When the true value of R is 0.548264, we have $ER(\widehat{R}_{BS}) < ER(\widehat{R}_{EBS}) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_U)$. On the other hand, when the true value of R is 0.925025, we have $ER(\widehat{R}_{BS}) < ER(\widehat{R}_U) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_{EBS})$. Moreover, it is observed that the average confidence interval lengths decrease as the sample size increases. When the true values of R are 0.548264 and 0.925025, we have $\overline{L}_{MLE} < \overline{L}_{AMLE}$ and $\overline{L}_{AMLE} < \overline{L}_{MLE}$, respectively, while the cp is around 0.95. The Bayesian intervals have the smallest cp and is far from 0.95. Sometimes, the cp for the Bayesian interval based on equation (4.204) is not reasonable, because it contains prior parameters. That is why, they are not reported in the table.

Table 4.12: Estimates of R using the Priors 1-2 when a is unknown.

(n,m)	R	$\widehat{R}$	$\widehat{R}_{BS}$	$\widehat{R}_{BL}$	CI_{AMLE}
(5,5)	0.501731	0.505022	0.501740	0.608414	(0.226617,0.783427)
		0.013672	0.007156	0.016118	0.556810/0.942400
(8,8)		0.501802	0.505712	0.605696	(0.275885,0.727719)
		0.009306	0.006937	0.015867	0.451834/0.964000
(10,10)		0.503607	0.508404	0.603941	(0.300070,0.707145)
		0.007870	0.006655	0.015828	0.407074/0.958400
(12,12)		0.500559	0.509592	0.603050	(0.313687,0.687430)
		0.006952	0.006593	0.015653	0.373742/0.967200
(5,5)	0.908896	0.874885	0.892723	0.410055	(0.737286,1.012484)
		0.007655	0.001766	0.106927	0.275197/0.887600
(8,8)		0.878471	0.888055	0.411905	(0.771516,0.985427)
		0.004302	0.001676	0.106247	0.213910/0.942000
(10,10)		0.882239	0.886410	0.412547	(0.789373,0.975104)
		0.003039	0.001604	0.106019	0.185731/0.950800
(12,12)		0.885163	0.885177	0.413023	(0.802291,0.968034)
		0.002315	0.001570	0.105845	0.165743/0.959200

Notes: The first row represents the average estimates and the second row represents corresponding ERsfor each choice of m. The last two columns, the first row represents a 95% confidence interval and the second row represents their lengths and cp's.

Table 4.13: Estimates of R using Priors 3-4 when a is known ($a = 1$).

(n,m)	R	$\widehat{R}_{MLE}$	$\widehat{R}_U$	$\widehat{R}_{BS}$	$\widehat{R}_{BL}$	$\widehat{R}_{BS,Lindley}$	$\widehat{R}_{BL,Lindley}$	$\widehat{R}_{EBS}$	$\widehat{R}_{EBL}$	$\overline{L}_{AMLE}/cp$	$\overline{L}_{MLE}/cp$	$\overline{L}_{Bayes}/cp$
(5,5)	0.548264	0.544393	0.548398	0.557316	0.552500	0.557118	0.551725	0.542989	0.538888	0.537462	0.518719	0.357006
		0.018763	0.022050	0.009936	0.004983	0.009945	0.004981	0.017737	0.008854	0.903600	0.963200	0.825200
(8,8)		0.543087	0.545404	0.562618	0.558775	0.562463	0.558268	0.542159	0.539442	0.434061	0.423938	0.294623
		0.013143	0.014515	0.008814	0.004427	0.008820	0.004423	0.012645	0.006323	0.916000	0.952400	0.828800
(10,10)		0.550509	0.552673	0.567317	0.563939	0.567182	0.563538	0.549606	0.547795	0.389725	0.382397	0.266601
		0.010863	0.011784	0.008384	0.004215	0.008388	0.004210	0.010509	0.005338	0.910800	0.944800	0.816000
(12,12)		0.545903	0.547493	0.567477	0.564453	0.567368	0.564132	0.545219	0.544703	0.359160	0.353201	0.246897
		0.009507	0.010134	0.007955	0.004000	0.007958	0.003996	0.009258	0.005256	0.911600	0.946400	0.802400
(5,5)	0.925025	0.911806	0.924280	0.920444	0.921894	0.920252	0.919861	0.906515	1.206026	0.187805	0.217568	0.135508
		0.003336	0.002786	0.001016	0.000508	0.001016	0.000515	0.003635	0.813203	0.919000	0.956000	0.830000
(8,8)		0.915128	0.922831	0.917583	0.917401	0.917454	0.917109	0.911635	1.152133	0.146492	0.162202	0.105474
		0.001722	0.001492	0.000965	0.000484	0.000966	0.000488	0.001862	0.554699	0.949000	0.962000	0.837000
(10,10)		0.914794	0.920968	0.915886	0.915650	0.915782	0.915462	0.911941	1.053810	0.132102	0.143450	0.094836
		0.001422	0.001237	0.000951	0.000478	0.000953	0.000481	0.001531	0.345217	0.938000	0.945000	0.804000
(12,12)		0.914524	0.919657	0.914662	0.916367	0.914577	0.914279	0.912120	0.971850	0.121081	0.129777	0.086744
		0.001264	0.001116	0.000908	0.000433	0.000909	0.000458	0.001350	0.173744	0.963000	0.938000	0.823000

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of m. The last three columns the first row, represents the average length and the second row represents the cp's.

Table 4.14: Estimates of R for the non informative prior when a is known ($a = 1$).

(n, m)	R	$\widehat{R}_{MLE}$	$\widehat{R}_U$	$\widehat{R}^*_{BS}$	$\widehat{R}^*_{BL}$	$\widehat{R}^*_{BS,Lindley}$	$\widehat{R}^*_{BL,Lindley}$	$\overline{L}_{AMLE}/cp$	$\overline{L}_{MLE}/cp$
(5,5)	0.25(0.33)	0.266423(0.341186)	0.247968(0.326509)	0.281039(0.352284)	0.274082(0.343648)	0.284963(0.355540)	0.276210(0.344467)	0.450388(0.514840)	0.453386(0.502883)
		0.014043(0.017174)	0.014965(0.019907)	0.013617(0.015422)	0.006594(0.007581)	0.013721(0.015052)	0.006438(0.007197)	0.898400(0.906000)	0.960000(0.966400)
	0.50(0.70)	0.498926(0.685468)	0.498778(0.701779)	0.499025(0.672944)	0.488998(0.664462)	0.499061(0.669370)	0.486032(0.659295)	0.566406(0.495294)	0.540191(0.488108)
		0.021534(0.016031)	0.026191(0.017984)	0.018434(0.014843)	0.009221(0.007607)	0.017520(0.014675)	0.008832(0.007770)	0.895600(0.912000)	0.958400(0.965200)
	0.90(0.92)	0.883283(0.909039)	0.899104(0.922619)	0.868514(0.895774)	0.870454(0.904289)	0.865887(0.893757)	0.863164(0.891947)	0.244504(0.197887)	0.276317(0.231064)
		0.004750(0.003040)	0.003934(0.002399)	0.005789(0.003915)	0.005073(0.008279)	0.006146(0.004172)	0.003177(0.002144)	0.909200(0.914000)	0.954800(0.968400)
(8,8)	0.25(0.33)	0.258155(0.342902)	0.246412(0.334032)	0.268267(0.350278)	0.263823(0.344579)	0.269846(0.351560)	0.264850(0.344911)	0.359416(0.418494)	0.362350(0.411523)
		0.008198(0.011890)	0.008528(0.012909)	0.008069(0.011158)	0.003957(0.005501)	0.008109(0.011070)	0.003923(0.005392)	0.929200(0.913600)	0.967200(0.953200)
	0.50(0.70)	0.500203(0.689251)	0.500203(0.699365)	0.500200(0.680734)	0.493515(0.675326)	0.500201(0.679309)	0.492308(0.673352)	0.460892(0.397948)	0.445768(0.394544)
		0.014846(0.011260)	0.016760(0.012021)	0.013380(0.010748)	0.006695(0.005441)	0.013116(0.010711)	0.006580(0.005493)	0.913600(0.908400)	0.951600(0.948000)
	0.90(0.92)	0.890575(0.913068)	0.900090(0.921253)	0.881358(0.904929)	0.882153(0.908535)	0.880492(0.904281)	0.879087(0.903326)	0.186129(0.152204)	0.203301(0.169266)
		0.002573(0.001801)	0.002258(0.001521)	0.003010(0.002187)	0.002223(0.002282)	0.003089(0.002244)	0.001575(0.001139)	0.921600(0.920000)	0.957600(0.956800)
(10,10)	0.25(0.33)	0.256511(0.339056)	0.247177(0.331922)	0.264786(0.345207)	0.261269(0.340642)	0.265795(0.346043)	0.261900(0.340865)	0.322379(0.374726)	0.324724(0.369908)
		0.006859(0.010372)	0.007054(0.011087)	0.006789(0.009830)	0.003332(0.004861)	0.006815(0.009779)	0.003321(0.004803)	0.922800(0.903600)	0.952800(0.945600)
	0.50(0.70)	0.499460(0.691922)	0.499449(0.700009)	0.499472(0.684879)	0.493939(0.680577)	0.499472(0.683957)	0.493135(0.679302)	0.419019(0.356717)	0.407135(0.354505)
		0.010976(0.009746)	0.012118(0.010276)	0.010061(0.009371)	0.005018(0.004724)	0.009928(0.009350)	0.004960(0.004745)	0.924400(0.906800)	0.953200(0.942400)
	0.90(0.92)	0.890843(0.914273)	0.898366(0.920758)	0.883502(0.907798)	0.883779(0.908572)	0.882975(0.907412)	0.881880(0.906693)	0.166387(0.135063)	0.178772(0.147460)
		0.002411(0.001396)	0.002165(0.001209)	0.002735(0.001654)	0.001758(0.001039)	0.002779(0.001682)	0.001404(0.000850)	0.903600(0.937200)	0.938800(0.955200)
(12,12)	0.25(0.33)	0.259423(0.337613)	0.251695(0.331606)	0.266398(0.342910)	0.263532(0.339048)	0.267102(0.343507)	0.263852(0.339234)	0.298084(0.344073)	0.299745(0.340346)
		0.005944(0.008643)	0.006029(0.009140)	0.005944(0.008259)	0.002914(0.004101)	0.005965(0.008230)	0.002924(0.004064)	0.930800(0.911200)	0.954800(0.942800)
	0.50(0.70)	0.499460(0.694379)	0.499417(0.701201)	0.499494(0.688311)	0.494822(0.684715)	0.499500(0.687654)	0.494266(0.683813)	0.384216(0.327169)	0.374950(0.325513)
		0.009910(0.007805)	0.010754(0.008150)	0.009209(0.007560)	0.004621(0.003826)	0.009124(0.007552)	0.004585(0.003837)	0.920000(0.912400)	0.948800(0.942400)
	0.90(0.92)	0.892941(0.913177)	0.899198(0.918626)	0.886793(0.907736)	0.886150(0.907885)	0.886442(0.907472)	0.885586(0.906874)	0.150435(0.125032)	0.160191(0.134584)
		0.001642(0.001235)	0.001488(0.001076)	0.001857(0.001446)	0.000971(0.000837)	0.001880(0.001464)	0.000951(0.000737)	0.930800(0.941600)	0.958000(0.950400)

Notes: The first row represents the average estimates and the second row represents corresponding ERs for each choice of (n, m). But, for the last two columns, the first row represents the average length and second row represents their cp's. The corresponding results are reported within bracket in each cell for $R = 0.33$, 0.70 and 0.92.

In Table 4.14, the Bayes estimates of R are also obtained for the non informative prior case. The ML, UMVU, Bayes estimaties and confidence intervals of R are computed for $R = 0.25,\ 0.33,\ 0.5,\ 0.7,\ 0.90,\ 0.92$. The Bayes estimaties under the SE and the LINEX $(v = 1)$ loss functions are obtained by using both series expansion and Lindley's methods as in Table 4.14. Moreover, the average length of approximate and exact confidence intervals and their cps of R are evaluated.

From Table 4.14, the ERs decrease for all the estimates when the sample size increases, as expected. It is clear that the Bayes estimates for the Jeffrey's non informative prior case are very similar to the corresponding MLEs. More specifically, the Bayes estimate given in equation (4.199) is very close to the ML estimate after some algebraic operation in which they have suitable form for comparison. For $R = 0.25,\ 0.33,\ 0.5,\ 0.7$ the UMVUE has the greatest ER and we have $ER(\widehat{R}_{BS}) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_U)$. For $R = 0.90, 0.92$, we have $ER(\widehat{R}_U) < ER(\widehat{R}_{MLE}) < ER(\widehat{R}_{BS})$. Moreover, the average lengths of the intervals also decrease as the sample size increases. When $\widehat{R}^*_{BS} < \widehat{R}_{MLE} < R$, this is the case for bigger values of R such as $0.90, 092$, it can be shown that $ER(\widehat{R}_{MLE}) < ER(\widehat{R}_{BS})$ for $n = m$. When $R = 0.25,\ 0.90, 0.92$, we have $\overline{L}_{AMLE} < \overline{L}_{MLE}$. On the other hand, when $R = 0.33$, $0.50, 0.70$, we have $\overline{L}_{AMLE} > \overline{L}_{MLE}$. The cp for exact and approximate is around 0.95.

4.4.4. Conclusion

In this section, the different methods of estimations of $R = P(X < Y)$ are compared when X and Y are two independent Kumaraswamy distributions with the common first shape parameters.

When the first shape parameter is unknown, it is observed that the Bayesian estimators have a smaller ER. And this result does not change for the different values of the prior parameters. Nominal coverage probabilities are attained for the asymptotic confidence intervals.

When the first shape parameter is known, the different estimates, namely MLE, UMVUE with Bayes and empirical Bayes estimates are compared. The Bayesian estimates of R are obtained by using series expansion and Lindley's approximation method for both conjugate and non informative prior cases. Under both of these methods the ER are quite similar. The different confidence intervals of R, namely approximate, exact and Bayesian are compared. Even though, the prior parameters are not known it is observed that the Bayesian interval discussed in equation (4.205) is quite satisfactory.

The ML, UMVU, Bayesian estimates as well as confidence interval for R are invariant with respect to a monotone transformation on (X, Y), see [Kotz et al., 2003]. If X is Kumaraswamy then $-\ln X$ is the two parameter generalized exponential distribution. Therefore, all the estimates for R, mentioned above, under the Kumaraswamy distribution is the same as the two parameter generalized exponential distribution.

The ML, UMVU, Bayesian estimates of R in random samples depends on all the observation, but in record case they only depend on the last record value. Moreover, we considered the non informative case (a is known) when the number of random samples and the number of record values are taken to be equal as in the work of [Ahmadi and Arghami, 2001]. In this case, Monte Carlo simulation reveals out that the record case produces smaller ER for the Bayes estimation of R (when cps are similar) for the large sample sizes.

On the other hand, we may use Theorem 3.1 in [Ahmadi and Arghami, 2001] to say that (Fisher) information in record values is no different from that of random samples case under the assumption of X_i, $i = 1, ..., n$ and Y_j, $j = 1, ..., m$ distributes as $Kum(1, b)$, and the number of record values are the same as the number of random samples. When distribution involves more than one parameters, comparing the information in records with random samples is a subject of future studies.

5. CONCLUSIONS

The estimation problem of the parameters of the distribution and stress-strength reliability are considered in this dissertation. The different methods of estimation based on record values or record values with their corresponding inter-record times are obtained when the underlying distribution is the Burr Type XII, the generalized exponential and the Kumaraswamy. Moreover, the prediction problem of the future record values is considered for some cases. The comparison of all obtained estimates is demonstrated by simulation study and real life examples. Detailed findings of the simulation results on inferences based on each distribution considered in this thesis are described at the end of each section.

REFERENCES

Ahmad K. E., Fakhry M. E., Jaheen, Z. F., (1997), “Empirical Bayes estimation of $P(Y < X)$ and characterizations of Burr-type X model”, Journal of Statistical Planning and Inference, 64, 297-308.

Ahmadi J., Arghami N. R., (2001), “On the Fisher information in record values”, Metrika, 53, 195-206.

Ahmadi J., Doostparast M., (2006), “Bayesian estimation and prediction for some life distributions based on record values”, Statistical Papers, 47 (3), 373-392.

Ahmadi J., Jozani M. J., Marchand E., Parsian A., (2009), “Prediction of k-records from a general class of distributions under balanced type loss functions”, Metrika, 70 (1),19-33.

Ahsanullah M., (1980), “Linear prediction of record values for the two parameter exponential distribution”, Annals of the Institute of Statistical Mathematics, 32 (1), 363-368.

Ahsanullah M., (1995), “Record Statistics”, 1st Edition, Nova Science Publishers.

Al-Hussaini E. K., Jaheen Z. F., (1992), “Bayesian estimation of the parameters, reliability and failure rate functions of the Burr Type XII failure model”, Journal of Statistical Computation and Simulation, 41 (1), 31-40.

Al-Hussaini E. K., Jaheen Z. F., (1995), “Bayesian prediction bounds for the Burr Type XII failure model”, Communications in Statistics-Theory and Methods, 24 (7), 1829-1842.

Ali Mousa M. A. M., (2001), “Inference and prediction for the Burr Type X model based on records”, Statistics: A Journal of Theoretical and Applied Statistics, 35 (4), 415-425.

Arnold B. C., Balakrishnan N., Nagaraja H. N., (1998), “Records”, 1st Edition, Wiley.

Arnold S. F., (1990), “Mathematical Statistics”, 1st Edition, Prentice-Hall.

Asgharzadeh A., Fallah A., (2011), “Estimation and prediction for exponentiated family of distributions based on records”, Communications in Statistics-Theory and Methods, 40 (1), 68-83.

Asgharzadeh A., Valiollahi R., Raqab M. Z., (2013), “Estimation of the stress–strength reliability for the generalized logistic distribution”, Statistical Methodology, 15, 73-94.

Asgharzadeh A., Valiollahi R., Raqab M. Z., (2014), “Estimation of $Pr(Y < X)$ for the two-parameter of generalized exponential records”, Communications in Statistics-Simulation and Computation, DOI: 10.1080/03610918.2014.964046.

Awad A. M., Gharraf M. K., (1986), “Estimation of $P(Y < X)$ in the Burr case: A comparative study”, Communications in Statistics-Simulation and Computation, 15 (2), 389-403.

Awad A. M., Raqab M. Z., (2000), “Prediction intervals for the future record values from exponential distribution: comparative study”, Journal of Statistical Computation and Simulation, 65, 325-340.

Baklizi A., (2008), “Likelihood and Bayesian estimation of $Pr(X < Y)$ using lower record values from the generalized exponential distribution”, Computational Statistics and Data Analysis, 52, 3468-3473.

Baklizi A., (2012), “Inference on $\Pr(X < Y)$ in the two-parameter Weibull model based on records”, ISRN Probability and Statistics, 2012, 1-11.

Baklizi A., (2014), “Bayesian inference for $\Pr(Y < X)$ in the exponential distribution based on records”, Applied Mathematical Modelling, 38, 1698-1709.

Balakrishnan N., Chan P. S., (1998), “On the normal record values and associated inference”, Statistics & Probability Letters, 39, 73-80.

Basak P., Balakrishnan N., (2003), “Maximum likelihood prediction of future record statistic”, Series on Quality, Reliability and Engineering Statistics, 7, 159-175.

Birnbaum Z. W., (1956), “On a use of Mann-Whitney statistics”, Proceeding Third Berkeley Symposium Mathematica Statistics and Probability, 13-17, Berkeley, California, USA, 2-5 Jully.

Birnbaum Z. W., McCarty B. C., (1958), “A distribution-free upper confidence bounds for $Pr(Y < X)$ based on independent samples of X and Y”, The Annals of Mathematical Statistics, 29 (2), 558-562.

Bolstad W. M., (2007), “Introduction to Bayesian Statistics”, 2nd Edition, Wiley.

Burr I. W., (1942), “Cumulative frequency functions”, The Annals of Mathematical Statistics, 13, 215-232.

Chandler K. N., (1952), “The distribution and frequency of record values”, Journal of the Royal Statistical Society Series B, 14, 220-228.

Chen M. H., Shao Q. M., (1999), “Monte Carlo estimation of Bayesian credible and HPD intervals”, Journal of Computational and Graphical Statistics, 8 (1), 69-92.

Cordeiro G. M., de Castro M., (2011), “A new family of generalized distributions”, Journal of Statistical Computation and Simulation, 81 (7), 883-898.

Crowder M. J., (2000), “Tests for a family of survival models based on extremes”, Recent Advances in Reliability Theory, 1st Edition, Birkhauser.

Dey S., Dey T., Salehi M., Ahmadi J., (2013), “Bayesian inference of generalized exponential distribution based on lower record values”, American Journal of Mathematical and Management Sciences, 32, 1-18.

Doostparast M., (2009), “A note on estimation based on record data”, Metrika, 69, 69-80.

Doostparast M., Ahmadi J., (2006), “Statistical analysis for geometric distribution based on records”, Computers & Mathematics with Applications, 52, 905-916.

Doostparast M., Akbari M. G., Balakrishnan N., (2010), “Bayesian analysis for the two-parameter Pareto distribution based on record values an times”, Journal of Statistical Computation and Simulation, 81 (11), 1393-1403.

Doostparast M., Balakrishnan N., (2010), “Optimal sample size for record data and associated cost analysis for exponential distribution”, Journal of Statistical Computation and Simulation, 80 (12), 1389-1401.

Doostparast M., Balakrishnan N., (2011), “Optimal record-based statistical procedures for the two-parameter exponential distribution”, Journal of Statistical Computation and Simulation, 81 (12), 2003-2019.

Doostparast M., Balakrishnan N., (2013), “Pareto analysis based on records”, Statistics: A Journal of Theoretical and Applied Statistics, 47 (5), 1075-1089.

Doostparast M., Deepak S., Zangoie A., (2013), “Estimation with the lognormal distribution on the basis of records”, Journal of Statistical Computation and Simulation, 83 (12), 2339-2351.

El-Deen M. M. S., Al-Dayian G. R., El-Helbawy A. A., (2014), “Statistical Inference for Kumaraswamy Distribution Based on Generalized Order Statistics with Applications”, British Journal of Mathematics & Computer Science, 4 (12), 1710-1743.

Feller W., (1966), “An introduction to probability theory and its applications Volume 2”, 2nd Edition, Wiley.

Fletcher S. C., Ponnamblam K., (1996), “Estimation of reservoir yield and storage distribution using moments analysis”, Journal of Hydrology, 182, 259-275.

Ganji A., Ponnambalam K., Khalili D., Karamouz, M., (2006), “Grain yield reliability analysis with crop water demand uncertainty”, Stochastic Environmental Research and Risk Assessment, 20, 259-277.

Garg M., (2009), “On Generalized Order Statistics from Kumaraswamy Distribution”, Tamsui Oxford Journal of Mathematical Sciences, 25 (2), 153-166.

Gelman A., Carlin J. B., Stern H. S., Rubin D. B., (2003), “Bayesian Data Analysis”, 2nd Edition, Chapman & Hall.

Ghitany M. E., Al-Awadhi S., (2002), “Maximum likelihood estimation of Burr XII distribution parameters under random censoring”, Journal of Applied Statistics, 29 (7), 955-965.

Ghitany M. E., Al-Jarallah R. A., Balakrishnan N., (2013), “On the existence and uniqueness of the MLEs of the parameters of a general class of exponentiated distributions”, Statistics: A Journal of Theoretical and Applied Statistics, 47 (3), 605-612.

Glick N., (1978), “Breaking records and breaking boards”, American Mathematical Monthly, 85, 2-26.

Gomes A. E., da-Silva C. Q., Cordeiro G. M., Ortega E. M. M., (2014), “A new lifetime model: the Kumaraswamy generalized Rayleigh distribution”, Journal of Statistical Computation and Simulation, 84 (2), 290-309.

Gradshteyn I. S., Ryzhik I. M., (1994), “Table of Integrals, Series and Products”, 5th Edition, Academic Press.

Gulati S., Padgett W. J., (1994), “Smooth nonparametric estimation of the distribution and density functions from record-breaking data”, Communications in Statistics-Theory and Methods, 23 (5), 1256-1274.

Gupta R. D., Kundu D., (1999), “Generalized exponential distributions”, Australian and New Zealand Journal of Statistics, 41, 173-188.

Gupta R. D., Kundu D., (2001), “Generalized exponential distributions: Different methods of estimation”, Journal of Statistical Computational and Simulation, 69 (4), 315-338.

Gupta R. D., Kundu D., (2002), “Generalized exponential distributions: Statistical inferences”, Journal of Statistical Theory and Applications, 1, 101-118.

Gupta R. D., Kundu D., (2007), “Generalized exponential distribution: Existing results and some recent developments”, Journal of Statistical Planning and Inference, 137, 3537-3547.

Hendi M. I., Abu-Youssef S. E., Alraddadi A. A., (2007), “A Bayesian analysis of record statistics from the Rayleigh model”, International Mathematical Forum, 2 (13), 619-631.

Hofmann G., Nagaraja H. N., (2003), “Fisher information in record data”, Metrika, 57, 177-193.

Jaheen Z. F., (2004), “Empirical Bayes Inference for Generalized Exponential Distribution Based on Records”, Communications in Statistics-Theory and Methods, 33 (8), 1851-1861.

Jaheen Z. F., (2005), “Estimation Based on Generalized Order Statistics from the Burr Model”, Communications in Statistics-Theory and Methods, 34 (4), 785-794.

Jones M. C., (2009), “Kumaraswamy’s distribution: a beta-type distribution with some tractability advantages”, Statistical Methodology, 6 (1), 70-81.

Kotz S., Lumelskii Y., Pensky M., (2003), “The Stress-Strength Model and its Generalizations: Theory and Applications”, 1st Edition, World Scientific.

Koutsoyiannis D., Xanthopoulos T., (1989), “On the parametric approach to unit hydrograph identification”, Water Resources Management, 3, 107-128.

Kumaraswamy P., (1976), “Sinepower probability density function”, Journal of Hydrology, 31, 181-184.

Kumaraswamy P., (1978), “Extended sinepower probability density function”, Journal of Hydrology, 37, 81-89.

Kumaraswamy P., (1980), “‘A generalized probability density function for double-bounded random processes”, Journal of Hydrology, 46, 79-88.

Kundu D., Gupta R. D., (2005), “Estimation of $P(Y < X)$ for generalized exponential distribution”, Metrika, 61, 291-308.

Kundu D., Gupta R. D., (2006), “Estimation of $P(Y < X)$ for the Weibull distribution”, IEEE Transactions on Reliability, 55 (2), 270-280.

Kundu D., Gupta R. D., (2008), “Generalized exponential distribution: Bayesian estimations”, Computational Statistics and Data Analysis, 52, 1873-1883.

Kundu D., Pradhan B., (2009), “Estimating the parameters of the generalized exponential distribution in presence of hybrid censoring”, Communications in Statistics-Theory and Methods, 38 (12), 2030-2041.

Lawless J. F., (2003), “Statistical Models and Methods for Lifetime Data”, 2nd Edition, Wiley.

Lehmann E. L., Casella G., (1998), “Theory of point estimation”, 2nd Edition, Springer.

Lemonte A. J., (2011), “Improved point estimation for the Kumaraswamy distribution”, Journal of Statistical Computation and Simulation, 81 (12), 1971-1982.

Lindley D. V., (1969), “Introduction to Probability and Statistics from a Bayesian Viewpoint Volume 1, 1st Edition, Cambridge University Press.

Lindley D. V., (1980), “Approximate Bayes method”, Trabajos de Estadistica, 3, 281-288.

Madadi M., Tata M., (2011), “Shannon information in record data”, Metrika, 74, 11-31.

Madi M. T., Raqab M. Z., (2007), “Bayesian prediction of rainfall records using the generalized exponential distribution”, Environmetrics, 18, 541-549.

Mitnik P. A., (2013), “New properties of the Kumaraswamy distribution”, Communications in Statistics-Theory and Methods, 42 (5), 741-755.

Mokhlis N. A., (2005), “Reliability of a Stress-Strength Model with Burr Type III Distributions”, Communications in Statistics-Theory and Methods, 34 (7), 1643-1657.

Nadar M., Papadopoulos A., (2011), “Bayesian analysis for the Burr Type XII distribution based on record values”, Statistica, 71 (4), 421-435.

Nadar M., Kızılaslan F., Papadopoulos A., (2014), “Classical and Bayesian estimation of $P(Y < X)$ for Kumaraswamy’s distribution”, Journal of Statistical Computation and Simulation, 84 (7), 1505-1529.

Nadarajah S., (2008), “On the distribution of Kumaraswamy”, Journal of Hydrology, 348, 568-569.

Nelson W. B., (1972), “Graphical analysis of accelerated life test data with the inverse power law model”, IEEE Transactions on Reliability, 21 (1), 2-11.

Nevzorov V., (2001), “Records: mathematical theory”, 1st Edition, American Mathematical Society.

Panahi H., Asadi S, (2010), “Estimation of $R = P(Y < X)$ for two-parameter Burr Type XII distribution”, World Academy of Science, Engineering and Technology, 4, 12-23.

Papadopoulos A. S., (1978), “The Burr distribution as a failure model from a Bayesian approach”, IEEE Transactions on Reliability, 27 (5), 369-371.

Paranaíba P. F., Ortega E. M. M., Cordeiro G. M., de Pascoa M. A. R., (2013), “The Kumaraswamy Burr XII distribution: theory and practice”, Journal of Statistical Computation and Simulation, 83 (11), 2117-2143.

Ponnambalam K., Seifi A., Vlach J., (2001), “Probabilistic design of systems with general distributions of parameters”, International Journal of Circuit Theory and Applications, 29, 527-536.

Pradhan B., Kundu D., (2009), “On progressively censored generalized exponential distribution”, Test, 18 (3), 497-515.

Press S. J., “Subjective and Objective Bayesian Statistics: Principles, Models, and Applications”, 2th Edition, Wiley.

Rényi A., (1962), “Théorie des éléments saillants d’une suite d’observations”, Colloquim on Combinatorial Methods in Probability Theory, 104-117.

Rao C. R., (1965), “Linear Statistical Inference and Its Applications”, 1st Edition, Wiley.

Raqab M. Z., (2002), “Inferences for generalized exponential distribution based on record statistics”, Journal of Statistical Planning and Inference, 104, 339-350.

Raqab M. Z., Ahmadi J., Doostparast M., (2007), “Statistical inference based on record data from Pareto model”, Statistics: A Journal of Theoretical and Applied Statistics, 41 (2), 105-118.

Raqab M. Z., Sultan K. S., (2014), “Generalized exponential records: existence of maximum likelihood estimates and its comparison with transforming based estimates”, Metron, 72, 65-76.

Rastogi M. K., Tripathi Y. M., (2012), “Estimating the parameters of a Burr distribution under progressive type II censoring”, Statistical Methodology, 9, 381-391.

Rezaei S., Tahmasbi R., Mahmoodi M., (2010), “Estimation of $P(Y < X)$ for generalized Pareto distribution”, Journal of Statistical Planning and Inference, 140, 480-494.

Samaniego F. J., Whitaker L. R., (1986), “On estimating popular characteristics from record breaking observations I.Parametric results”, Naval Research Logistics Quarterly, 33, 531–543.

Sarhan M. A.,Tadj L., (2008), “Inference using record values from generalized exponential distribution with application”, Bulletin of Statistics Economics, 2, 72-85.

Seifi A., Ponnambalam K., Vlach J., (2000), “Maximization of manufacturing yield of systems with arbitrary distributions of component values”, Annals of Operations Research, 99, 373-383.

Sindhu T. N., Feroze N., Aslam M., (2013), “Bayesian Analysis of the Kumaraswamy Distribution under Failure Censoring Sampling Scheme”, International Journal of Advanced Science and Technology, 51, 39-58.

Soliman A. A., (2005), “Estimation of parameters of life from progressively censored data using Burr-XII model”, IEEE Transactions on Reliability, 54 (1), 34-42.

Soliman A. A., Abd-Ellah A. H., Abou-Elheggag N. A., Ahmed E. A., (2011), “Reliability estimation in stress–strength models: an MCMC approach”, Statistics: A Journal of Theoretical and Applied Statistics, 47 (4): 715-728.

Soliman A. A., Abd-Ellah A. H., Sultan K. S., (2006), “Comparison of estimates using record statistics from Weibull model: Bayesian and non-Bayesian approaches”, Computational Statistics and Data Analysis, 51 (1), 2065-2077.

Sundar V., Subbiah K., (1989), “Application of double bounded probability density function for analysis of ocean waves”, Ocean Engineering, 16, 193-200.

Tarvirdizade B., (2013), “Estimation of $\Pr(X > Y)$ for exponentiated gumbel distribution based on lower record values”, İstatistik Journal of Turkish Statistical Association, 6 (3), 103-109.

Tierney L., (1994), “Markov chains for exploring posterior distributions”, The Annals of Statistics, 22 (4), 1701-1728.

Tierney L., Kadane J. B., (1986), “Accurate approximations for posterior moments and marginal densities”, Journal of the American Statistical Association, 81, 82-86.

Varian H. R., (1975), “A Bayesian approach to real estate assessment”, Studies in Bayesian Econometrics and Statistics, 1, 195-208.

Wang L., Shi Y., (2010), “Empirical Bayes inference for the Burr model based on records”, Applied Mathematical Sciences, 34, 1663-1670.

Wang L., Shi Y., (2013), “Reliability analysis of a class of exponential distribution under record values”, Journal of Computational and Applied Mathematics, 239, 367-379.

Wang B. X., Ye Z. S., (2015), “Inference on the Weibull distribution based on record values”, Computational Statistics and Data Analysis, 83, 26-36.

Web 1, (2015), http://home.iitk.ac.in/ kundu/2MSSBD.pdf, (Acces Time: 16/03/2015).

Wong A. C. M., Wu Y. Y., (2009), “A note on interval estimation of $P(X < Y)$ using lower record data from the generalized exponential distribution”, Computational Statistics and Data Analysis, 53, 3650-3658.

BIOGRAPHY

Fatih Kızılaslan was born in İstanbul, Turkey in 1983. He graduated from Üsküdar Burhan Felek High School in 2000. He received the B.Sc. and M.Sc. degrees in Mathematics from the Gebze Institute of Technology, Turkey in 2007 and 2009. He has been working as a research assistant at Gebze Institute of Technology since 2007. His research interests include record values and reliability.

APPENDICES

Appendix A: Publications Based on the Thesis

Kızılaslan F., Nadar M., (2013), “Bayesian and Non-Bayesian analysis for the Burr Type XII distribution based on record values and inter-record times ”, 7th International Days of Statistics and Economics, 601-610, Prague, Czech Republic, 19-21 September.

Kızılaslan F., Nadar M., (2013), “Classical and Bayesian Analysis for The Generalized Exponential Distribution Based on Record Values and Times ”, Anadolu University Journal of Science and Technology-B: Theoretical Sciences, 2 (2), 111-120.

Kızılaslan F., Nadar M., (2015), “Estimation with the generalized exponential distribution based on record values and inter-record times”, Journal of Statistical Computation and Simulation, 85 (5), 978-999.

Nadar M., Kızılaslan F., (2014), “Classical and Bayesian estimation of $P(X < Y)$ using upper record values from Kumaraswamy’s distribution”, Statistical Papers, 55 (3), 751-783.

Nadar M., Kızılaslan F., (2014), “Estimation and prediction of the Burr Type XII distribution based on record values and inter-record times ”, Journal of Statistical Computation and Simulation, DOI:10.1080/00949655.2014.970554.

Nadar M., Papadopoulos A., Kızılaslan F., (2013), “Statistical analysis for Kumaraswamy’s distribution based on record data”, Statistical Papers, 54 (2), 355-369.

Appendix B: Conference Presentations Based on the Thesis

Kızılaslan F., (2011), “Bayesian and Non-Bayesian estimation of Kumaraswamy distribution ”, 7th International Statistics Congress, Antalya, Turkey, 28 April- 01 May.

Kızılaslan F., Nadar M., (2012), “Classical and Bayesian Analysis for The Generalized Exponential Distribution Based on Record Values and Times ”, 8th International Symposium of Statistics, Eskisehir, Turkey, 11-13 October.

Kızılaslan F., Nadar M., Papadopoulos A., (2011), “Classical and Bayesian estimation of $P(X < Y)$ using upper record values from Kumaraswamy’s distribution ”, Yeditepe International Research Conference on Bayesian Learning, Istanbul, Turkey, 15-17 June.