وزارة التعليم العالي والبحث العلمي

University 20 August 1955 of Skikda
Faculty of Science
Department of Mathematic

جامعة 20 أوت 1955 سكيكدة
كلـــــية العلـــوم
قسم الرياضيات

N[0] : U.S/F.S/D.M/……/2023.

Faculty of Science
Department of Mathematic

Memory

Presented for the graduation of Master in Mathematics

Estimation of the Gompertz Distribution Parameters under Joint Progressive Censoring Data

Option : ANEDP

By: ***BOUTOBZA Roqiya & KETTOUCHE Karima***

Supervisor: BOULKEROUA Fouzia M.C.B U.SKIKDA

In front of the jury:

President: MALLEM Khadidja M.C.B U. SKIKDA
Examiner: KAREK Chafia M.C.B U. SKIKDA

Year: 2022/2023

DEDICATION

I dedicate this a achievement to the symbol of love and the balm of Misery: My mother **Fatima**, and to one who reaped the thorns from my path and to Pave the path of knowledge for me dear father **Laidi**. And this achievement is small for them for all the moral and material they gave me and their standing beside me and offering them what they could small and large until finished this difficult journey that lasted for many years.

I would like to dedicate this achievement to my brother **Ayoub**, And my sister **Assia**, **Rayane**, **Mouna**.

To all my family members end everyone who stood by me in complete my academic career, And we pray to god to complete success for us.

Roqiya

DEDICATION

Praise be to god, enough, and prayers be upon the beloved chosen one ,his family, and those who fulfilled .As for what follows : Praise be to god, who enabled us to value this step in our academic journey , with this note of ours, the fruit what of effort and success, by his grace ,dedicated to the honorable **parents** , may god protect them, and their performance as a light for my path .

To all the honorable family who helped me, **sisters** and **brothers** and their buds.

To my esteemed husband who contributed to my encouragement the rest of the journey .

To the children of the future god willing especially those who shared with me the rest of the journey.

To my **husbands** family and her buds.

To the companions of the fence that shared his moments.

To every one who helped me and supported me in the academic journey.

I ask god for success, payment and success hopefully.

Karima

ACKNOWLEDGEMENTS

Befor presenting this work , First of all , we would like to thank Our god for believing in him, enlightening our path, and guiding us in the best and right ways .

First of all , we would like to thank our supervisor, **Dr. FOUZIA BOULKEROUA** that followed us and guided us throughout the completion of this work ,that which has been a very valuable help on the scientific and moral level, we thank the jury members: **Dr.MALLEM KHADIDJA** and **Dr. KAREK CHAFIA**

We would like to express our sincere thanks to all the teachers in the Mathematics department , who have helped us throughout our university course.

We would also like to thank all the people who helped us during our teach us.

We also thank our colleagues for their support and encouragement.

We are very grateful to our parent s, brothers and sisters , who encouraged and helped me to reach this formation.

Finally , we would like to express our sincere gratitude to all those who participated directly Or indirectly in the realization of this thesis.

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

GO: Gompertz distribution.

PDF: Probability Density Function.

CDF: Cumulative Distribution Function.

JPC: Joint Progressive Censoring.

CI: Confidence Interval.

Cr: Credible Interval.

SE: Sequare Error.

LINEX: Linear-Exponential.

MLEs: Maximum Likelihood Estimators.

M-H: Metropolis-Hastings.

BEs: Bayesian Estimators.

ABSTRACT

This thesis aims to estimate the Gompertz distributions parameters using the joint progressive type-II censoring scheme. For the estimate problem, the likelihood, Bootstrap, and Bayesian approaches are used. We employ various numerical methods, such Newton-Raphson, to solve the likelihood equations because the resulting maximum likelihood estimators are not written in closed forms. Additionally, we take into account the Bayesian method for estimating the unknown parameters while using independent gamma priors for the scale and shape parameters. We employ importance sampling and Metropolis-Hastings approaches in the Bayesian analysis, which are based on symmetric and asymmetric loss functions. Furthermore, credible intervals based on the Bayesian method and confidence intervals based on asymptotic normality are presented. To compare the performance of the proposed methods, a Monte Carlo simulation is run, and real-life data is examined for illustrative purposes.

RÉSUMÉ

Le but de cette mémoire est d'évaluer les paramétres de la distribution Gompertz dans le cadre du schéma de contrôle progressif commun type-II. Pour le probléme d'estimation, les méthodes de probabilité, de bootstrap et de Bayesiennes sont utilisées. Afin de résoudre les équations de probabilité, nous utilisons des techniques numériques telles que celles de Newton-Raphson, car les estimateurs de probabilité maximaux qui ont été obtenus ne sont pas présentés en formes fermées. Nous prenons également en compte l'approche Bayessienne pour estimer les parameters inconnus sous des priores gamma indépendantes sur la taille et la forme des parameters. L'échantillonnage d'importance et les méthodes de Metropolis-Hastings basées sur deux types de fonctions de perte sont utilisés dans l'analyse bayésienne. Les intervalles de confiance basés sur la normalité asymptotique et les intervalles de confiance basés sur la méthode Bayessienne sont également abordés. Pour comparer les méthodes proposées, une simulation Monte Carlo est effectuée et des données de la vie réelle sont analysées.

ملخص

الهدف من هذه الأطروحة هو تقدير معاملات توزيع جومبرتز لنوع من البيانات المشتركة المبتورة بشكل لقد تم استخدام طرق الإحتمالية العظمى والبوتستراب والبيز لحل مشاكل تقدير المعلمات. نظراً لأن .تدريجي تقديرات الإحتمالية القصوى التي تم الحصول عليها لا يتم التعبير عنها بشكل مبسط ، فإننا نستخدم بعض الطرق العددية مثل نيوتن رافسون . وكذلك نأخذ بعين الاعتبار نهج بيز للمعلمات المجهولة والتي تتبع توزيعات جاما. ولإستخدام تحليل بيز، نستخدم طرق التقدير العيني وطرق ميتروبوليس-هاستينجس واعتماد دوال الخسارة (الدوال المتماثلة وغير المتماثلة). ولأغراض المقارنات بين تقادير المعلمات فقد أجرينا دراسة محاكاة وذلك لمعرفة الطريقة الأفضل. وبالنتيجة فقد توصلنا الى أن طريقة ميتروبوليس هي الطريقة الأفضل في إيجاد تقدير المعلمات سواء التقدير النقطي أو التقدير بفترة.

CHAPTER 1

INTRODUCTION

The censoring scheme is critical in many laboratory life testing trials for cost and time management, and it aids researchers in reducing the time, it takes to track the lifespan of all units in the trial or test. Furthermore, censorsing occurs when some individuals have little information about the time frame. Naturally, there are various kinds of censored tests that have been researched in the literature. The most common kind is the type-II censoring scheme. In this scheme, n units are tested, but instead of continuing until all units fail, the test stops at the $m-th$ failure time $(1 \leq m \leq n)$. Various distributions have been analyzed in a number of studies involving this method of censoring. The above-mentioned scheme will be useless if the experimenter wants to remove live units at points other than the final termination point of the life test, this provision will be desirable, as with the case of accidental breakage of experimental items, in which the failure of test items will be unavoidable at points other than the termination point. A generalization of type-II censoring is called the progressive type II scheme which has drawn a lot of attention in statistic research over the past ten to fifteen years. A number of authors, such as Cohen [8], Mann [21], Viveros and Balakrishnan [32], and Balakrishnan and Aggarwala [4], have studied the statistical inference on time failure parameters under progressive type II censoring. Almost all of these types of data are focused only on one sample issues. However, in several situations the experimenter attempts to compare different populations. In this context, a joint censoring scheme has been investigated in the literature. Lately, it has presented by Rasouli and Balakrishnan [28] for a comparative study of products of two populations coming from different units within the same facility.

The literature related to this issue under research like, Balakrishnan et al. [5] whose extended the JPC scheme for more than two exponential populations. Mondal and Kundu [24] considered the problem of point and interval estimation of the unknown parameters of Weibull populations under JPC scheme. Doostparast et al. [9] obtained the Bayes estimates of the parameters for a general class of distributions with respect to the SE and LINEX loss functions under the JPC scheme. Parsi et al. [27]) provided the conditional maximum likelihood and interval estimation of the parameters of two Weibull distributions. Parsi and Bairamov [26] determined the expected number of failures in life testing experiment under the JPC scheme for different families. See also, Ashour and Abo-Kasem [3], Mondal and Kundu [25].

In this work, we analyze the joint progressively censored data when the lifetime distributions of the experimental units of the two populations follow two-parameter Gompertz distribution, which is a commonly used as continuous distribution. It is a type of probability distribution that is skewed both to the right or to left. This family was introduced at first by Gompertz [15]. It is known as a generalization of the exponential distribution. It has some interesting connections with well-known distributions like exponential, double exponential, Weibull, Gumbel, and generalized logistic distributions. Demographers and actuaries have paid a lot of attention to this distribution. The GO distribution has been used in a variety of applications, including survival analysis and certain sciences such as: gerontology, computer science, genetics, medicine, and marketing science. More applications and survey of the Gompertz model can be found in Ahuja and Nash [2]. They discussed the generalized Gompertz Verhulst family of distributions.

The estimation of GO parameters under different censoring schemes is considered in Chen [6], Ismail [17] and Ghitany et al. [14].

The organization of this thesis is as follows: we start by an introduction in Chapter 1. Then, in Chapter 2, we review and describe some basic definitions relevant to the topic under study in this dissertation. In Chapter 3, based on joint progressive type II censored sample from GO(α_1, θ) and GO(α_2, θ), the maximum likelihood approach is used to estimate the shape and scale parameters α_1, α_2 and θ, where we can not obtained the MLE in closed form for this we suggest using approximation numerical Newton-Raphson to calculate mmaximum likelihood estimators. Also, the Bayesian approaches are used to estimate α_1, α_2 and θ, or some function of α_1, α_2 and θ, say $\lambda = g(\alpha_1, \alpha_2, \theta)$, under different loss functions. The Markov Chain Monte Carlo method has been used to compute the bayesian estimators and also to construct symmetric credible intervals for α_1, α_2 and θ. Finaly, a numerical study based on simulation analysis and a real data analyses are demonstrated in Chapter 4, all accounts using the program R. We conclude our findings of this thesis in Chapter 5.

CHAPTER 2

BASIC DEFINITIONS

This chapter provides a reminder of the statistical properties and probability distributions that we will use later.

2.1 Random variable

The notion of random variable is the main concept of probability theory and statistics.

2.1.1 Real Random Variable

Let (Ω, S, P) be a probability space. A real random variable is an application T of Ω in R

$$T : \ \omega \in \ \Omega \longrightarrow T(\omega) \in R.$$

When the set of possible realizations of the random variable T is finite or countable, the real random variable T is said to be discrete random variable. Otherwise, we say that the variable random T is continuous random variable.

2.1.2 Real Random Vector

when studying a random experiment one uses most of the time several random variables $T_1, ..., T_k$ which are defined on the same probability space since they relate to the same experiment. Equivalently, we can consider that the real random variables $T_1, ..., T_k$ define a random variable T with values in R^k by posing

$$T = (T_1, ..., T_k).$$

2.2 Probability Density Function

For a continuous random variable $T = t$, a probability density function is a functin such that:

1. $f(t) \geq 0$.
2. $\int\limits_{-\infty}^{\infty} f(t)dt = 1$.
3. $P(a \leq T \leq b) = \int_a^b f(t)dt =$ area under $f(t)$ from a to b for any a and b in R.

2.3 Cumulative Distribution Function

The distribution function or cumulative distribution function (CDF) of a random variable T is the F_T function such that

$$\begin{aligned} F_T: \quad R &\longrightarrow [0,1] \\ t &\longmapsto F_T(t) = P(T \leq t). \end{aligned}$$

2.4 Mean and Variance

Suppose $T = t$ is a continuous random variable with density function $f(t)$.

- The mean or the expected value of T denoted as μ or $E(T)$ is

$$\mu = E(T) = \int_{-\infty}^{\infty} tf(t)dt.$$

- The variance of T denoted as $V(T)$ or σ^2 is

$$\sigma^2 = V(T) = \int_{-\infty}^{\infty} (t-\mu)^2 f(t)dx = \int_{-\infty}^{\infty} t^2 f(t)dt - \mu^2.$$

The standard deviation of T is $\sigma = \sqrt{\sigma^2} = \sqrt{V(T)}$.

2.5 Gompertz Distribution

Definition 2.5.1. *A continuous random variable T is said to have a Gompertz distribution with shape parameter $\alpha > 0$, and scale parameter $\theta > 0$, shown as $T \sim GO(\alpha, \theta)$, if its probability density function (PDF) is given by*

$$f(t;\theta,\alpha) = \alpha\ \theta\ e^{\theta\, t - \alpha\, (e^{\theta\, t} - 1)},\ t > 0, \tag{2.1}$$

and its cumulative distribution function (CDF) is given by

$$F(t;\theta,\alpha) = 1 - e^{-\alpha\, (e^{\theta\, t} - 1)},\ t > 0. \tag{2.2}$$

Figure (2.1) shows the density function of the Gompertz distribution for some selected choices of α and θ.

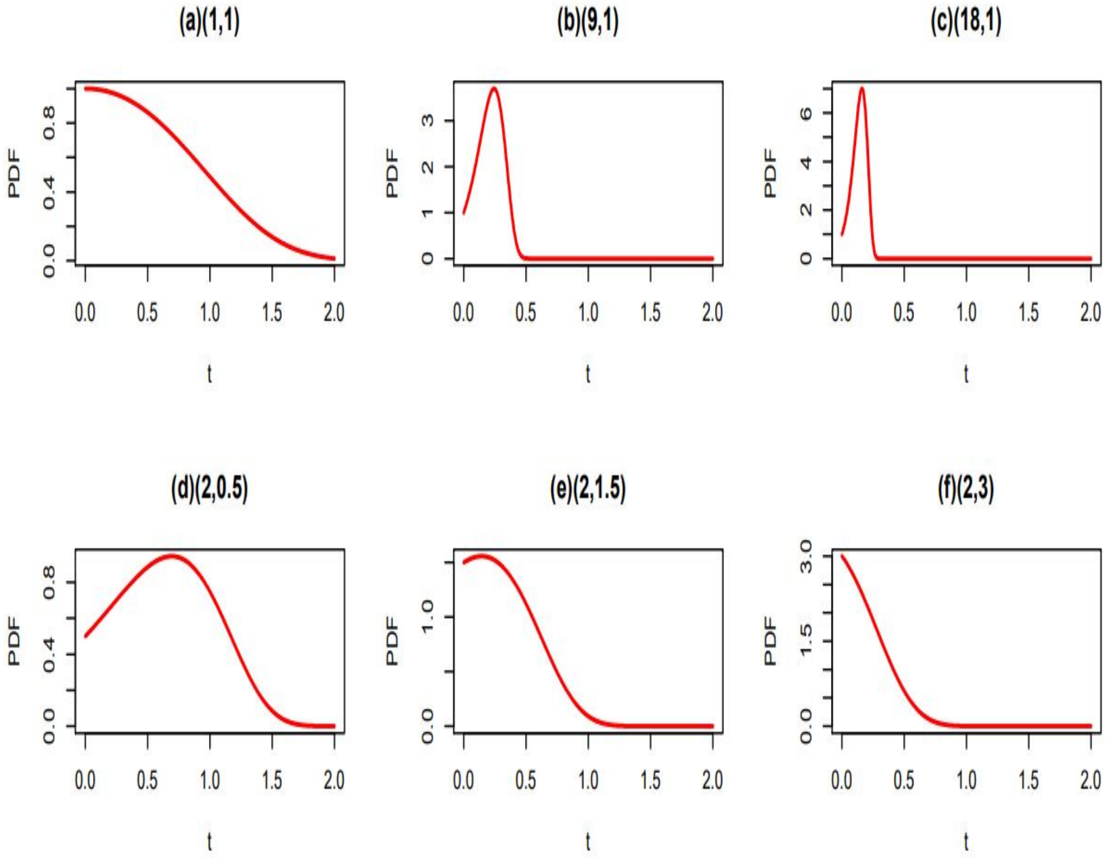

Figure 2.1: graphical representation of the density function of Gompertz distribution $Go(\alpha, \theta)$ for some values.

2.6 Progressively Censored Data

As we can tell **previossly** that the progressive censoring scheme is a generalization of type-II censoring, it has a simple description given by: assume n identical items are put on a life-testing experiment, and it is decided to observe only k $(k < n)$ failures, then, given a censoring scheme $\boldsymbol{R} = (R_1, ..., R_m)$, the remaining $(n - k)$ items are censored progressively as follows. At the time of the first failure, R_1 of the $(n - 1)$ surviving components are removed randomly from the experiment; at the time of the next failure, R_2 items of the remaining $(n - 2 - R_1)$ surviving components are removed randomly from the experiment;

repeating this censoring up to the final stage, at the time of the $k-th$ failure, all remaining $R_k = n - k - R_1 - \ldots - R_{k-1}$ are removed. n, m and $\boldsymbol{R}$ are prefixed constants. This censoring scheme can be depicted pictorially as in Figure 2.2 (see Balakrishnan and Aggarwala [4]).

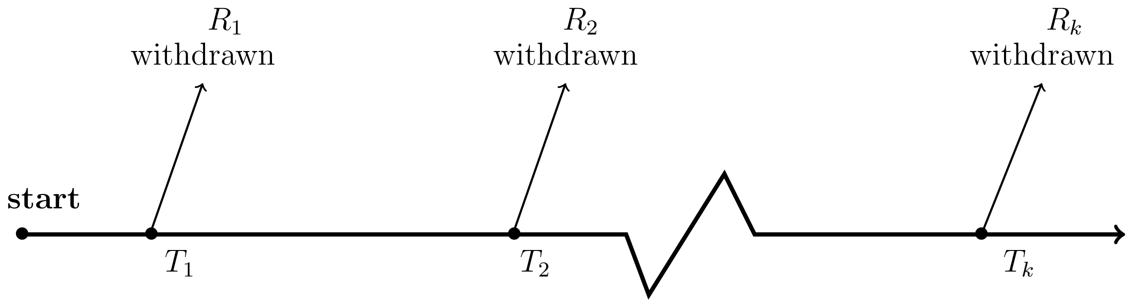

Figure 2.2: Progressive censoring of life test with intervention times $T_1, ..., T_k$ and censoring scheme $R_1, ..., R_k$, $T_1 < T_2, ..., T_k$.

The likelihood function under this progressive censoring scheme is defined by

$$L(\boldsymbol{t}) = C \prod_{i=1}^{k} f(t_i)(1 - F(t_i))^{R_i},$$

where

$$C = n(n - R_1 - 1)(n - R_1 - R_2 - 2)...(n - \sum_{j=1}^{k-1} R_j - m + 1).$$

2.7 The Joint Progressive Censoring Data

Almost, all traditional censoring schemes are based on one sample problem, but when comparative study among more than one population is of interest, we can rely on different joint censoring schemes, such as the **joint progressive censoring** (JPC) scheme. Us, to reduce the cost and completion time of the experiment, the experimenter uses a JPC scheme and terminates the life testing comparative experiment as soon as a pre-specified number of failures (say k) occurs. Here, we give a description about this censoring scheme and its properties with a schematic diagram below (Figure 2.3). The JPC scheme was first introduced by Rasouli and Balakrishnan [28]. More precisely, suppose two independent samples of sizes m and n are combined and placed on a life testing experiment simultaneously, which are selected from two different groups of production Group 1 (Gr-1) and Group 2 (Gr-2) under the same operating situations. The size of the combined sample is $N = m + n$. Let $k < N$ and $R_1, R_2, ..., R_k$ are non-negative integers satisfying $\sum_{i=1}^{k} R_i + k = m + n$, where $R_i = S_i + W_i$ with S_i and W_i being the number of removals at the i-th stage from Gr-1 and Gr-2, respectively. Based on the combined sample, at the first failure time T_1, $R_1 = S_1 + W_1$ units are randomly withdrawn from the remaining $(N - 1)$ surviving units where S_1 and W_1 are the number of removed units from Gr-1 and Gr-2, respectively. Proceeding similarly, at the second time of failure, T_2, $R_2 = S_2 + W_2$ items are chosen randomly and withdrawn from the remaining combined $N - 2-$ R_1 surviving units, and so on. In the final stage (at the time of the k-th failure), all the remaining $R_k = N - k - \sum_{i=1}^{k} R_i$ surviving units are withdrawn (see Balakrishnan and Cramer (2014)). The JPC scheme includes the complete sample and type-II censoring schemes as special cases when $R_1 = R_2 = ... = R_k = 0$, for all $i = 1, 2, ..., k$ and $R_1 = R_2 = ... = R_{k-1} = 0$, so that $R_k = N - k$, respectively.

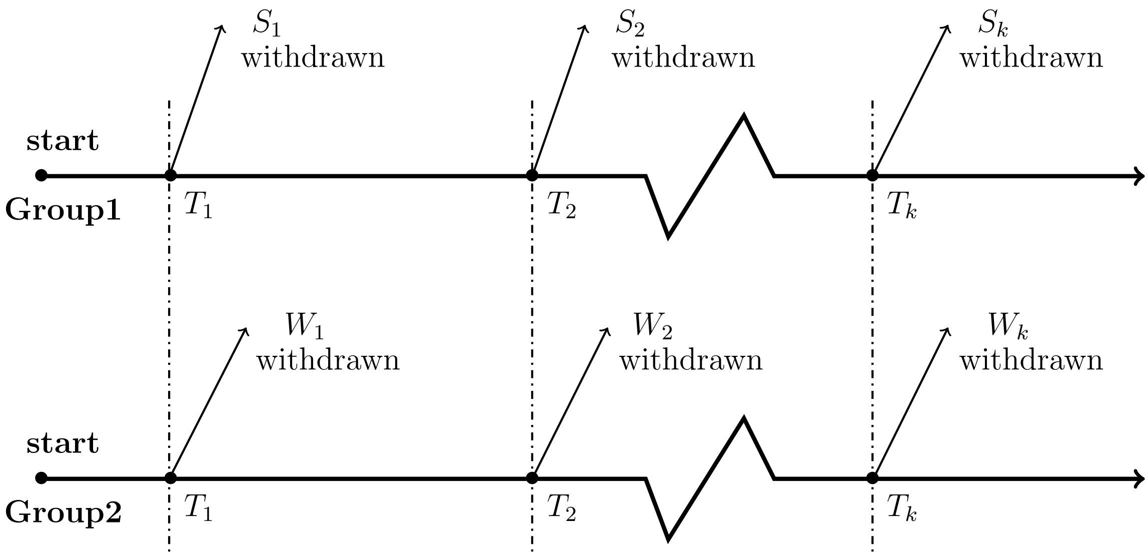

Figure 2.3: Joint progressive censoring of life test with intervention times $T_1, ..., T_k$ and censoring scheme $R_1 = S_1 + W_1, ..., R_k = S_k + W_k$.

Let $X_1, X_2, ..., X_m$ independent and identically (iid) distributed m units from Gr-1 of GO lifetime distribution with CDF $F(x; \alpha_1, \theta)$, and PDF $f(x; \alpha_1, \theta)$. Furthermore, let $Y_1, Y_2, ..., Y_n$ be iid n units from Gr-2 of GO lifetime distribution with CDF $F(y; \alpha_2, \theta)$ and PDF $f(y; \alpha_2, \theta)$. For a given $(R_1, \ldots, R_k)$, where k

is a non-negative integer such that ($k \geq 0$) satisfying $\sum_{i=1}^{k} R_i = N - k$, let $(\boldsymbol{t}, \boldsymbol{\delta}, \boldsymbol{S}) = \{(t_1, \delta_1, S_1), (t_2, \delta_2, S_2), \ldots, (t_k, \delta_k, S_k)\}$, be the JPC data from the combined population. Here for $j = 1, \ldots, k$, we have $\delta_j = 1$ if the failure at t_j occurs from Gr-1, and $\delta_j = 0$ if the failure at t_j occurs from Gr-2. S_j and then $W_j = R_j - S_j$, denotes the number of items removed at t_j from Gr-1 and Gr-2, respectively.

The likelihood function based on JPC sample is given by

$$
\begin{aligned}
L(\boldsymbol{t} \mid \alpha_1, \alpha_2, \theta) &= C \prod_{i=1}^{k} \Big[f(t_i, \alpha_1, \theta) \Big]^{\delta_i} \Big[f(t_i, \alpha_2, \theta) \Big]^{1-\delta_i} \Big[1 - F(t_i, \alpha_1, \theta) \Big]^{S_i} \\
& \quad \Big[1 - F(t_i, \alpha_2, \theta) \Big]^{W_i}, \qquad (2.3)
\end{aligned}
$$

where

$$
k_1 = \sum_{i=1}^{k} \delta_i, \ k_2 = k - k_1, \ S_k = m - k_1 - \sum_{i=1}^{k-1} S_i, \ C = D_1 D_2,
$$

with D_1 and D_2 being

$$
D_1 = \prod_{j=1}^{k} \left[\left(m - \sum_{i=1}^{j-1} \delta_i - \sum_{i=1}^{j-1} S_i \right) \delta_j + \left(n - \sum_{i=1}^{j-1} (1 - \delta_i) - \sum_{i=1}^{j-1} W_i \right) (1 - \delta_j) \right],
$$

and

$$
D_2 = \prod_{j=1}^{k-1} \frac{\binom{m - \sum_{i=1}^{j-1} \delta_i - \sum_{i=1}^{j-1} S_i}{S_j} \binom{n - \sum_{i=1}^{j-1} (1-\delta_i) - \sum_{i=1}^{j-1} W_i}{W_j}}{\binom{N - j - \sum_{i=1}^{j-1} R_i}{R_j}}.
$$

2.8 Quantile Function

The quantile function of a random variable (or probability distribution) is the inverse of its distribution function. We call the quantile function of T the function, denoted Q_T, from $]0, 1[$ in R by

$$
Q_T(q) = \inf\{F_T(t) \geq q\}, \ 0 < q < 1.
$$

2.9 Estimation

An estimator of an unknown parameter of a model or probability distribution is a function which corresponds to a sequence of observations $t_1, t_2, ..., t_n$ from the model or the law probability, the value $\hat{\theta}$ called an estimator or an estimate

$$\hat{\theta} = f(t_1, t_2, ..., t_n).$$

2.10 Quality of an Estimator

2.10.1 Bias

A random variable fluctuates around his expexted value. It is therefore to be hoped that The expectation of $\hat{\theta}$ is equal to θ. $Bias(\hat{\theta}) = E(\hat{\theta}) - \theta$. When $Bias(\hat{\theta}) = 0$, the estimator is said to be unbiased, and if $Bias(\hat{\theta}) > 0$, the estimator is said positively biased.

2.10.2 Mean Squared Error

The Mean Squared Error (MSE) also called quadratic risk is the expectation of the square of the error between the true value and its value Estimated.

$$MSE(\hat{\theta}) = E(\hat{\theta} - \theta)^2.$$

If the risk is low, the estimator $\hat{\theta}$ is close to θ.

2.10.3 Convergence in Distribution

For random variables $T_n \in R$ and $T \in R$, T_n converges in distribution to X (note $T_n \xrightarrow{\text{D}} T$), if for all t such that $t \longrightarrow P(T \leq t)$ is continuous,

$$P(T_n \leq t) \xrightarrow{\text{D}} P(T \leq t) \text{ as } n \longrightarrow \infty.$$

2.11 Asymptotic Normality

The following central limit theorem and Delta method will be accepted.

2.11.1 Central Limit Theorem

Theorem 2.11.1. *(Central Limit Theorem (TCL))*

Let $T_1, T_2, ..., T_n$ be a sequence of i.i.d random variables of the same distribution (thus of same expectation μ and same standard deviation σ).

$$Z_n = \frac{T_n - \mu}{\frac{\sigma}{\sqrt{n}}} \longrightarrow N(0, 1).$$

The random variable Z_n converges in distribution to the reduced centered normal distribution. It is a consequence of the TCL which ensures that

$$\sqrt{n}(T_n - \mu) \xrightarrow{D} N(0, \sigma^2).$$

2.11.2 Delta Method

Let $T_1, T_2, ..., T_n$ be a sequence of random variables of expectation θ and standard deviation σ. Si $\sqrt{n}(T_n - \theta) \xrightarrow{\text{D}} N(0, \sigma^2)$, and a differentiable function g such that $g'(\theta) \neq 0$. In this case the delta method gives

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

2.12 Goodness-of-Fit Tests

Here, we present two statistics to measure the distance between the CDF of the parent model and the empirical distribution of the sample. The first test statistic is Kolmogorov-Smirnov (K-S) test. The K-S test is a non-parametric test that measures the distance between the empirical distribution function say $F_n(t)$ of the sample and the cumulative distribution function of the reference distribution say $F^*(t)$. The empirical distribution function is defined as

$$F_n(t) = \frac{\#(i : t_i \leq t)}{n}.$$

The K-S distance D_n can be given as:

$$D_n = \sup_{t \in R} |F_n(t) - F^*(t)|, \tag{2.4}$$

Small values of D_n indicate goodness of fit. For more details about K-S test, see Massey [22] and Durbin [10]. Another test is Cramer-von Mises (CvM) test. Let $F_n(t)$ be the empirical distribution function and $F^*(t)$ be the cumulative distribution function of a variable t. The CvM statistic is a classical goodness-of-fit statistic that characterizes the distance between $F_n(t)$ and $F^*(t)$ in l^2-norm. It is defined as

$$w^2 = \int_{-\infty}^{+\infty} \left[F_n(t) - F^*(t)\right]^2 \, dF^*(t). \tag{2.5}$$

Note that supremum (2.5) must occur at one of the observed values t_i or to the left of t_i. Small values of w^2 indicate goodness of fit.

2.13 Maximum Likelihood Method

Let T_i for $i = 1, ...n$, n random variables i.i.d, the likelihood function is written as:

$$L(T_1, ..., T_n; \theta) = \begin{cases} \prod_{i=1}^{n} P(T_i = t_i; \theta), & \text{If the } T_i \text{ are discerete.} \\ \prod_{i=1}^{n} f(t_i; \theta), & \text{If the } T_i \text{ are continuous.} \end{cases}$$

The maximum likelihood estimate of θ is the value $\hat{\theta}$, of which maximizes the likelihood function $L(T_1, ..., T_n; \theta)$. Done $\hat{\theta}_n$, will generally be calculated by maximizing the log-likelihood

$$\hat{\theta} = arg\max \ln L(T_1, ..., T_n; \theta) = arg\max \, l(T_1, ..., T_n; \theta),$$

when $\theta = (\theta_1, ..., \theta_d) \in \Theta$ and all partial derivatives below exist, is solution of the system of equations called likelihood equations

- $\frac{\partial}{\partial \theta_j}\, l(T_1, ..., T_n, \theta) = 0,\ \forall\, j \in 1, ..., d.$
- $\frac{\partial^2}{\partial \theta_j^2}\, l(T_1, ..., T_n; \theta) < 0.$

2.14 Newton Raphson Method

Newton's method for the solution of the equation $f(t) = 0$ is defined by

$$t_{k+1} = t_k - \frac{f(t_k)}{f'(t_k)}, k = 0, 1, 2, ...,$$

with prescribed starting value t_0 .We implicitly assume in the defining formula that $f'(t_k) \neq 0$ for all $k \geq 0$.

2.15 Fisher Information Matrix

The (symmetric) Fisher Information Matrix $I(\theta)$ is the $d \times d$ matrix with entries

$$-E\left[\frac{\partial^2 l(\theta)}{\partial\theta_i \partial\theta_j}\right],\ 1 \leq i, j \leq d.$$

The observed Fisher information matrix is

$$J(\theta) = \frac{\partial^2 l(\theta)}{\partial\theta_i \partial\theta_j},\ 1 \leq i, j \leq d.$$

2.16 Confidense Interval

A confidence interval, in statistics, refers to the probability that a population parameter will fall between a set of values for a certain proportion of times. For repeated random samples from a distribution with unknown parameter θ, a γ. 100% confidence interval will cover θ in γ. 100% of all cases.

2.17 Bootstrap Method

The bootstrap inference was developed by Efron [11]. This method is based on repeated random sampling with replacement from the original sample, $\boldsymbol{t} = (t_1, t_2, ..., t_n)$. Mainly, we generate random samples of the same size n as the original sample, which is known as a bootstrap sample, denoted $\boldsymbol{t^*}$. The term "with replacement" refers to the fact that each bootstrap sample will select any observation several times. To calculate a $100(1-\gamma)\%$ $(0 < \gamma < 1)$ confidence interval (CI) for unknown parameter θ from a sample $\boldsymbol{t} = (t_1, t_2, ..., t_n)$, we take a random sample $\boldsymbol{t^*}$, with replacement from data of the same size as the original sample, and calculate the estimate of the parameter θ (say $\hat{\theta}^*$) from this bootstrap random sample. We do this repeatedly, say B times. So we now have B bootstrap samples $\boldsymbol{t^*} = (\boldsymbol{t_1^*}, \boldsymbol{t_2^*}, ..., \boldsymbol{t_B^*})$, and B estimates of θ, one from each sample $(\hat{\theta}_1^*, \hat{\theta}_2^*, ..., \hat{\theta}_B^*)$. If these are ordered in increasing value, $(\hat{\theta}_{(1)}^*, \hat{\theta}_{(2)}^*, ..., \hat{\theta}_{(B)}^*)$, a bootstrap $100(1-\gamma)\%$ CI for θ would be from the $(\frac{\gamma}{2})B\text{-}th$ to the $(1-\frac{\gamma}{2})B\text{-}th$ largest values.

2.18 Bayesian Estimation and Related Techniques

One of the most popular techniques in analyzing a wide variety of models is the Bayesian approach. In Bayesian principle, the unknown parameter θ is considered as a random variable and it is assumed to have a prior probability distribution, which is denoted by $\Pi(\theta)$. Based on the observed data $\boldsymbol{t} = (t_1, t_2, ..., t_n)$ from a statistical model $f(\boldsymbol{t}|\theta)$ and the prior distribution $\Pi(\theta)$, we update our belief about θ using the posterior distribution $\Pi(\theta|\boldsymbol{t})$ as:

$$\Pi(\theta|\boldsymbol{t}) = \frac{f(\boldsymbol{t}, \theta)}{f(\boldsymbol{t})} = \frac{f(\boldsymbol{t}|\theta)\ \Pi(\theta)}{f(\boldsymbol{t})} = \frac{f(\boldsymbol{t}|\theta)\ \Pi(\theta)}{\int f(\boldsymbol{t}|\theta)\ \Pi(\theta)\ d\theta}.$$

The quantity

$$f(\boldsymbol{t}) = \int f(\boldsymbol{t}|\theta)\ \Pi(\theta)\ dt,$$

is normalizing constant of the posterior distribution.

Bayes estimation depends on both the prior distribution(s) of the parameter(s) of a statistical model and on the loss function used. Hence, it is necessary to define the prior type as well as the loss function that is used in Bayesian estimation. Below are two different kinds of priors.

- **The Non-informative Prior:** If a probability density function $\Pi(\theta)$ of θ provides no details about θ, it is said to be a non-informative prior. Clearly, $\Pi(\theta) = 1$ and $\Pi(\theta) = \frac{1}{\theta}$ are two basic examples of non-informative priors.
- **The Informative Prior:** The informative prior distribution has its own parameters called hyper-parameters. In the general Bayesian method, the conjugate distribution is widely used as an informative prior. If the prior and posterior distributions are all from the same family, the prior is considered to be a conjugate prior for that family of distributions.

In statistics, a loss function represents the loss associated with an error in estimation. For estimating any parameter (say, θ) by a decision rule d, the following loss functions are often considered.

- **Square Error (SE) Loss Function:** The SE loss function is symetric functions, it is associated with losses of equal magnitude for overestimation and underestimation. It has a simple defnition as

$$L_S = (\theta - d)^2. \tag{2.6}$$

The Bayes estimator (BE) of θ (say $\hat{\theta}_S$) under this type of loss function, is the posterior mean, i.e.

$$\hat{\theta}_S = E[\theta|\boldsymbol{t}]. \tag{2.7}$$

- **Linear-Exponential (LINEX) Error Loss Function:** This loss function is an assymetric function and it defined as

$$L_L = e^{\nu(d-\theta)} - \nu(d - \theta) - 1, \nu \neq 0, \tag{2.8}$$

where sign of the shape parameter ν reflects the direction of assymmetry, and its magnitude reflects the degree of assymmetry. The LINEX loss function was introduced by Varian [31]. Under the LINEX loss function, the BE (denoted by $\hat{\theta}_L$) is given by

$$\hat{\theta}_L = -\frac{1}{\nu} \log \left(E[e^{-\nu\,\theta} | \boldsymbol{t}] \right). \tag{2.9}$$

2.18.1 Markov Chain Monte Carlo Simulation

The Markov Chain Monte Carlo (MCMC) method is a general simulation method used to sample posterior distributions and compute posterior quantities of interest. This technique is often applied to solve integration (such as the posterior mean) and optimization problems in large dimensional spaces. However, in most instances, the integration does not have a closed structure. The idea of Monte Carlo simulation is to draw an i.i.d. set of samples $\{\theta_i\}_{i=1}^N$ from a target density $p(\theta)$ defined on a high-dimensional space Z. By using Markov chain samples and Monte Carlo integration, we are able to approximate the integration with no closed form. This integration could be described as

$$\int_Z g(\theta)\; p(\theta)\; d\theta = \frac{1}{N} \sum_{j=1}^{N} g(\theta_j),$$

where $g(.)$ is a function of interest, $p(\theta)$ is the posterior distribution of θ and $\{\theta_j\}$ are MCMC samples from $p(\theta)$ on its support Z, and N is the number of desired samples. For more details of the MCMC method see, Smith and Roberts [29]. There are several common methods for defining MCMC, including Metropolis-Hastings (M-H) method.

2.18.1.1 Metropolis-Hastings Method

The M-H algorithm is used to implement posterior simulation in almost any problem that allows for pointwise evaluation of the prior distribution and likelihood function. The M-H algorithm is one of the most widely used MCMC methods, with applications in reliability analysis, statistical physics, and machine learning, among others. This algorithm is a very general MCMC method created by Metropolis et al. [23] and later expanded by Hastings [16]. The M-H algorithm allows for sampling from an analytic target complicated distribution (denoted $p(.)$) by filtering samples from a proposal distribution (denoted $q(.|.)$) that is also given in the analytic form. It operates by sampling a chain of correlated samples that converge in distribution to the target (see M-H Algorithm).

Algorithm 1:

- Step 1: Select an initial guess t_0;
- Step 2: For $j = 1, 2, \ldots, M$, repeat:

  (a) Draw candidate t^* with its pdf $q(t|t_{j-1})$ and u from $U(0, 1)$;

  (b) Compute the acceptance probability

$$\varepsilon = \frac{p(t^*)/q(t^*|t_{j-1})}{p(t_{j-1})/q(t_{j-1}|t^*)} = \frac{p(t^*)}{p(t_{j-1})}\frac{q(t_{j-1}|t^*)}{q(t^*|t_{j-1})},$$

  (c) If $u < \min(1, \varepsilon)$, then

  accept the proposal: set $t_j = t^*$,

  else reject the proposal: $t_j = t_{j-1}$ and go to Step (a).

2.18.1.2 Importance Sampling Method

Importance sampling is a classical option that dates back to the 1940s; for example, see Geweke [13]. It is a common solution when numerical integration is difficult, and it has become the foundation for many importance sampling-based Bayesian network algorithms. The concept behind value sampling is to rewrite the mean in the following way. Let $p(t)$ be a probability density of t over domain $Z \subset R^n$ where R is the set of real numbers, and let $H(t)$ be another probability density function on R^n such that its support includes the support of $p(t)$, where $H(T)$ should be easy to sample from. Then, we can rewrite $I(f)$ as follows:

$$I(f) = \int_Z f(t)p(t)dt = \int_Z \frac{f(t)p(t)}{H(t)}H(t)dt = \int_Z f(t)w(t)H(t)dt, \qquad (2.10)$$

where $w(t) = p(t)/H(t)$ is the importance weight. Also, we can write the last expression as

$$I(f) = E_H\left[f(t)w(t)\right], \qquad (2.11)$$

where E_H is the expectation for which the distribution of t is $H(t)$ rather than $p(t)$. The density $p(t)$ is called the nominal or target distribution, and $H(t)$ is called the importance or proposal function. In order to estimate the integral, we follow the next algorithm

Algorithm 2:

- Generate samples $(t_1, , t_n)$ according to the distribution $H(t)$.
- Use the generated values in the sample-mean formula, then the estimator for $I(f)$ is
$$\hat{I}_H(f) = \frac{1}{n}\sum_{i=1}^{n}\frac{f(t_i)p(t_i)}{H(t_i)} = \frac{1}{n}\sum_{i=1}^{n} f(t_i)w(t_i).$$

CHAPTER 3

METHODS OF ESTIMATION

In this chapter, we consider JPC scheme based on two combined samples from GO distribution. The MLEs of the unknown parameters are derived in Section 3.1. The asymptotic normality of the MLEs and bootstrap methods are applied to establish CIs for the parameters in this section also. Further, the BEs of the parameters using the SE and LINEX loss functions are discussed in Section 3.2. The M-H and Lindley's approximation methods are proposed to obtain these BEs and then construct the corresponding CrIs.

3.1 Maximum Likelihood Estimation

Here in this section, we consider the likelihood and bootstrap method for estimating the unknown parameters and corresponding confidence intervals. by combining (2.1),(2.2), and (2.3), we write the likelihood function of JPC GO data which can be viewed as a function of α_1,α_2 and θ as:

$$
\begin{aligned}
L(\boldsymbol{t}|\alpha_1,\alpha_2,\theta) &= C\prod_{i=1}^{k}[\alpha_1 e^{\theta t_i-\alpha_1(e^{\theta t_i}-1)}]^{\delta_i}[\alpha_2 e^{\theta t_i-\alpha_2(e^{\theta t_i}-1)}]^{\delta_i-1} \\
&\times [1-(1-e^{-\alpha_1(e^{\theta t_i}-1)})]^{Si}[1-(1-e^{-\alpha_2(e^{\theta t_i}-1)})]^{wi} \\
&= C\prod_{i=1}^{k}[\alpha_1 e^{\theta t_i-\alpha_1(e^{\theta t_i}-1)}]^{\delta_i}[\alpha_2 e^{\theta t_i-\alpha_2(e^{\theta t_i}-1)}]^{\delta_i-1} \\
&\times [e^{-\alpha_1(e^{\theta t_i}-1)}]^{Si}[e^{-\alpha_2(e^{\theta t_i}-1)}]^{wi} \\
&= C\alpha_1^{\delta_i}\alpha_2^{1-\delta_i}\theta e^{\theta t_i(\delta_i+1-\delta_i)} \times e^{-\alpha_1(\delta_i(e^{\theta t_i}-1)+S_i(e^{\theta t_i}-1))} \\
&= C\alpha_1^{k_1}\alpha_2^{k_2}\theta^k e^{\theta\sum_{i=1}^{k} t_i} \times e^{-\alpha_1\sum_{i=1}^{k} A_\theta(t_i)(\delta_i+S_i)} \times e^{-\alpha_2\sum_{i=1}^{k} A_\theta(t_i)(1-\delta_i+w_i)} \\
&= C\alpha_1^{k_1}\alpha_2^{k_2}\theta^k e^{\theta\sum_{i=1}^{k} t_i-\alpha_1\sum_{i=1}^{k} u_i A_\theta(t_i)-\alpha_2\sum_{i=1}^{k} V_i A_\theta(t_i)} \qquad (3.1)
\end{aligned}
$$

where, $A_\theta(t_i) = e^{\theta t_i} - 1, u_i = \delta_i + S_i$ and $V_i = 1 - \delta_i + w_i$

The corresponding log-likelihood function can be written as

$$
\begin{aligned}
l(\boldsymbol{t}|\alpha_1,\alpha_2,\theta) &\propto k_1\ log\ \alpha_1 + k\ log\ \theta + k_2\ log\ \alpha_2 + \theta\sum_{i=1}^{k} t_i - \alpha_1\sum_{i=1}^{k} u_i A_\theta(t_i) \\
&- \alpha_2\sum_{i=1}^{k} v_i A_\theta(t_i). \qquad (3.2)
\end{aligned}
$$

The MLEs can be obtained by taking the first partial derrivatives of Eq (3.2) concerning α_1, α_2 and θ then equating each to zero. That is, for $k_1 > 0$ and $k_2 > 0$, the likelihood equations can be obtained as follows :

$$
\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_1} = \frac{\partial\left(k_1\ log\ \alpha_1 - \alpha_1\sum_{i=1}^{k} u_i A_\theta(t_i)\right)}{\partial\alpha_1} = \frac{k_1}{\alpha_1} - \sum_{i=1}^{k} u_i A_\theta(t_i) = 0 \qquad (3.3)
$$

$$
\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_2} = \frac{\partial\left(k_2\ log\ \alpha_2 - \alpha_2\sum_{i=1}^{k} u_i A_\theta(t_i)\right)}{\partial\alpha_2} = \frac{k_2}{\alpha_2} - \sum_{i=1}^{k} v_i A_\theta(t_i) = 0 \qquad (3.4)
$$

$$
\begin{aligned}
\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\theta} &= \frac{\partial\left(k\ log\ \theta+\theta\sum_{i=1}^{k}t_i-\alpha_1\sum_{i=1}^{k}u_iA_\theta(t_i)-\alpha_2\sum_{i=1}^{k}v_iA_\theta(t_i)\right)}{\partial\theta} \\
&= \frac{k}{\theta}+\sum_{i=1}^{k}t_i-\alpha_1\sum_{i=1}^{k}u_it_ie^{\theta t_i}-\alpha_2\sum_{i=1}^{k}v_it_ie^{\theta t_i}=0 \qquad (3.5)
\end{aligned}
$$

It follows from (3.3) and (3.4) that the MLEs of α_1 and α_2 can be obtained, respectively,

$$
\hat{\alpha}_1(\theta)=\frac{k_1}{\sum_{i=1}^{k}u_iA_\theta(t_i)},\ \hat{\alpha}_2(\theta)=\frac{k_2}{\sum_{i=1}^{k}v_iA_\theta(t_i)} \qquad (3.6)
$$

Upon plugging $\hat{\alpha}_1$ and $\hat{\alpha}_2$ into Eq.(3.5), we immediately have

$$
\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\theta}=\frac{k}{\theta}+\sum_{i=1}^{k}t_i-\frac{k_1\sum_{i=1}^{k}u_it_ie^{\theta t_i}}{\sum_{i=1}^{k}u_iA_\theta(t_i)}-\frac{k_2\sum_{i=1}^{k}v_it_ie^{\theta t_i}}{\sum_{i=1}^{k}v_iA_\theta(t_i)}=0 \qquad (3.7)
$$

It can be checked that for given θ, the log-likelihood function in (3.2) is unimodal. This in turns out the estimates given in (3.6) are the MLEs of α_1 and α_2. Further, for given $\alpha_j(j=1,2)$, $\frac{\partial l}{\partial\theta}$ is a monotone decreasing function starting from ∞ at 0 to a negative constant when $\theta\longrightarrow\infty$. By using this fact and $\frac{\partial^2 l}{\partial\theta^2}<0$, we conclude that the MLE of θ exists and unique. The non-linear equation (3.7) cannot be solved analytically and a numerical method such as Newton-Raphson method can be applied. Once the MLE of θ, $\hat{\theta}$ is coputed, the MLEs $\hat{\alpha}_1$ and $\hat{\alpha}_2$ are obtained directly using (3.6). For special cases where $R_i=0$, for all $i=1,2,...,k$, and then $S_i=W_i=0$, the tems $\sum_{i=1}^{k}S_iA_\theta(t_i)$, and $\sum_{i=1}^{k}S_it_ie^{\theta t_i}$ in (3.3), (3.4), and (3.7)reduce to 0. This in turns out that the likelihood equations become

$$
\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_1}=\frac{k_1}{\alpha_1}-\sum_{i=1}^{k}\delta_iA_\theta(t_i)=0,
$$

$$
\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_2}=\frac{k_2}{\alpha_2}-\sum_{i=1}^{k}(1-\delta_i)A_\theta(t_i)=0,
$$

and

$$\frac{\partial l(\alpha_1,\alpha_2,\theta)}{\partial\theta} = \frac{k}{\theta} + \sum_{i=1}^{k} t_i - \alpha_1 \sum_{i=1}^{k} \delta_i t_i e^{\theta t_i} - \alpha_2 \sum_{i=1}^{k} (1-\delta_i) t_i e^{\theta t_i} = 0.$$

It follows that

$$\hat{\alpha}_1 = \frac{k_1}{\sum_{i=1}^{k} \delta_i A_\theta(t_i)}, \hat{\alpha}_2 = \frac{k_2}{\sum_{i=1}^{k} (1-\delta_i) A_\theta t_i},$$

and

$$\hat{\theta} = \frac{k}{\sum_{i=1}^{k} t_i - \hat{\alpha}_1 \sum_{i=1}^{k} \delta_i t_i e^{\theta t_i} - \hat{\alpha}_2 \sum_{i=1}^{k} (1-\delta_i) t_i e^{\theta t_i}}.$$

Clearly, the MLEs are not expressed in closed forms and their respective variaces cannot be obtained. Because of that, we propose to use the asymptotic variances of $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\theta}$. From the log-likelihood function in Eq(3.2), we have

$$\begin{aligned}
\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_1^2} &= -\frac{k_1}{\alpha_1^2} \\
\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_1\partial\alpha_2} &= -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_2\partial\alpha_1} = 0, \\
\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_2^2} &= -\frac{k_2}{\alpha_2^2}, \\
\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_1\partial\theta} &= \frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\theta\partial\alpha_1} = -\sum_{i=1}^{k} u_i t_i e^{\theta t_i}, \\
\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\alpha_2\partial\theta} &= \frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\theta\partial\alpha_2} = -\sum_{i=1}^{k} v_i t_i e^{\theta t_i}, \\
\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial\theta^2} &= -\frac{k}{\theta^2} - \alpha_1 \sum_{i=1}^{k} u_i t_i^2 e^{\theta t_i} - \alpha_2 \sum_{i=1}^{k} v_i t_i^2 e^{\theta t_i}. \qquad (3.8)
\end{aligned}$$

To this end, let us consider $Q = (\alpha_1, \alpha_2, \theta)$. Under the usual regularity conditions (Lehmann and Casella [20] and large k, the asymptotic normality of $\hat{Q}$=$(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\theta})$ that based on the convergence in distribution ($\xrightarrow{\text{D}}$) of $\hat{Q}$ can be described as

$$\hat{Q} \xrightarrow{\text{D}} N_3(Q, I^{-1}(Q)),$$

where $I^{-1}(Q)$ is the inverse of the Fisher information matrix of the unknown parameters Q [we find it by taking the negative expectation of the expressions in (3.8)]. In partical applications, one may use the approximation,

$$\hat{Q} \xrightarrow{\text{D}} N_3(Q, J^{-1}(Q)),$$

denotes the inverse of the observed information matrix of Q. Hence, we can define the observed information matrix $J(Q)$ by taking the negative of the expressions in (3.8)as follows:

$$\begin{aligned} J(Q) &= \begin{pmatrix} -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \alpha_1^2} & -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \alpha_1 \partial \alpha_2} & -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \alpha_1 \partial \theta} \\ -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \alpha_2 \partial \alpha_1} & -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \alpha_2^2} & -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \alpha_2 \partial \theta} \\ -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \theta \partial \alpha_1} & -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \theta \partial \alpha_2} & -\frac{\partial^2 l(\alpha_1,\alpha_2,\theta)}{\partial \theta^2} \end{pmatrix} \qquad (3.9) \\ &= -\begin{pmatrix} \frac{k_1}{\alpha_1^2} & 0 & \sum_{i=1}^{k} u_i t_i e^{\theta t_i} \\ 0 & \frac{k_2}{\alpha_2^2} & \sum_{i=1}^{k} v_i t_i e^{\theta t_i} \\ \sum_{i=1}^{k} u_i t_i e^{\theta t_i} & \sum_{i=1}^{k} v_i t_i e^{\theta t_i} & \frac{k}{\theta^2} - \alpha_1 \sum_{i=1}^{k} u_i t_i^2 e^{\theta t_i} - \alpha_2 \sum_{i=1}^{k} v_i t_i^2 e^{\theta t_i} \end{pmatrix}. \end{aligned}$$

Therefore, the $100(1-\gamma)$ CIs for α_1, α_2 and θ are

$$(\hat{\alpha}_1 - z_{1-\frac{\gamma}{2}}\sqrt{V_{11}}, \hat{\alpha}_1 + z_{1-\frac{\gamma}{2}}\sqrt{V_{11}}), (\hat{\alpha}_2 - z_{1-\frac{\gamma}{2}}\sqrt{V_{22}}, \hat{\alpha}_2 + z_{1-\frac{\gamma}{2}}\sqrt{V_{22}}),$$

and

$$(\hat{\theta} - z_{1-\frac{\gamma}{2}}\sqrt{V_{33}}, \hat{\theta} + z_{1-\frac{\gamma}{2}}\sqrt{V_{33}}),$$

respectively, where V_{11}, V_{12} and V_{33} are the elements of the main diagonal of $J^{-1}(\hat{Q})$ and z_γ is $100\gamma^{th}$ percentile of the standard normal distribution. Usually, the CI based on the asymptotic results do not perform quite well for small sample size. For this reason, it is more appropriate to propose alternative CI, say, Bootstrap-t method (Boot-t method), see for example, Ahmed [1]. The following algorithm describes the steps for obtaining Boot-t CIs.

- **Step 1:** Estimate α_1, α_2 and θ using the maximum likelihood based on the observed informative sample (say $\hat{\alpha}_1, \hat{\alpha}_2$ and $\hat{\theta}$).

- **Step 2:** Using $\hat{\alpha}_1,\ \hat{\alpha}_2$ and $\hat{\theta}$ obtained in **Step 1**, generate a Bootstrap sample and then obtain the first k observed censored units, $B_1, B_2, ..., B_k$ under the GO model. Then compute the corrresponding MLEs $\hat{\alpha}_1^*$, $\hat{\alpha}_2^*$ and $\hat{\theta}^*$ of $\alpha_1,\ \alpha_2$ and θ and the elements $(V_{11}^*, V_{22}^*, V_{33}^*)$ of the main diagonal of $J^{*-1}(\alpha_1, \alpha_2, \theta)$.

- **Step 3:** Based on the Bootstrap sample in **Step 2**, define an estimated version

$$Q_1^* = \frac{\hat{\alpha}_1^* - \hat{\alpha}_1}{\sqrt{V_{11}^*}},\ Q_2^* = \frac{\hat{\alpha}_2^* - \hat{\alpha}_2}{\sqrt{V_{22}^*}} \text{ and } Q_3^* = \frac{\hat{\theta}^* - \hat{\theta}}{\sqrt{V_{33}^*}}$$

- **Step 4:** Generate M=1000 Bootstrap samples and vesions of Q_1^*, Q_2^* and Q_3^* then we obtain the $100\gamma^{th}$ sample quatiles of Q_1^*, Q_2^* and Q_3^* (say $q_{1,\gamma}^*$, $q_{2,\gamma}^*$ and $q_{3,\gamma}^*$.

- **Step 5:**Compute the approximate $100(1-\gamma)$ CIs for α_1, α_2 and θ as

$$(\hat{\alpha}_1 - q_{1,1-\frac{\gamma}{2}}^*\sqrt{V_{11}}, \hat{\alpha}_1 + q_{1,\frac{\gamma}{2}}^*\sqrt{V_{11}}), (\hat{\alpha}_2 - q_{2,1-\frac{\gamma}{2}}^*\sqrt{V_{22}}, \hat{\alpha}_2 + q_{2,\frac{\gamma}{2}}^*\sqrt{V_{22}}),$$

and

$$(\hat{\theta} - q_{3,1-\frac{\gamma}{2}}^*\sqrt{V_{33}}, \hat{\theta} + q_{1-\frac{3,\gamma}{2}}^*\sqrt{V_{33}}),$$

3.2 Bayesian Estimation of the Parameters

In this section, we formulate the posterior densities of the parameters α_1, α_2, and θ based on JPC sample coming from from two-parameter GO distribution and then obtain the corresponding Bayes estimators (BEs) of these unknown parameters as well as the credible intervals (CrIs) under different error loss functions, L_1 and L_2, with respect to the priors as described in Chapter 1.
Now, we specify the prior distributions of α_1, α_2, and θ. It is desirable that the model parameters are independent such that all prior and posterior densities belong to similar families. These prior choices allow the posterior distribution to be analytically tractable and computationally efficient. A natural choice for the priors of α_1, α_2, and θ would be to assume that the three quantities are independent gamma $G(a_i, b_i), i = 1, 2$ and $G(a_0, b_0)$ distributions, respectively, with the following densities:

$$g_{a_i,b_i}(\alpha_i) = \frac{b_i^{a_i}}{\Gamma(a_i)}\, \alpha_i^{a_i-1} e^{-b_i\alpha_i},\ i = 1, 2, \tag{3.10}$$

and

$$g_{a_0,b_0}(\theta) = \frac{b_0^{a_0}}{\Gamma(a_0)}\, \theta^{a_0-1} e^{-b_0\theta}, \tag{3.11}$$

where α_i, and $\theta > 0$ and a_i, b_i, and a_0, b_0, are chosen to reflect prior knowledge about α_i and θ.

By combining the prior distributions given in (3.10) and (3.11) and likelihood function in (3.1), we can express the joint posterior density of α_1, α_2 and θ as

$$\pi\left(\alpha_1, \alpha_2, \theta \mid \boldsymbol{t}\right) \propto \alpha_1^{k_1+a_1-1} e^{-\alpha_1\left(b_1+\sum_{i=1}^{k} u_i A_\theta(t_i)\right)} \times \alpha_2^{k_2+a_2-1} e^{-\alpha_2\left(b_2+\sum_{i=1}^{k} v_i A_\theta(t_i)\right)}$$
$$\times\, \theta^{k+a_0-1} e^{-\theta\left(b_0-\sum_{i=1}^{k} t_i\right)},$$

by setting

$$\Delta_\theta(\mathbf{t}) = \sum_{i=1}^{k} u_i A_\theta\left(t_i\right) \text{ and } \bar{\Delta}_\theta(\tilde{t}) = \sum_{i=1}^{k} v_i A_\theta\left(t_i\right)$$

we may rewrite the joint posterior distribution as

$$\pi\left(\alpha_1, \alpha_2, \theta \mid \mathbf{t}\right) \propto \mathbf{p}_1\left(\alpha_1 \mid \theta, \mathbf{t}\right) \times \mathbf{p}_2\left(\alpha_2 \mid \theta, \mathbf{t}\right) \times \mathbf{p}_3(\theta \mid \mathbf{t}) \tag{3.12}$$

where $\mathbf{p}_1(\alpha_1 \mid \theta, \mathbf{t})$, and $\mathbf{p}_2(\alpha_2 \mid \theta, \mathbf{t})$, are PDFs of $G(k_1 + a_1, b_1 + \Delta_\theta(\mathbf{t}))$, and $G\left(k_2 + a_2, b_2 + \bar{\Delta}_\theta(\mathbf{t})\right)$, respectively, while $\mathbf{p}_3(\theta \mid \mathbf{t})$ is defined by

$$\mathbf{p}_3(\theta \mid \mathbf{t}) = \mathbf{g}_1(\theta \mid \mathbf{t}) \times \mathbf{g_2}(\theta, \mathbf{t}) \tag{3.13}$$

with $\mathbf{g}_1(\theta \mid \mathbf{t})$ being the PDF of $G(k + a_0, b_0)$ and

$$\mathbf{g}_2(\theta, \mathbf{t}) = \frac{e^{\theta \sum_{i=1}^{k} t_i}}{(b_1 + \Delta_\theta(\mathbf{t}))^{k_1 + a_1} \left(b_2 + \bar{\Delta}_\theta(\mathbf{t})\right)^{k_2 + a_2}}.$$

Under the error loss functions, L_1 and L_2, the BE of any function of α_1, α_2 and θ (say, $\lambda(\alpha_1, \alpha_2, \theta)$), respectively, takes the following form:

$$\hat{\lambda}_S(\alpha_1, \alpha_2, \theta) = \frac{\int_0^\infty \int_0^\infty \int_0^\infty \lambda(\alpha_1, \alpha_2, \theta)\, \mathbf{p}_1(\alpha_1 \mid \theta, \mathbf{t})\, \mathbf{p}_2(\alpha_2 \mid \theta, \mathbf{t})\, \mathbf{p_3}(\theta \mid \mathbf{t})\; d\alpha_1\, d\alpha_2\, d\theta}{\int_0^\infty \int_0^\infty \int_0^\infty \mathbf{p}_1(\alpha_1 \mid \theta, \mathbf{t})\, \mathbf{p_2}(\alpha_2 \mid \theta, \mathbf{t})\, \mathbf{p_3}(\theta \mid \mathbf{t})\; d\alpha_1\, d\alpha_2\, d\theta} \tag{3.14}$$

and

$$\hat{\lambda}_L(\alpha_1, \alpha_2, \theta) = -\frac{1}{v} \log\left(\frac{\int_0^\infty \int_0^\infty \int_0^\infty e^{-v\lambda(\alpha_1, \alpha_2, \theta)} \mathbf{p}_1(\alpha_1 \mid \theta, \mathbf{t})\, \mathbf{p_2}(\alpha_2 \mid \theta, \mathbf{t})\, \mathbf{p_3}(\theta \mid \mathbf{t})\; d\alpha_1\, d\alpha_2\, d\theta}{\int_0^\infty \int_0^\infty \int_0^\infty \mathbf{p}_1(\alpha_1 \mid \theta, \mathbf{t})\, \mathbf{p}_2(\alpha_2 \mid \theta, \mathbf{t})\, \mathbf{p_3}(\theta \mid \mathbf{t})\; d\alpha_1\, d\alpha_2\, d\theta}\right) \tag{3.15}$$

It is clear to note that Eq.'s (3.14) and (3.15) cannot be simplified into closed-form expressions and produce the BEs of α_1, α_2 and θ and must be calculated numerically. Among the various methods suggested to approximate the joint posterior density, alternative method called importance sampling will be proposed to provide sample based estimates of the parameters.

3.2.1 Bayesian Sample Based Methods

In light of the JPC sample, one can update the prior information about the shape and scale parameters via the posterior model, it is possible to provide approximate BEs of α_1, α_2 and θ based on importance sampling technique. In the context of the approach introduced by Geman and Geman [12], we generate the posterior distribution of α_1, α_2 and θ based on the conditional arguments as follows:

Algorithm of Bayesian Sample-Based Method:

-Step 1: Generate θ from $\mathbf{g}_1(\theta \mid \mathbf{t})$.

-Step 2: Generate α_1 from $\mathbf{p}_1(\alpha_1 \mid \theta, \mathbf{t})$ and α_2 from $\mathbf{p}_2(\alpha_2 \mid \theta, \mathbf{t})$.

-Step 3: Repeat steps 1-3, M times and then obtain $(\theta_1, \alpha_{11}, \alpha_{21}), \ldots, (\theta_M, \alpha_{1M}, \alpha_{2M})$.

-Step 4: Compute

$$w_i(\theta_i, \alpha_{1i}, \alpha_{2i}) = \frac{\mathbf{g}_2(\theta_i, \mathbf{t})}{\sum_{i=1}^{M} \mathbf{g}_2(\theta_i, \mathbf{t})}$$

and then compute an approximate BE of a function of α_1, α_2 and θ (say, $\lambda(\alpha_1, \alpha_2, \theta)$) under a SE loss function as

$$\hat{\lambda}_S = \sum_{i=1}^{M} w_i(\alpha_{1i}, \alpha_{2i}, \theta_i)\, \lambda(\alpha_{1i}, \alpha_{2i}, \theta_i).$$

The other Bayesian estimate under L_2 can be computed readily based on the following expression:

$$\hat{\lambda}_L = -\frac{1}{v} \log\left[\sum_{i=1}^{M} w_i(\alpha_{1i}, \alpha_{2i}, \theta_i)\, e^{-v\lambda(\alpha_{1i}, \alpha_{2i}, \theta_i)}\right],$$

Using the percentile-based argument used in Chen and Shao [7], we can provide CrI of any function of the parameters, $\lambda(\alpha_1, \alpha_2, \theta)$. The 100γ th $(0 < \gamma < 1)$ quantile of λ is λ_γ such that $P(\lambda \leq \lambda_\gamma) = \gamma$. Let us assume that $\lambda_{(1)}, \ldots, \lambda_{(M)}$ be order statistics of $\lambda_1, \ldots, \lambda_M$ and $w^{(1)}, \ldots, w^{(M)}$ be the values associated with $\lambda_{(1)}, \ldots, \lambda_{(M)}$. This implies that the consistent sample based estimate of λ_γ is $\hat{\lambda}_\gamma = \lambda_{(\kappa)}$, where κ is an integer satisfying

$$\sum_{i=1}^{\kappa-1} w^{(i)} \leq \gamma \leq \sum_{i=1}^{\kappa} w^{(i)}$$

This in turns that $(1-\gamma)100\%$CrI of λ can be computed as $\left(\hat{\lambda}_{\frac{\gamma}{2}}, \hat{\lambda}_{1-\frac{\gamma}{2}}\right)$, say, C-S CrI. It is important to point out that the CrIs obtained by the previous approach does not specify whether the values of λ within these intervals have highest probability than that of the values outside the intervals. It is more desirable to have CrI of $\lambda\,(\alpha_1, \alpha_2, \lambda)$ with the highest posterior density (HPD). For any probability content, $1-\gamma$, the HPD interval is of the shortest width and the posterior density for every point outside the interval is less than that for every point inside the interval. For M sufficiently large, the $100(1-\gamma)\%$ HPD interval for λ may be chosen as the shortest of the intervals $C_\kappa(M), \kappa = 1, 2, \ldots, M - [(1-\gamma)M]$, where $[x]$ is the largest integer that is less than or equal to x, with

$$C_\kappa(M) = \left(\lambda_{(\kappa)}, \lambda_{(\kappa+[(1-\gamma)M])}\right)$$

Therefore, the HPD intervals of the three parameters are computed in this way.

Another Bayesian sampling algorithm is to estimate the posterior distribution based on Gibbs sampler approach. This approach requires being able to sample from the full conditional distributions from each parametric quantity involved. This can be applied for α_1 and α_2 but not for θ.

3.2.2 Metropolis-Hastings Method

Metropolis-Hastings (M-H) steps are proposed into the Gibbs sampler so that α_1 and α_2 are sampled directly from their full conditional distributions, whereas θ can be updated via a M-H steps as explained in Tierney [30], using $G\,(k + a_0, b_0)$ as a proposal distribution. The M-H steps proceed as follows:

Algorithm of M-H method:

-Step 1: Select an initial guess θ_0;

-Step 2: For $t = 1, 2, \ldots, M$, repeat:

(a) Draw candidate θ^* from $G\,(k + a_0, b_0)$ with its pdf $g_{k+a_0,b_1}\,(\theta^* \mid \theta_{t-1})$ and u from $U(0, 1)$;

(b) Compute the acceptance probability

$$\varepsilon = \frac{\mathbf{p}_3(\theta^*)/g_{k+a_0,b_1}(\theta^*|\theta_{t-1})}{\mathbf{p}_3(\theta_{t-1})/g_{k+a_0,b_0}(\theta_{t-1}|\theta^*)} = \frac{\mathbf{p}_3(\theta^*)}{\mathbf{p}_3(\theta_{t-1})}\frac{g_{k+a_0,b_0}(\theta_{t-1})}{g_{k+a_0,b_0}(\theta^*)} = h\,(\theta^*, \theta_{t-1}) \times e^{(\theta^*-\theta_{t-1})\sum_{i=1}^{k} t_i}$$

where

$$h\left(\theta^{*}, \theta_{t-1}\right)=\frac{\left(b_{1}+\Delta_{\theta_{t-1}}(\mathbf{t})\right)^{k_{1}+a_{1}}\left(b_{2}+\bar{\Delta}_{\theta_{t-1}}(\mathbf{t})\right)^{k_{2}+a_{2}}}{\left(b_{1}+\Delta_{\theta^{*}}(\mathbf{t})\right)^{k_{1}+a_{1}}\left(b_{2}+\bar{\Delta}_{\theta^{*}}(\mathbf{t})\right)^{k_{2}+a_{2}}}$$

(c) If $u < \min(1, \varepsilon)$, then set $\theta_t = \theta^*$, else go to (a).

Once the posterior samples have been obtained, the simulation consistent Bayes estimates under different loss functions can be computed. The associated CrIs of the parameters can also be constructed.

CHAPTER 4

SIMULATION STUDY

Here in this chapter specially in Section 4.1, we perform a comprehensive simulation study to assess the performance of the estimates developed in the previous chapters. Also in Section 4.2, we analyze a real data set using GO distribution for illustrative purposes. All computations are performed using R software.

4.1 Results for Two-Parameter Gompertz Model

In this section, we present results for the simulation study including JPC based on two-parameter GO model. Where we compare the performance of the different methods of estimation based on Monte Carlo simulation. We compare performance of the MLEs and BEs in terms of Bias and mean square error (MSE). The Bias and MSE of an estimate $\hat{\theta}$ of θ are defined in Chapter 1 as:

$$Bias(\hat{\theta}) = \frac{1}{I}\sum_{i=1}^{n}\left(\hat{\theta}_i - \theta_i\right), \text{ and } MSE(\hat{\theta}) = \frac{1}{I}\sum_{i=1}^{n}(\hat{\theta}_i - \theta_i)^2,$$

where I is the number of iterations in the simulation process. In this simulation, the values of GO parameters are considered as $\alpha_1 = 2$, $\alpha_2 = 1.5$ and $\theta = 2$. Here, we consider different effective sample sizes, $k = 20, 25$ and different censoring schemes. For conducting the Bayesian analysis, we assume two priors. For the first prior (Prior 0), we assume that $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$. Then, we assume an additional prior; Prior 1 with $a_1 = a_2 = a_3 = 1$, $b_1 = b_2 = b_3 = 2$. We use the following notation for a particular progressive censoring scheme. For example, $k = 6$ and $\boldsymbol{R} = (4, 0_{(5)})$ means $R_1 = 4$, $R_2 = R_3 = R_4 = R_5 = R_6 = 0$.

To conduct the comparison, we have randomly generated four different joint progressive censored schemes from the GO distribution for $k = 20$ and for $k = 25$, summarized in Table 4.1. We then compute Biases, MSEs for the estimates and CIs for GO parameters. Under the non-informative and informative priors, Biases and MSEs of the estimates over 10000 replications are computed for various censoring schemes of $k = 20$ and $k = 25$.

Table 4.1: Censoring schemes.

Notation	k	C.S.
R_1	20	$R = (7, \underbrace{0}_{18\ times}, 15)$
R_2	20	$R = (\underbrace{0}_{18\ times}, 7, 15)$
R_3	25	$R = (7, \underbrace{0}_{23\ times}, 10)$
R_5	25	$R = (\underbrace{0}_{11\ times}, 7, \underbrace{0}_{12\ times}, 10)$

Biases and MSEs of the MLEs and BEs of α_1, α_2 and θ under SE and LINEX loss functions with Prior 0 and Prior 1 are computed over 10000 replications and displayed in Tables 4.2 and 4.3. MLEs are computed by maximizing the likelihood function and so by solving the likelihood Equations (3.6) and (3.7).

For the LINEX loss function, the BEs are computed using different values of ν (say, $\nu = -0.5, 0.5$).

Table 4.2: Biases and MSEs of the MLEs and BEs of α_1, α_2 and θ under Prior 0.

C. S.		MLE		BEs (SE)	BEs (LINEX)	
					$\nu=-0.5$	$\nu=0.5$
R_1	α_1	-2.2299(2.21170)	Importance	-1.1463(1.0433)	-0.7778(0.6839)	-0.8332(0.7366)
			M-H	-0.9865(0.9551)	-0.7496(0.6056)	-0.7945(0.6948)
	α_2	-0.8938(0.3016)	Importance	0.1475(0.0237)	0.1138(0.0199)	0.1211(0.0220)
			M-H	0.1275(0.0223)	0.1129(0.0159)	0.1184(0.0192)
	θ	2.3451(1.0226)	Importance	-0.5623(0.3181)	-0.5614(0.3171)	-0.5615(0.3172)
			M-H	-0.5493(0.3043)	-0.5468(0.3017)	-0.5472(0.3021)
R_2	α_1	-2.8802(2.6612)	Importance	-1.9933(1.4117)	-1.4519(1.2045)	-1.5022(1.2253)
			M-H	-1.6882(1.2854)	-1.4295(1.1707)	-1.4511(1.1818)
	α_2	2.0606(0.3579)	Importance	0.2365(0.0677)	0.2138(0.0613)	0.2199(0.0647)
			M-H	0.1831(0.0339)	0.1581(0.0253)	0.1717(0.0298)
	θ	2.7829(1.2112)	Importance	-0.5688(0.3256)	-0.5681(0.3248)	-0.5682(0.3249)
			M-H	-0.5509(0.3062)	-0.5487(0.3038)	-0.5491(0.3042)
R_3	α_1	-2.00034(1.9505)	Importance	-1.0891(0.9215)	-0.6279(0.4572)	-0.6510(0.5651)
			M-H	-0.7669(0.6280)	-0.2102(0.1626)	-0.3279(0.3204)
	α_2	-0.5653(0.2172)	Importance	0.1034(0.0147)	0.0648(0.0124)	0.0750(0.0136)
			M-H	0.0759(0.0109)	0.0608(0.0058)	0.0693(0.0085)
	θ	2.2216(0.9134)	Importance	-0.4122(0.1725)	-0.4116(0.1720)	-0.4117(0.1721)
			M-H	-0.4006(0.1644)	-0.3968(0.1615)	-0.3975(0.1621)
R_4	α_1	-2.4630(1.9543)	Importance	-1.2816(1.1374)	-0.9002(0.8845)	-0.9420(0.9514)
			M-H	-1.1971(1.0740)	-0.8814(0.8097)	-0.9345(0.9070)
	α_2	0.5822(0.2396)	Importance	0.1076(0.0187)	0.0830(0.0156)	0.0949(0.0173)
			M-H	0.0987(0.0118)	0.0790(0.0065)	0.0919(0.0092)
	θ	2.3007(0.9792)	Importance	-0.4174(0.1768)	-0.4168(0.1764)	-0.4169(0.1765)
			M-H	-0.4040(0.1670)	-0.4007(0.1644)	-0.4013(0.1649)

* The entries in parentheses are MSEs.

Table 4.3: Biases and MSEs of the MLEs and BEs of α_1, α_2 and θ under Prior 1.

C. S.			BEs (SE)	BEs (LINEX)	
				$\nu = -0.5$	$\nu = 0.5$
R_1	α_1	Importance	-1.1289(1.0428)	-0.7405(0.6344)	-0.8273(0.7128)
		M-H	-0.9552(0.7871)	-0.7311(0.5352)	-0.7921(0.6871)
	α_2	Importance	0.1145(0.0188)	0.1041(0.0137)	0.1126(0.0160)
		M-H	0.0869(0.0160)	0.0753(0.0129)	0.0819(0.0149)
	θ	Importance	-0.5586(0.3169)	-0.5293(0.3164)	-0.5348(0.3165)
		M-H	-0.5363(0.2910)	-0.5172(0.2836)	-0.5305(0.2847)
R_2	α_1	Importance	-1.8672(1.3662)	-1.2824(1.1321)	-1.3197(1.1485)
		M-H	-1.6128(1.2472)	-1.1025(1.0841)	-1.2163(1.0960)
	α_2	Importance	0.2273(0.0608)	0.2109(0.0531)	0.2184(0.0573)
		M-H	0.1558(0.0249)	0.1311(0.0177)	0.1446(0.0215)
	θ	Importance	-0.5638(0.3201)	-0.5633(0.3196)	-0.5634(0.3197)
		M-H	-0.5498(0.3049)	-0.5458(0.3006)	-0.5464(0.3013)
R_3	α_1	Importance	-0.9891(0.9186)	-0.6173(0.3998)	-0.6489(0.5071)
		M-H	-0.6449(0.5509)	-0.2120(0.1608)	-0.2170(0.2581)
	α_2	Importance	0.0631(0.0087)	0.0464(0.0065)	0.0558(0.0077)
		M-H	0.0449(0.0023)	0.0208(0.0007)	0.0344(0.0015)
	θ	Importance	-0.3964(0.1602)	-0.3960(0.1598)	-0.3961(0.1599)
		M-H	0.3833(0.1507)	0.3810(0.1490)	0.3814(0.1493)
R_4	α_1	Importance	-1.2632(1.1237)	-0.8684(0.7944)	-0.9130(0.8753)
		M-H	-1.2099(1.0621)	-0.7549(0.7474)	-0.8342(0.8061)
	α_2	Importance	0.0869(0.0091)	0.0609(0.0068)	0.0754(0.0080)
		M-H	0.0636(0.0078)	0.0461(0.0039)	0.0561(0.0059)
	θ	Importance	-0.3969(0.1608)	-0.3966(0.1605)	-0.3967(0.1606)
		M-H	-0.3834(0.1513)	-0.3814(0.1480)	-0.3818(0.1494)

* The entries in parentheses are MSEs.

As seen in Tables 4.2 and 4.3, the BEs perform well for $k = 20$ and $k = 25$ in the sense of Bias and MSE. As expected, the BEs under Prior 1 are better than under Prior 0. Further, we can easily notice that the M-H method beats the importance sampling method in the sense of MSEs and Biases for all parameters under all two error loss functions. Under LINEX error loss function, the estimates based on M-H algorithm show the least MSEs and Biases values over SE error loss function in all parameters and all methods of estimation.

Table 4.4: ALs and CPs of 95% approximate and Boot-t CIs of α_1, α_2 and λ when $m = 20$ and $n = 25$.

C. S.			Approx.	Boot-t
R_1	α_1	AL	1.8292	1.6115
		CP	0.6285	0.6652
	α_2	AL	1.3666	0.7526
		CP	0.6126	0.6865
	θ	AL	1.5442	1.3756
		CP	0.6153	0.7111
R_2	α_1	AL	2.0057	1.8869
		CP	0.7180	0.7560
	α_2	AL	1.5039	1.0467
		CP	0.7126	0.7622
	θ	AL	1.8848	1.7137
		CP	0.7112	0.7153
R_3	α_1	AL	1.6714	1.4578
		CP	0.5395	0.6308
	α_2	AL	0.9944	0.7411
		CP	0.5472	0.5619
	θ	AL	0.9153	0.8792
		CP	0.5429	0.5486
R_4	α_1	AL	1.6867	1.5641
		CP	0.5468	0.6432
	α_2	AL	1.0007	0.7412
		CP	0.5407	0.5808
	θ	AL	1.2853	0.9249
		CP	0.5479	0.6301

Table 4.5: ALs and CPs of 95% CIs of α_1, α_2 and θ when $m = 20$ and $n = 25$ under Prior 0.

C. S.			HPD (Importance)			HPD M-H		
			SE	LINEX		SE	LINEX	
				$\nu = -0.5$	$\nu = 0.5$		$\nu = -0.5$	$\nu = 0.5$
R_1	α_1	AL	0.9935	0.7431	0.9156	0.1251	0.1097	0.1126
		CP	0.8424	0.8957	0.8941	0.9502	0.9721	0.9683
	α_2	AL	0.3610	0.3559	0.3566	0.0660	0.0620	0.0615
		CP	0.8955	0.9124	0.9107	0.9760	0.9808	0.9806
	θ	AL	0.1935	0.1909	0.1923	0.1722	0.1716	0.1718
		CP	0.8571	0.8633	0.8617	0.9603	0.9754	0.9640
R_2	α_1	AL	1.8401	1.1961	1.4714	0.1689	0.1215	0.1326
		CP	0.9090	0.9157	0.9140	0.9725	0.9847	0.9805
	α_2	AL	0.4195	0.4102	0.4166	0.1005	0.0967	0.0983
		CP	0.9074	0.9141	0.9124	0.9877	0.9887	0.9885
	θ	AL	0.2357	0.2330	0.2344	0.2298	0.2294	0.2295
		CP	0.8940	0.9160	0.9040	0.9800	0.9807	0.9805
R_3	α_1	AL	0.7672	0.5516	0.7318	0.1219	0.0556	0.0849
		CP	0.8243	0.8743	0.8727	0.8608	0.9107	0.8957
	α_2	AL	0.3579	0.3556	0.3563	0.0631	0.0607	0.0608
		CP	0.8752	0.8955	0.8921	0.9563	0.9606	0.9605
	θ	AL	0.1907	0.1887	0.1898	0.1692	0.1686	0.1687
		CP	0.8433	0.8617	0.8602	0.9600	0.9643	0.9607
R_4	α_1	AL	1.0998	0.8349	0.9685	0.1333	0.0972	0.1078
		CP	0.8608	0.8974	0.8941	0.9527	0.9730	0.9720
	α_2	AL	0.3672	0.3566	0.3599	0.0751	0.0699	0.0703
		CP	0.8888	0.8955	0.8938	0.9600	0.9685	0.9683
	θ	AL	0.2184	0.2156	0.2171	0.1775	0.1771	0.1772
		CP	0.8453	0.8639	0.8612	0.9680	0.9688	0.9687

Table 4.6: ALs and CPs of 95% CIs of α_1, α_2 and θ when $m = 20$ and $n = 25$ under Prior 1.

C. S.			HPD (Importance)			HPD M-H		
			SE	LINEX		SE	LINEX	
				$\nu = -0.5$	$\nu = 0.5$		$\nu = -0.5$	$\nu = 0.5$
R_1	α_1	AL	0.7496	0.6535	0.6581	0.0940	0.0710	0.0756
		CP	0.8544	0.8727	0.8547	0.9325	0.9447	0.9403
	α_2	AL	0.2748	0.2637	0.2709	0.0609	0.0574	0.0587
		CP	0.8703	0.8812	0.8752	0.9400	0.9408	0.9405
	θ	AL	0.1917	0.1896	0.1906	0.1692	0.1686	0.1687
		CP	0.8243	0.8453	0.8438	0.9400	0.9407	0.9405
R_2	α_1	AL	1.2695	0.9335	1.0856	0.0983	0.0800	0.0840
		CP	0.8576	0.8974	0.8775	0.9405	0.9752	0.9607
	α_2	AL	0.3843	0.3678	0.3809	0.0908	0.0867	0.0876
		CP	0.9498	0.9520	0.9510	0.9600	0.9752	0.9607
	θ	AL	0.2295	0.2266	0.2282	0.2080	0.2069	0.20757
		CP	0.8392	0.8633	0.8438	0.9600	0.9754	0.9605
R_3	α_1	AL	0.7136	0.6014	0.6575	0.0937	0.0636	0.0744
		CP	0.8000	0.8424	0.8226	0.9163	0.9440	0.9241
	α_2	AL	0.2670	0.2630	0.2646	0.0478	0.0458	0.0467
		CP	0.8703	0.8768	0.8752	0.9252	0.9407	0.9405
	θ	AL	0.1829	0.1802	0.1816	0.1675	0.1669	0.1670
		CP	0.7857	0.7913	0.7899	0.9004	0.9207	0.9205
R_4	α_1	AL	0.8861	0.7315	0.8155	0.0945	0.0723	0.0760
		CP	0.8196	0.8608	0.8409	0.9365	0.9561	0.9405
	α_2	AL	0.2927	0.2616	0.2742	0.0697	0.0674	0.0675
		CP	0.8518	0.8955	0.8752	0.9200	0.9407	0.9252
	θ	AL	0.1951	0.1922	0.1938	0.1721	0.1716	0.1718
		CP	0.8035	0.8038	0.8037	0.9200	0.9407	0.9205

Tables 4.4-4.6 present the average lengths (ALs) and coverage probabilities (CPs) of 95% CIs for α_1, α_2 and θ based on Boot-t, asymptotic maximum likelihood and Bayesian methods with Prior 0 and Prior 1 under SE and LINEX loss functions. From Tables 4.4-4.6, it is observed that the HPD CrIs are shorter than the asymptotic and Boot-t CIs under all priors and all loss functions for $k = 20$ and $k = 25$. It can also be noticed that M-H CrIs are the shortest over all other intervals. Furthermore, the CrIs under LINEX loss function are better than SE in all types of estimation used. The performances of HPD CrI tend to be high under the informative prior when compared to HPD CrI under non-informative prior, asymptotic and Boot-t CIs. While the Boot-t method performs well when compared to asymptotic method for estimating of all parameters. It can be also

observed that all CIs are shorter for $k = 25$ when compared to $k = 20$.

In summary, it is clear that the BEs based on importance sampling and M-H methods under different loss functions and priors work better than the MLEs in all the cases considered. MSEs and Biases of the BEs obtained by M-H algorithm are smaller than that of the BEs computed from the importance sampling method under the two types of loss functions. When comparing the BEs under SE and LINEX loss functions, we can notice the LINEX loss function provides better results than SE. It is realized that the HPD intervals based on M-H method compete the ones based on importance sampling method in terms of ALs and CPs criteria under SE and LINEX loss functions. It is also checked that the ALs and CPs of HPD CrIs based on LINEX loss function tend to be close.

4.2 Data Analysis

In this section, we present the analysis of real data sets to illustrate the performance of the obtained methods. The data represent the survival time in months of stage 4 Melanoma patients based on their gender who received treatment at the University of Oklahoma Health Sciences Center from 1974-1978. The data were originally reported by Lee et al. [18] and for more details, see for example Lee and Wang [19]. The sets of data are:

Data set 1 (female patients):

```
1.3  2.7  3.8  4.2  7.4  9.3  10.5  11.4  13.3  13.8 13.8  20
22.2
```

Data set 2 (male patients):

```
0.4  0.9  1.2  1.5  1.6  1.7  2.5  2.5   3.9   3.9  4  4.2 4.5  5.8
5.9  6.3  7.3  7.4  8.3  9.8  11   11.1  16.1  20.5
```

Table 4.7: MLEs, K-S and CvM goodness-of-fit tests.

Data Set	Scale Parameter	Shape Parameter	K-S(p-value)	CvM(p-value)
1	0.9900692	1.0227568	0.13925933(0.9931)	0.03012802(0.899)
2	0.9900692	0.4378163	0.09860765(0.9368)	0.0285854(0.9954)

Table 4.7 presents the ML estimate of the unknown parameters, the goodness-of-fit tests based on Kolmogrov-Smirnov (K-S) and Cramer-von Mises (CvM) statistics. It is easily seen that the GO model fits both data sets very well. This conclusion is also supported by diagnostic plots of the empirical and fitted distribution functions in Figures 4.1 and 4.2. In addition, it is of interest to study the null hypothesis $H_0 : \theta_1 = \theta_2 = \theta$ (i.e., the scale parameters are equal) versus the alternative hypothesis $H_1 : \theta_1 \neq \theta_2$ using the likelihood ratio test. For the given data, the test statistic is computed as $\Delta = L1/L2 = 0.465558$, $-2\log(\Delta) = 1.529034$ and the p-value of the test is $P(\chi^2_{(1)} > 0.465558) = 0.495065$. Hence, the assumption of equality of the scale parameters cannot be rejected.

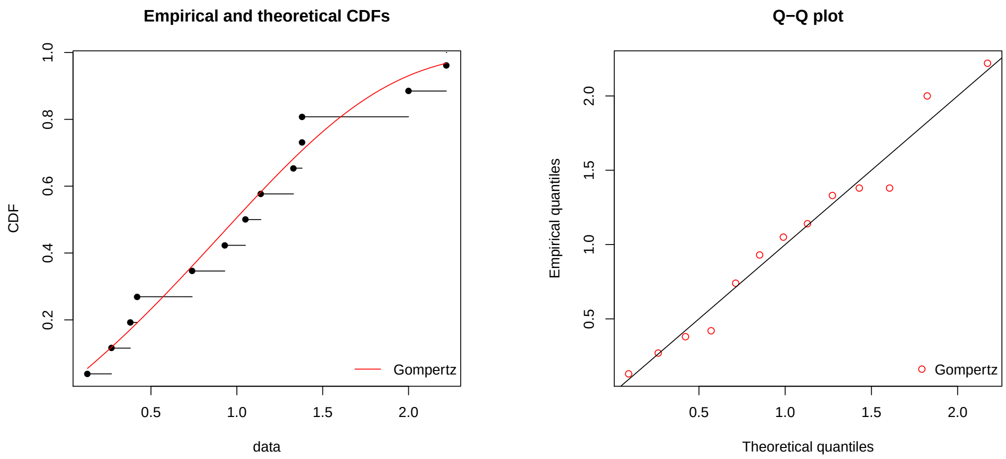

Figure 4.1: Empirical and fitted distribution functions and Q-Q plots for data set 1.

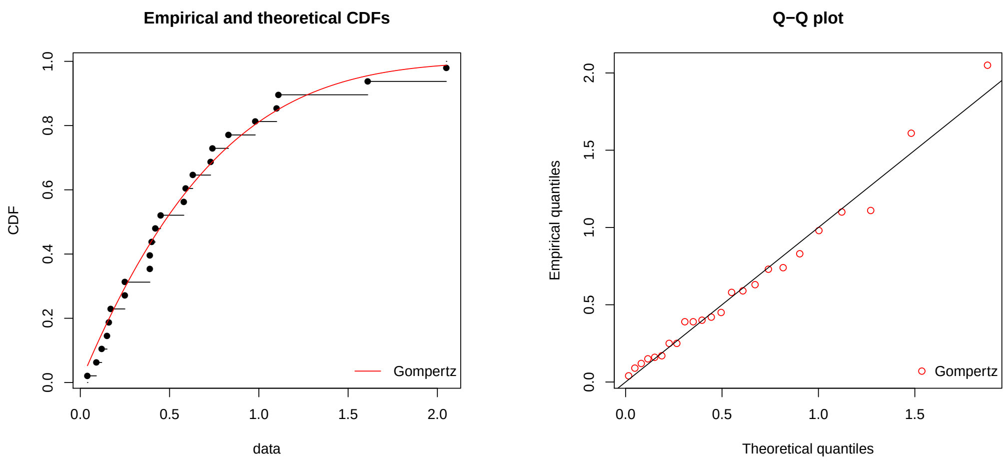

Figure 4.2: Empirical and fitted distribution functions and Q-Q plots for data set 2.

For explanation purposes, we suggest the following joint progressive type-II censored sample with $m = 13$ and $n = 24$, $k = 10$, $R_i = 2$, $i = 1, ..., 3$ and $R_i = 3$, $i = 4, \ldots, 10$. The resulting data set is recorded as follows:

```
(0.4,0,0), (0.9,0,2), (1.2, 0,1), (1.3,1,0), (1.5,0,1), (1.6,0,2),
(1.7,0,1), (2.5,0,1), (2.7,1,1), (3.8,1,2).
```

Based on the above observed joint progressive type II censored data, we obtain the MLEs and BEs of α_1, α_2 and θ under SE and LINEX loss functions. We have generated 50000 observations to compute the BEs of α_1, α_2 and θ based on the importance sampler after discarding the initial 5000 burn-in samples. Note that, for computing the BEs and HPD CrIs, we assume that the priors of α_1, α_2 and θ are improper, i.e. $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$, since we do not have any prior information. The M-H algorithm is also used to compute the BEs of θ where we use the gamma distribution as a proposal distribution. Here, the best fitted model for the full conditional distribution can be concluded by managing the choice of the parameters for the proposal distribution. Therefore, to generate numbers from the target probability distribution, we use the M-H algorithm with gamma proposal distribution. We assumed the initial value of θ to be its MLE, $\hat{\theta}$ which is computed using Newton-Raphson method. Here, we generated 50000 random variates and we checked the acceptance rate for this choice of variance to be 68.27% which is quite satisfactory. We discarded the initial 5000 burn-in samples and computed the BEs based on the remaining observations.

Graphical diagnostics tools involving trace and Autocorrelation function (ACF) plots are used to check the convergence of M-H algorithm.

Figure 4.3 shows the trace and ACF plots for θ. From the trace plot, we can easily observe a random scatter about some mean value represented by a solid line with a fine mixing of the chains for the simulated values of θ. The ACF plot shows that chains have low autocorrelations. As a result, these plots indicate the rapid convergence of the M-H algorithm based on the proposed gamma distribution.

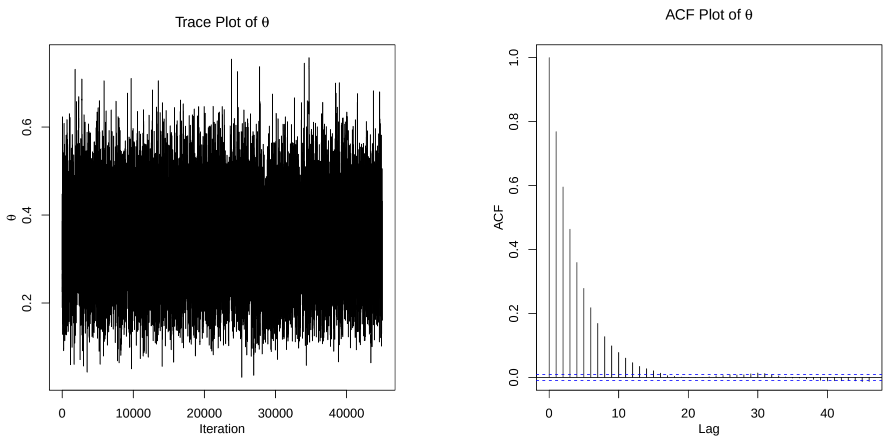

Figure 4.3: Plots of Metropolis-Hastings Markov chains for θ.

The results for MLEs and BEs using importance and M-H samplers along with the 95% Boot-t CI, asymptotic CI and HPD CrIs for α_1, α_2 and θ are presented in Tables 4.8 and 4.9.

Table 4.8: MLEs and BEs of α_1, α_2 and θ based on joint progressive censored data.

	MLE	Method	SE	LINEX	
				ν= -0.5	ν=0.5
α_1	0.1366	Importance	0.1213	0.1216	0.1210
		M-H	0.1685	0.1707	0.1665
α_2	0.2940	Importance	0.2640	0.2648	0.2632
		M-H	0.3274	0.3311	0.3238
θ	0.4498	Importance	0.4515	0.4518	0.4511
		M-H	0.3992	0.3993	0.3991

Table 4.9: The corresponding 95% CIs for α_1, α_2 and θ.

	Approx.	Boot-t	Loss		HPD (Importance)	HPD (M-H)
α_1	(0.0528,0.7261)	(0.0245,0.6258)	SE		(0.0268,0.2650)	(0.1126,0.3113)
			LINEX	ν= - 0.5	(0.0268,0.2050)	(0.0929,0.2332)
				ν= 0.5	(0.0268,0.1925)	(0.0932,0.2213)
α_2	(0.0901,1.2781)	(0.0413,0.9146)	SE		(0.0827,0.3226)	(0.1773,0.3781)
			LINEX	ν= - 0.5	(0.1086,0.3187)	(0.1816,0.3535)
				ν= 0.5	(0.1089,0.2912)	(0.1784,0.3354)
θ	(0.1295,1.1700)	(0.2788,0.7127)	SE		(0.2128,0.5202)	(0.3308,0.5694)
			LINEX	ν= - 0.5	(0.2352,0.4735)	(0.3411,0.5412)
				ν= 0.5	(0.2928,0.4702)	(0.3584,0.5186)

CHAPTER 5

CONCLUSIONS

In this work, the estimation problem of the parameters based on joint type-II progressive censoring scheme when their lifetimes follow Gompertz distributions with different shape parameters. It is shown that the maximum likelihood estimators of the model parameters and their asymptotic confidence intervals can be obtained. We have also proposed different Bayesian procedures for estimate the parameters involved, namely, importance sampling procedure and Metropolis-Hastings algorithm. The corresponding credible intervals are also discussed. The performance of all methods presented in this thesis are evaluated and compared via Monte Carlo simulations. It is observed that the Bayes estimates under Metropolis-Hastings method outperform the frequentist methods as well as the Bayes ones under the importance sampling method in the sense of Bias and mean square error for all parameters under the error loss functions applied. Under LINEX error loss function, the MSEs and Biases of Metropolis-Hastings based estimates tend to be smaller over square error lose. By considering the average length and coverage probability as optimality criteria for the credible intervals of the parameters, it is also evident that the highest posterior density credible intervals using Metropolis-Hastings compete the approximate, Bootstrap-t confidence intervals, and the posterior density credible intervals based on importance sampling. Although, we have mainly restricted our attention to the joint type-II progressive censoring scheme produced from the two populations, but the so developed procedures can be extended to more than two populations as well. More investigation is needed along this line.

REFERENCES

[1] Ahmed, E. A. (2014). Bayesian estimation based on progressive type-II censoring from two-parameter bathtub-shaped lifetime model: an Markov Chain Monte Carlo Approach, *Journal of Applied Statistics*, 41(4), 752-768.

[2] Ahuja, J. C. and Nash, S.W. (1967). The generalized Gompertz-Verhulst family of distributions, *Sankhya*, 29-A, 141156.

[3] Ashour, S.K. and Abo-Kasem, O.E. (2017). Statistical inference for two exponential populations under joint progressive type-I censored scheme, *Communication in Statistics- Theory and Methods*, 46(7), 3479-3488.

[4] Balakrishnan, N. and Aggarwala, R. (2000). *Progressive Censoring: Theory, Methods and Applications*, Birkhauser, Boston.

[5] Balakrishnan, N, Su, F. and Liu, K. Y. (2015). Exact likelihood inference for k exponential populations under joint progressive type-II censoring, *Communications in Statistics-Simulation and Computation*, 44(4), 902-23.

[6] Chen, J. (1997). Parameter estimation of the Gompertz population, *Biometrical Journal*, 39(1), 117-124.

[7] Chen, M. H. and Shao, Q. M. (1999). Monte Carlo estimation of Bayesian credible and HPD intervals, *Journal of Computational and Graphical Statistics*, 8(1), 69-92.

[8] Cohen, A. C. (1963). Progressively censored samples in life testing, *Technometrics*, 5(3), 327-339

[9] Doostparast, M., Vali, M., Ahmadi, M. and Ahmadi, J. (2013). Bayes estimation based on joint progressive type II censored data under LINEX loss function, *Communication in Statistics- Simulation and Computation*, 42(8), 1865-1886.

[10] Durbin, J. (1975). Kolmogorov-Smirnov tests when parameters are estimated with applications to tests of exponentiality and tests on spacings, *Biometrika*, 62(1), 5-22.

[11] Efron, B. (1979). Bootstrap methods: Another look at the jackknife, *The Annals of Statistics*, 7(1), 1-26.

[12] Geman, S. and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the Bayesian restoration of images, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6(6), 721741.

[13] Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration, *Econometrica: Journal of the Econometric Society*, 1317-1339.

[14] Ghitany, M. E., Alqallaf, F. and Balakrishnan, N. (2014). On the likelihood estimation of the parameters of Gompertz distribution based on complete and progressively type-II censored samples, *Statistical Computation and Simulation*, 84(8), 1803-1812.

[15] Gompertz, B. (1825). On the nature of the function expressive of the law of human mortality and on the new mode of determining the value of life contingencies, *Philosophical Transactions of the Royal Society of London*, 115-A, 513580.

[16] Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications, *Biometrika*, 57(1), 97-109.

[17] Ismail, A. A. (2010). Bayes estimation of Gompertz distribution parameters and acceleration factor under partially accelerated life tests with type-I censoring, *Statistical Computation and Simulation*, 80(11), 1253-1264.

[18] Lee, E. T., Ishmael, D. R., Bottomley, R. H., and Murray, J. L. (1982). An Analysis of skin tests and their relationship to recurrence and survival in stage III and stage IV Melanoma patients, *Cancer*, 49(11), 23362341.

[19] Lee, E. T. and Wang, J. W. (2003). *Statistical Methods for Survival Data Analysis*, Third Edition, John Wiley & Sons, Inc., Hoboken, New Jersey.

[20] Lehmann, E. L. and Casella, G. (1998). *Theory of Point Estimation*, Second Edition, Springer, Berlin.

[21] Mann, N. R. (1971). Best linear invariant estimation for Weibull parameters under progressive censoring, *Technometrics*, 13, 521-533.

[22] Massey Jr, F. J. (1951). The Kolmogorov-Smirnov test for goodness of fit, *Journal of the American Statistical Association*, 46(253), 68-78.

[23] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines, *The Journal of Chemical Physics*, 21(6), 1087-1092.

[24] Mondal, S. and Kundu, D. (2019). Point and interval estimation of Weibull parameters based on joint progressively censored data, *Sankhya*, 81(1)-B, 1-25.

[25] Mondal, S. and Kundu, D. (2020). On the joint type-II progressive censoring scheme, *Communications in Statistics-Theory and Methods*, 49(4), 958-976.

[26] Parsi, S. and Bairamov, I. (2009). Expected values of the number of failures for two populations under joint type-II progressive censoring, *Computational Statistics and Data Analysis*, 53(10), 3560-3570.

[27] Parsi, S., Ganjali, M. and Farsipour, S. (2010). Conditional maximum likelihood and interval estimation for two Weibull populations under joint type-II progressive censoring, *Communication in Statistics- Theory and Methods*, 40(12), 2117-2135.

[28] Rasouli, A. and Balakrishnan, N. (2010). Exact likelihood inference for two exponential populations under joint progressive type-II censoring, *Communications in Statistics-Theory and Methods*, 39(12), 2172-2191.

[29] Smith, A. F. and Roberts, G. O. (1993). Bayesian computation via the Gibbs sampler and related Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society: Series B (Methodological)*, 55(1), 3-23.

[30] Tierney, L. (1994). Markov chains for exploring posterior distributions, *Annals of Statistics*, 1701-1728.

[31] Varian, H. R. (1975). A Bayesian approach to real estate assessment, In: Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage, Stephen E. Fienberg and A. Zellner, Eds., pp. 195208, North-Holland Publishing Company, Amsterdam, The Netherlands.

[32] Viveros, R. and Balakrishnan, N. (1994). Interval estimation of parameters of life from progressively censored data, *Technometrics*, 36, 84-91.