ISSN 1645-6726

REVSTAT

Statistical Journal

vol. 20 - n. 1 - January 2022

Instituto Nacional de Estatística
Statistics Portugal

REVSTAT-Statistical Journal, vol.20, n. 1 (January 2022)

vol.1, 2003- . - Lisbon : Statistics Portugal, 2003- .

Continues: Revista de Estatística = ISSN 0873-4275.

ISSN 1645-6726 ; e-ISSN 2183-0371

Editorial Board (2019-2023)

Editor-in-Chief – *Isabel FRAGA ALVES*

Co-Editor – *Giovani L. SILVA*

Associate Editors

Marília ANTUNES

Barry ARNOLD

Narayanaswamy BALAKRISHNAN

Jan BEIRLANT

Graciela BOENTE

Paula BRITO

Valérie CHAVEZ-DEMOULIN

David CONESA

Charmaine DEAN

Fernanda FIGUEIREDO

Jorge Milhazes FREITAS

Alan GELFAND

Stéphane GIRARD

Marie KRATZ

Victor LEIVA

Artur LEMONTE

Shuangzhe LIU

Maria Nazaré MENDES-LOPES

Fernando MOURA

John NOLAN

Paulo Eduardo OLIVEIRA

Pedro OLIVEIRA

Carlos Daniel PAULINO

Arthur PEWSEY

Gilbert SAPORTA

Alexandra M. SCHMIDT

Manuel SCOTTO

Julio SINGER

Lisete SOUSA

Milan STEHLÍK

María Dolores UGARTE

Executive Editor – *José A. PINTO MARTINS*

Assistant Editors – *José Cordeiro | Olga Bessa Mendes*

Publisher – *Statistics Portugal*

Layout-Graphic Design – *Carlos Perpétuo* | **Cover Design*** – *Helena Nogueira*

Edition - 140 copies | **Legal Deposit Registration** - 191915/03 | **Price** [VAT included] - € 9,00

Creative Commons Attribution 4.0 International (CC BY 4.0)

© Statistics Portugal, Lisbon. Portugal, 2022

**image*: stain glass window by Abel Manta (1888-1982)

Editorial:
In memoriam: Sir David Cox [1924-2022]

We deeply regret to hear that Sir David Cox passed away on January 18th, 2022, at age 97. His name is associated with one of the most influential statisticians of all time. REVSTAT - Statistical Journal is honoured with his collaboration in the Editorial Board since its foundation. David Cox served as REVSTAT Associate Editor from 2002 to 2018. He has been very active until recently, and his last article in our journal was published in October 2020 [Matched pairs with binary outcomes, *REVSTAT - Statistical Journal*, Vol. 18, No. 5, 581-592, 2020, jointly with Christiana Kartsonaki].

Cox Regression Model and its extensions have been widely used in survival analysis since the publication [Regression Models and Life-Tables, *Journal of the Royal Statistical Society - Series B (Methodological),* Vol. 34, No. 2, 187-220, 1972] by David Cox in 1972, one of the most cited papers of all time.

Along his life in the United Kingdom, the long career in Statistics includes the position of Editor of *Biometrika* from 1966 to 1991 and President of the Royal Statistical Society (RSS) in the period 1980-1982. He was knighted by Queen Elizabeth in 1985. The RSS awarded him with Copley Medal in 2010.

Among his long list of interesting talks, we mention 2 in special: In 2014, David Cox gave a lecture in celebrating the 180th anniversary of RSS, 'Statistics past, present and future', and in 2018 the talk 'In gentle praise of significance tests', at the RSS Annual Conference.

David Cox Research Prize is the recently renamed RSS Prize in 2021, which is aimed to award annually a RSS fellow, near the beginning of their research career, for outstanding published contributions to statistical theory or application.

Among other positions and awards, Sir David Cox was Professor of Statistics Head of Department of Mathematics in Imperial College, and Department of Statistics, University of Oxford, and served as Warden and Honorary Fellow of Nuffield College, Oxford.

At last, but not least, it is a pleasure to hear that David Roxbee Cox was able to remain active as a researcher until the end of his life.

January 21, 2022

Isabel Fraga Alves
Giovani L. Silva

INDEX

Abstracted/indexed in: *Current Index to Statistics, Google Scholar, Mathematical Reviews, Science Citation Index Expanded®, Zentralblatt für Mathematic, Scimago Journal & Country Rank* and *SCOPUS.*

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 1–14

SOME RELIABILITY ESTIMATES FOR GENERALIZED EXPONENTIAL DISTRIBUTION WITH PRESENCE OF k-OUTLIERS

Authors: Seyed Mahdi Mahmoudi *
– Department of Statistics, Semnan University,
Semnan, Iran
mahmoudi@semnan.ac.ir

Reza Zarei
– Faculty of Mathematical Sciences, Department of Statistics, University of Guilan,
Rasht, Iran
r.zarei@guilan.ac.ir; rezazarei.r@gmail.com

Received: June 2019 Revised: October 2019 Accepted: October 2019

Abstract:

- This paper studies the problem of parameters estimation for the two-parameter generalized exponential distribution in presence of outliers to estimate the reliability (R) as a measure of system performance. The maximum likelihood and Bayes estimators of R are obtained when the scale parameter λ is known and unknown. Monte Carlo simulations and bootstrap approach are performed to compare the different proposed methods.

Keywords:

- *Bayes estimator; bootstrap method; generalized exponential distribution; maximum likelihood estimator; Monte Carlo simulation; outliers.*

AMS Subject Classification:

- 49A05, 78B26.

*Corresponding author.

1. INTRODUCTION

Over the past decades, different statistical distributions and related models have been proposed for treating randomness and uncertainty, among which the exponentiated Weibull distribution models is a key one [20]. Meanwhile, two-parameter generalized exponential distribution (denoted by GED) has also been proposed as a sub-model in the exponentiated Weibull distribution model which model the real data in a more realistic manner. Several researchers have concentrated on applying this distribution in various fields and studied the problem of parameters estimation for GED [11]–[14], [18], [22], [23], [25], [27]–[30], [35].

Inferences about stress-strength model is an important and interesting fields in the reliability theory. In the mechanical reliability of a system, if we denote X as the strength of a component which is subject to the stress Y, then $R = P(Y < X)$ is known as a measure of system performance. The problem of estimating R for certain family of probability distributions, has been widely studied in the literature. In the following, we review the main studies in this context in an attempt to display the motivation for this paper.

The MLE of $P(Y < X)$, when X and Y have bivariate exponential distribution, has been considered by Awad *et al.* [1]. Church and Harris [2], Downton [6], Woodward and Kelley [34] and Owen *et al.* [26] considered the estimation of $P(Y < X)$, when X and Y are normally distributed. Similar problem for the multivariate normal distribution has been considered by Gupta and Gupta [10]. Kelley *et al.* [16] and Sathe and Shah [32] considered the estimation of $P(Y < X)$ when X and Y are independent exponential random variables. Constantine and Karson [4] considered the estimation of $P(Y < X)$, when X and Y are independent Gamma random variables. Sathe and Dixit [31] have been estimate of $P(Y < X)$ in the negative binomial distribution. Surles and Padgett [33] considered the estimation of $P(Y < X)$, where X and Y are Burr Type random variables. Finally, Nasiri and Pazira [24] have done the estimation of $P(Y < X)$ in exponential case.

The drawback of the above mentioned models is their lack of a supporting the sample data which contain outliers due to human error in measuring or erroneous procedures. To the best of our knowledge, a few researchers investigated the statistical inference about R based on samples contain outlier observation(s). Kim and Chung [17] and Jeevanand and Nair [15] have considered the Bayesian estimation of R based on samples containing outlier from the Burr-X distribution and exponential distribution, respectively. Li and Hao [19] studied the Bayesian and maximum likelihood estimation of R when X and Y are two independent generalized exponential distributions containing one outlier. Pazira and Nasiri [28] and Nasiri [21] consider the estimating parameters of R for generalized exponential distribution and Lomax distribution with presence k-outliers, respectively. Ghanizadeh [8] and Ghanizadeh *et al.* [9] studied the estimation of R in the presence of k-outlier for Rayleigh and Exponentiated Gamma distribution, respectively.

In the present work, the Bayes and maximum likelihood approaches to estimate the $P(Y < X)$ are incorporated into the samples containing outliers. This paper is organized as follows: First, in Section 2, we recall the concept of GED and then formulated the problem.

Then, we investigate the MLE and the Bayes estimators of R when the scale parameter is known and unknown, respectively in Section 3 and 4. The different proposed methods have been compared using Monte Carlo simulations and bootstrap methods and their results have been reported in Section 5. An numerical example is illustrated in Section 6. Finally, a brief conclusion presented in Section 7.

2. MATHEMATICAL FORMULATION

The two-parameter GED has the following density function

$$f(x,\alpha,\lambda) = \alpha\lambda e^{-\lambda x}\left(1-e^{-\lambda x}\right)^{\alpha-1}, \qquad x>0\,, \tag{2.1}$$

where $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters, respectively. We denote the two-parameter GED with the shape parameter α and scale parameter λ will be denoted by GE(α,λ).

For different values of the shape parameter, the density function can take different shape. If the scale parameter λ is equal to one, for $\alpha \le 1$, the density function 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 (α), it is not symmetric. For this density function (2.1), the mode is at $\log\alpha$ for $\alpha > 1$ and for $\alpha \le 1$, the mode is at α. It has the median at $-\ln(1-0.5^{1/\alpha})$. The mean, median and mode are non-linear functions of the shape parameter α and as this parameter goes to infinity all of them tend to infinity. For large values of α, the mean, median and mode are approximately equal to α but they converge at different rates. Figure 1 shows the shape of $f(x,\alpha)$ for different values of α when $\lambda = 1$ (for more details refer to Gupta and Kundu [11]).

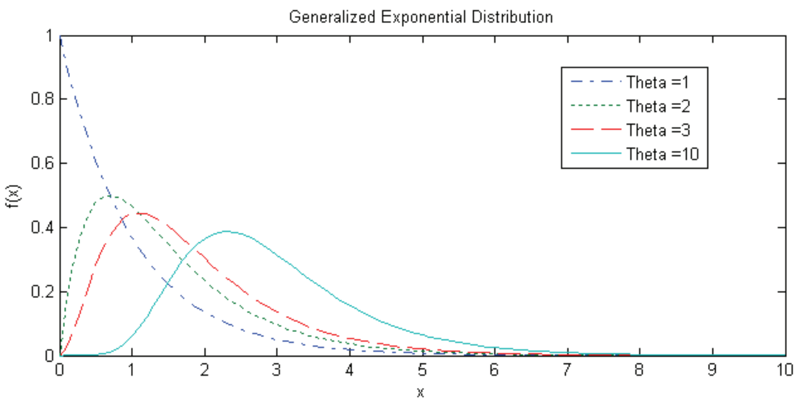

Figure 1: pdf of GE$(\alpha,1)$ for different values of α.

The main aim of this paper is to focus on the inference of $R = P(Y < X)$, where $Y \sim$ GE(α,λ), with pdf denoted in Equation (2.1) and X has GED with presence of k outliers, with pdf

$$f(x,\beta_1,\beta_2,\lambda) = \frac{k}{n}\,f(x,\beta_1,\lambda) + \frac{n-k}{n}\,f(x,\beta_2,\lambda)\,, \qquad x>0\,, \tag{2.2}$$

where function $f(\cdot)$ is given in Equation (2.1). For more details see Dixit [5] and Nasiri and Pazira [23]–[24].

To this end, suppose that $Y_1, Y_2, ..., Y_m$ be a random sample for Y with pdf

$$g(y, \alpha, \lambda) = \alpha\lambda e^{-\lambda y}\left(1-e^{-\lambda y}\right)^{\alpha-1}, \qquad y>0, \tag{2.3}$$

and $X_1, X_2, ..., X_n$ be random sample for X with pdf

$$f(x, \beta_1, \beta_2, \lambda) = \frac{k}{n}\, g(x, \beta_1, \lambda) + \frac{n-k}{n}\, g(x, \beta_2, \lambda), \qquad x>0, \tag{2.4}$$

with presence of k outliers. The function $g(\cdot)$ is given in Equation (2.3).
Then, based on the definition of R, we have that

$$\begin{aligned} R \;=\; P(Y<X) \;&=\; \int_0^\infty\!\!\int_0^x g(y,\alpha,\lambda)\, f(x,\beta_1,\beta_2,\lambda)\, dy\, dx \\ &= \frac{k}{n}\cdot\frac{\beta_1}{\alpha+\beta_1} + \frac{n-k}{n}\cdot\frac{\beta_2}{\alpha+\beta_2}\,. \end{aligned} \tag{2.5}$$

Thus, in order to estimate the R, it is sufficient that we estimate the parameters α, β_1 and β_2.

3. MAXIMUM LIKELIHOOD ESTIMATOR OF R

In this section, we study the maximum likelihood estimation of the R. In order to compute the MLE of R, first we consider the joint distribution of $X_1, X_2, ..., X_n$ with presence of k outliers as follows:

$$\begin{aligned} f(x_1, x_2, ..., x_n) \;&= \\ &= \frac{1}{C(n,k)} \prod_{i=1}^n \left[\beta_2\lambda e^{-\lambda x_i}\left(1-e^{-\lambda x_i}\right)^{\beta_2-1}\right] \sum_{\underline{A}} \prod_{r=1}^k \left(\frac{\beta_1\lambda e^{-\lambda x_{A_r}}\left(1-e^{-\lambda x_{A_r}}\right)^{\beta_1-1}}{\beta_2\lambda e^{-\lambda x_{A_r}}\left(1-e^{-\lambda x_{A_r}}\right)^{\beta_2-1}}\right) \\ &= \frac{1}{C(n,k)}\, \beta_1^k\beta_2^{n-k}\, \lambda^n e^{-\lambda\sum x_i} \prod_{i=1}^n \left[\left(1-e^{-\lambda x_i}\right)^{\beta_2-1}\right] \sum_{\underline{A}} \prod_{r=1}^k \left(1-e^{-\lambda x_{A_r}}\right)^{\beta_1-\beta_2}, \end{aligned} \tag{3.1}$$

where $C(n,k) = \binom{n}{k}$ and $\sum_{\underline{A}} = \sum_{A_1=1}^{n-k+1} \sum_{A_2=A_1+1}^{n-k+2} \cdots \sum_{A_k=A_{k-1}+1}^{n}$. (For more details see [28]).
Using Equation (3.1), the likelihood function based on two observed sample is given as follows:

$$L(\alpha, \beta_1, \beta_2, \lambda) \;=\; g(y_1, y_2, ..., y_m)\, f(x_1, x_2, ..., x_n)\,.$$

The Log-likelihood function of the observed sample is

$$\begin{aligned} \ln L(\alpha,\beta_1,\beta_2,\lambda) \;&=\; m\ln(\alpha\lambda) - \lambda\sum_{i=1}^m y_i + (\alpha-1)\sum_{i=1}^m \ln\left(1-e^{-\lambda y_i}\right) \\ &\quad + \ln\left[\frac{\beta_1^k\beta_2^{n-k}}{C(n,k)}\, \lambda^n e^{-\sum_{i=1}^n \lambda x_i} \prod_{i=1}^n \left[\left(1-e^{-\lambda x_i}\right)^{\beta_2-1}\right] \sum_{\underline{A}} \prod_{r=1}^k \left(1-e^{-\lambda x_{A_r}}\right)^{\beta_1-\beta_2}\right]. \end{aligned} \tag{3.2}$$

It is well known that, in order to compute the The MLE's of α say $\hat{\alpha}$, we must obtain the solution of following equation

$$\frac{\partial \ln L}{\partial \alpha} = \frac{m\lambda}{\alpha\lambda} + \sum_{i=1}^{m} \ln\left(1 - e^{-\lambda y_i}\right) = 0\,,$$

or

$$\frac{m}{\alpha} = -\sum_{i=1}^{m} \ln\left(1 - e^{-\lambda y_i}\right).$$

Hence,

$$\hat{\alpha} = \frac{-m}{\sum_{i=1}^{m} \ln\left(1 - e^{-\hat{\lambda} y_i}\right)}\,. \tag{3.3}$$

In similar way, the MLE's of β_1, β_2 and λ, say $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\lambda}$ respectively, obtained as the solutions of

$$\begin{aligned}\frac{\partial \ln L}{\partial \beta_1} &= \frac{k}{\beta_1} + \frac{\frac{\partial}{\partial \beta_1} \sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0 \\ &= \frac{k}{\beta_1} + \frac{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2} \ln\left(1 - e^{-\lambda x_{A_r}}\right)}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0\,,\end{aligned} \tag{3.4}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \beta_2} &= \frac{n-k}{\beta_2} + \sum_{i=1}^{n} \ln\left(1 - e^{-\lambda x_i}\right) + \frac{\frac{\partial}{\partial \beta_2} \sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0 \\ &= \frac{n-k}{\beta_2} + \sum_{i=1}^{n} \ln\left(1 - e^{-\lambda x_i}\right) - \frac{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2} \ln\left(1 - e^{-\lambda x_{A_r}}\right)}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0\,,\end{aligned} \tag{3.5}$$

$$\begin{aligned}\frac{\partial \ln L}{\partial \lambda} &= \frac{m}{\lambda} - \sum_{i=1}^{m} y_i + \frac{n}{\lambda} - \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{m} \frac{y_i\, e^{-\lambda y_i}}{1 - e^{-\lambda y_i}} \\ &\quad + (\beta_2 - 1) \sum_{i=1}^{n} \frac{x_i\, e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} + \frac{\frac{\partial}{\partial \lambda} \sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0 \\ &= \frac{m}{\lambda} - \sum_{i=1}^{m} y_i + \frac{n}{\lambda} - \sum_{i=1}^{n} x_i + (\alpha - 1) \sum_{i=1}^{m} \frac{y_i\, e^{-\lambda y_i}}{1 - e^{-\lambda y_i}} \\ &\quad + (\beta_2 - 1) \sum_{i=1}^{n} \frac{x_i\, e^{-\lambda x_i}}{1 - e^{-\lambda x_i}} + \frac{\sum_{\underline{A}} \prod_{r=1}^{k} (\beta_1 - \beta_2)\; x_{A_r} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2 - 1}}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0\,.\end{aligned} \tag{3.6}$$

From Equations (3.4)–(3.6), we obtain the $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\lambda}$ as the solution of non-linear equations.

Since ML estimators are invariant, so the MLE of R becomes

$$\hat{R} = \frac{k}{n} \frac{\hat{\beta}_1}{\hat{\alpha} + \hat{\beta}_1} + \frac{n-k}{n} \frac{\hat{\beta}_2}{\hat{\alpha} + \hat{\beta}_2}\,. \tag{3.7}$$

Note 3.1. For $\beta_1 = \beta_2 = \beta$ in case of no outliers presence, $\hat{\alpha}$ and $\hat{\beta}$ can be obtained as

$$\hat{\alpha} = \frac{-m}{\sum_{i=1}^{m} \ln\left(1 - e^{-\hat{\lambda} y_i}\right)}, \qquad \hat{\beta} = \frac{-n}{\sum_{i=1}^{n} \ln\left(1 - e^{-\hat{\lambda} x_i}\right)}$$

and $\hat{\lambda}$ can be obtained as the function of the non-linear equation

$$\begin{aligned} g(\lambda) &= \frac{m+n}{\lambda} - \frac{n}{\sum_{i=1}^{n} \ln\left(1 - e^{-\lambda x_i}\right)} \sum_{i=1}^{n} \frac{x_i\, e^{-\lambda x_i}}{\left(1 - e^{-\lambda x_i}\right)} \\ &\quad - \frac{m}{\sum_{i=1}^{m} \ln\left(1 - e^{-\lambda y_i}\right)} \sum_{i=1}^{m} \frac{y_i\, e^{-\lambda y_i}}{\left(1 - e^{-\lambda y_i}\right)} - \sum_{i=1}^{n} \frac{x_i}{\left(1 - e^{-\lambda x_i}\right)} - \sum_{i=1}^{m} \frac{y_i}{\left(1 - e^{-\lambda y_i}\right)} = 0 \end{aligned}$$

are given by Kundu and Gupta [18].

Note 3.2. The estimation of R when λ is known was studied by Pazira and Nasiri [28]. In this case, the MLE estimation of R is given as Equation (3.7) in which $\hat{\alpha}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ given as follows:

$$\hat{\alpha} = \frac{-m}{\sum_{i=1}^{m} \ln\left(1 - e^{-y_i}\right)}, \tag{3.8}$$

$$\frac{\partial \ln L}{\partial \beta_1} = \frac{k}{\beta_1} + \frac{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-x_{A_r}}\right)^{\beta_1 - \beta_2} \ln\left(1 - e^{-x_{A_r}}\right)}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0, \tag{3.9}$$

$$\frac{\partial \ln L}{\partial \beta_2} = \frac{n-k}{\beta_2} + \sum_{i=1}^{n} \ln\left(1 - e^{-x_i}\right) - \frac{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-x_{A_r}}\right)^{\beta_1 - \beta_2} \ln\left(1 - e^{-x_{A_r}}\right)}{\sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-x_{A_r}}\right)^{\beta_1 - \beta_2}} = 0. \tag{3.10}$$

3.1. Bootstrap method

In this subsection, we propose the percentile bootstrap method based on the idea of Efrom [7] in two cases of parameter λ is known and unknown. The algorithms for estimating the R in these cases are illustrated below.

When λ is unknown

`Step 1:` From the sample $\{y_1, ..., y_m\}$ and $\{x_1, ..., x_n\}$, compute $\hat{\alpha}$, $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\lambda}$ from equations (3.3), (3.4) and (3.5) and (3.6) respectively.

`Step 2:` Using $\hat{\alpha}$ and $\hat{\lambda}$, we generate a bootstrap sample $\{y_1^*, ..., y_m^*\}$ and similarly using $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\lambda}$, generate a bootstrap sample $\{x_1^*, ..., x_n^*\}$. Based on $\{y_1^*, ..., y_m^*\}$ and $\{x_1^*, ..., x_n^*\}$ compute R.

`Step 3:` Repeat `step 2`, NBOOT times.

When λ is known

`Step 1:` From the sample $\{y_1, ..., y_m\}$ and $\{x_1, ..., x_n\}$, compute $\hat{\alpha}$, $\hat{\beta}_1$ and $\hat{\beta}_2$ from Equations (3.8), (3.9) and (3.10) respectively.

`Step 2:` Using $\hat{\alpha}$, we generate a bootstrap sample $\{y_1^*, ..., y_m^*\}$ and similarly using $\hat{\beta}_1$ and $\hat{\beta}_2$, generate a bootstrap sample $\{x_1^*, ..., x_n^*\}$. Based on $\{y_1^*, ..., y_m^*\}$ and $\{x_1^*, ..., x_n^*\}$ compute R.

`Step 3:` Repeat `step 2`, NBOOT times.

4. BAYES ESTIMATOR OF R

In this section, we obtain the Bayes estimation of R under assumption that the parameters β_1, β_2, α and λ are random variables. We mainly obtain the Bayes estimate of R under the squared error loss. It is assumed that the parameters β_1, β_2, α and λ have independent gamma priors with the parameters $\beta_1 \sim \text{Gamma}(a_1, b_1)$, $\beta_2 \sim \text{Gamma}(a_2, b_2)$, $\alpha \sim \text{Gamma}(a_3, b_3)$ and $\lambda \sim \text{Gamma}(a_4, b_4)$. Based on the above assumptions, the joint density of the data, β_1, β_2, α and λ can be obtained as

$$\begin{aligned} L(\text{data}, \beta_1, \beta_2, \alpha, \lambda) &= L(\text{data}; \beta_1, \beta_2, \alpha, \lambda) \,.\, \pi(\beta_1) \,.\, \pi(\beta_2) \,.\, \pi(\alpha) \,.\, \pi(\lambda) \\ &= C_1\, \beta_1^{k+a_1-1}\, \beta_2^{n-k+a_2-1}\, \alpha^{m+a_3-1}\, \lambda^{n+m+a_4-1}\, h(\beta_1, \beta_2, \alpha)\, h(\beta_1, \beta_2, \alpha, \lambda)\,, \end{aligned}$$

where

$$\begin{aligned} C_1 &= \prod_{i=1}^{4} \left(\frac{b_i^{a_i}}{\Gamma(a_i)} \right) \frac{1}{C(n,k)}\,, \\ h\,(\beta_1, \beta_2, \lambda) &= \sum_{\underline{A}} \prod_{r=1}^{k} \left(1 - e^{-\lambda x_{A_r}}\right)^{\beta_1 - \beta_2}, \\ h\,(\beta_1, \beta_2, \alpha, \lambda) &= e^{-b_1\beta_1 - \beta_2\left(b_2 - \sum\limits_{i=1}^{n} \ln\left(1 - e^{-\lambda x_i}\right)\right) - \alpha\left(b_3 - \sum\limits_{j=1}^{m} \ln\left(1 - e^{-\lambda y_j}\right)\right)\, e^{-\lambda(n\bar{x} + m\bar{y})}}\,. \end{aligned}$$

Therefore, the joint posterior density of given the data is

$$L(\beta_1, \beta_2, \alpha, \lambda \,|\, \text{data}) = \tag{4.1}$$

$$= \frac{\beta_1^{k+a_1-1} \beta_2^{n-k+a_2-1}\, \alpha^{m+a_3-1}\, \lambda^{n+m+a_4-1}\, h(\beta_1, \beta_2, \alpha)\;\, h(\beta_1, \beta_2, \alpha, \lambda)}{\int_0^\infty\!\int_0^\infty\!\int_0^\infty\!\int_0^\infty \beta_1^{k+a_1-1} \beta_2^{n-k+a_2-1}\, \alpha^{m+a_3-1}\, \lambda^{n+m+a_4-1}\, h(\beta_1, \beta_2, \alpha)\;\, h(\beta_1, \beta_2, \alpha, \lambda)\; d\beta_1\, d\beta_2\, d\alpha\, d\lambda}\,.$$

Finally, the Bayes estimator of R, denoted by $\hat{R}_{\mathrm{B}}$, is given as follows

$$\hat{R}_{\mathrm{B}} = \tag{4.2}$$

$$= \frac{k}{n} \frac{\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty u(\alpha,\beta_1)\,\beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,\lambda^{n+m+a_4-1}\,h(\beta_1,\beta_2,\alpha)\;h(\beta_1,\beta_2,\alpha,\lambda)\;d\beta_1\,d\beta_2\,d\alpha\,d\lambda}{\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty \beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,\lambda^{n+m+a_4-1}\,h(\beta_1,\beta_2,\alpha)\;h(\beta_1,\beta_2,\alpha,\lambda)\;d\beta_1\,d\beta_2\,d\alpha\,d\lambda}$$
$$+ \frac{n-k}{n} \frac{\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty u(\alpha,\beta_2)\,\beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,\lambda^{n+m+a_4-1}\,h(\beta_1,\beta_2,\alpha)\;h(\beta_1,\beta_2,\alpha,\lambda)\;d\beta_1\,d\beta_2\,d\alpha\,d\lambda}{\int_0^\infty\int_0^\infty\int_0^\infty\int_0^\infty \beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,\lambda^{n+m+a_4-1}\,h(\beta_1,\beta_2,\alpha)\;h(\beta_1,\beta_2,\alpha,\lambda)\;d\beta_1\,d\beta_2\,d\alpha\,d\lambda},$$

where $u(\alpha,\beta_i) = \frac{\beta_i}{\alpha+\beta_i}$, $i = 1, 2$.

Furthermore, in the case of λ known, the Bayes estimator of R is given by

$$\hat{R}_{\mathrm{B}} = \tag{4.3}$$

$$= \frac{k}{n} \frac{\int_0^\infty\int_0^\infty\int_0^\infty u(\alpha,\beta_1)\,\beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,g(\beta_1,\beta_2)\;g(\beta_1,\beta_2,\alpha)\;d\beta_1\,d\beta_2\,d\alpha}{\int_0^\infty\int_0^\infty\int_0^\infty \beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,g(\beta_1,\beta_2)\;g(\beta_1,\beta_2,\alpha)\;d\beta_1\,d\beta_2\,d\alpha}$$
$$+ \frac{n-k}{n} \frac{\int_0^\infty\int_0^\infty\int_0^\infty u(\alpha,\beta_2)\,\beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,g(\beta_1,\beta_2)\;g(\beta_1,\beta_2,\alpha)\;d\beta_1\,d\beta_2\,d\alpha}{\int_0^\infty\int_0^\infty\int_0^\infty \beta_1^{k+a_1-1}\beta_2^{n-k+a_2-1}\,\alpha^{m+a_3-1}\,g(\beta_1,\beta_2)\;g(\beta_1,\beta_2,\alpha)\;d\beta_1\,d\beta_2\,d\alpha},$$

where

$$\begin{aligned} g(\beta_1,\beta_2) &= \sum_{\underline{A}} \prod_{r=1}^{k} \left(1-e^{-\lambda x_{A_r}}\right)^{\beta_1-\beta_2}, \\ g(\beta_1,\beta_2,\alpha) &= e^{-b_1\beta_1 - \beta_2\left(b_2 - \sum_{i=1}^{n} \ln\left(1-e^{-\lambda x_i}\right)\right) - \alpha\left(b_3 - \sum_{j=1}^{m} \ln\left(1-e^{-\lambda y_j}\right)\right)}. \end{aligned}$$

Since Equations (4.2) and (4.3) can not be obtained analytically, we adopt the Gibbs sampling technique to compute the Bayes estimate of R. Moreover, to compute different Bayes estimates, we prefer to use the non-informative prior, because we do not have any prior information on R. On the other hand, the non-informative prior provides prior distributions which are not proper, we adopt the suggestion of Congdon [3] and Kundu and Gupta [18].

5. SIMULATION RESULTS

In this section, we present some results based on Monte Carlo simulations to compare the performance of the different methods. We consider two cases separately to draw inference on R, namely when:

(i) the common scale parameter λ is known;

(ii) the common scale parameter λ is unknown.

In both cases we consider the following small sample size

$$(n, m) = (15, 15), (20, 20), (25, 25), (15, 20), (20, 15), (15, 25), (25, 15), (20, 25), (25, 20).$$

Moreover, in both cases we take $\alpha = 1.50$, $\beta_1 = 2.50$ and $\beta_2 = 2.75$. Without loss of generality we take $\lambda = 1$ in the case λ is known. Here we present a complete analysis of a simulated data, and the results are given in Tables 1 to 4 for $k = 1$ and Tables 5 to 8 for $k = 2$.

It is observed that the maximum likelihood estimator of R, when λ is known and unknown works quite well. We report the average estimates and the MSEs based on 5000 replications. The results are reported in Tables 1 and 2 for $k = 1$, and 5 and 6 for $k = 2$. In this case, as we expected, when $m = n$ and m, n increase then the average biases and the MSEs decrease. For fixed m as n increase the MSEs decrease and also for fixed n as m increases the MSEs decrease.

Based on obtained results, it is clear that the estimator of R using bootstrap method, when λ is known and unknown works quite well. We report the average estimates and the MSEs based on 100 replications. The results are reported in Tables 3 and 4 for $k = 1$, and 7 and 8 for $k = 2$. In this case, as we expected, when $m = n$ and m, n increase then the average biases and the MSEs decrease. For fixed m as n increase the MSEs decrease and also for fixed n as m increases the MSEs decrease.

Table 1: MLE when $k = 1$, $\alpha = 1.5$, $\beta_1 = 2.5$, $\beta_2 = 2.75$ and $\lambda = 1$.

(n, m)	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15, 15)	1.8444	2.5000	3.3135	0.6278	−0.0178	0.0217
(20, 20)	1.6075	2.5000	2.7086	0.6237	−0.0222	0.0063
(25, 25)	1.8233	2.5000	2.7277	0.6074	−0.0388	0.0082
(15, 20)	1.6851	2.5000	3.1127	0.6445	−0.0011	0.0023
(20, 15)	1.4864	2.5000	3.4041	0.6959	0.0500	0.0034
(15, 25)	1.7807	2.5000	2.6832	0.6071	−0.0385	0.0088
(25, 15)	1.4213	2.5000	2.7206	0.6490	0.0028	0.0033
(20, 25)	1.6360	2.5000	2.8331	0.6333	−0.0126	0.0030
(25, 20)	1.5888	2.5000	2.6093	0.6249	−0.0213	0.0073

Table 2: MLE when $k = 1$, $\alpha = 1.5$, $\beta_1 = 2.5$ and $\beta_2 = 2.75$.

(n, m)	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15, 15)	0.9877	1.5666	2.5000	2.7500	0.6443	−0.0012	0.0051
(20, 20)	0.9899	1.6447	2.5000	2.7500	0.6338	−0.0122	0.0050
(25, 25)	1.0223	1.6172	2.5000	2.7500	0.6344	−0.0118	0.0036
(15, 20)	1.0108	1.6242	2.5000	2.7500	0.6365	−0.0091	0.0050
(20, 15)	1.0209	1.6831	2.5000	2.7500	0.6291	−0.0168	0.0060
(15, 25)	1.0165	1.6402	2.5000	2.7500	0.6359	−0.0097	0.0052
(25, 15)	1.0037	1.6527	2.5000	2.7500	0.6324	−0.0138	0.0054
(20, 25)	0.9974	1.5571	2.5000	2.7500	0.6425	−0.0034	0.0032
(25, 20)	1.0251	1.6440	2.5000	2.7500	0.6325	−0.0137	0.0044

Table 3: Bootstrap method when $k = 1$, $\alpha = 1.5$, $\beta_1 = 2.5$, $\beta_2 = 2.75$ and $\lambda = 1$.

(n, m)	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15, 15)	1.6774	14.6622	59.6942	0.9076	0.2620	0.0735
(20, 20)	1.7030	9.7609	180.1635	0.8881	0.2421	0.0780
(25, 25)	0.3793	2.5000	12.5714	0.9010	0.2548	0.0796
(15, 20)	4.5208	2.5000	13.3324	0.6994	0.0539	0.0105
(20, 15)	1.5245	6.5037	89.0338	0.9411	0.2952	0.0899
(15, 25)	3.4078	6.1308	272.3902	0.8519	0.2063	0.0503
(25, 15)	2.6388	2.5401	112.7504	0.8501	0.2039	0.0489
(20, 25)	1.6082	2.5000	7.5632	0.7984	0.1525	0.0324
(25, 20)	0.5908	2.5000	2.1251	0.8065	0.1603	0.0692

Table 4: Bootstrap method when $k = 1$, $\alpha = 1.5$, $\beta_1 = 2.5$ and $\beta_2 = 2.75$.

(n, m)	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15, 15)	0.0942	47.7459	2.5000	2.7500	0.3846	−0.2609	0.1497
(20, 20)	0.1831	4.0096	2.5000	2.7500	0.4931	−0.1528	0.0571
(25, 25)	1.8805	9.7125	2.5000	2.7500	0.3571	−0.2890	0.1195
(15, 20)	0.7036	8.4306	2.5000	2.7500	0.4055	−0.2401	0.0980
(20, 15)	1.9512	6.6062	2.5000	2.7500	0.3598	−0.2861	0.1075
(15, 25)	0.8380	2.2379	2.5000	2.7500	0.6275	−0.0181	0.0474
(25, 15)	0.2228	10.3235	2.5000	2.7500	0.3642	−0.2820	0.1196
(20, 25)	1.4792	2.4191	2.5000	2.7500	0.5761	−0.0698	0.0277
(25, 20)	0.4040	1.7287	2.5000	2.7500	0.6455	−0.0006	0.0198

Table 5: MLE when $k = 2$, $\alpha = 1.5$, $\beta_1 = 2.5$, $\beta_2 = 2.75$ and $\lambda = 1$.

(n, m)	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15, 15)	1.8567	2.8847	2.2264	0.5198	−0.1244	0.0217
(20, 20)	1.5968	2.5165	2.5660	0.5848	−0.0601	0.0071
(25, 25)	1.6174	1.9278	2.4801	0.6004	−0.0449	0.0056
(15, 20)	1.5593	3.9744	2.8183	0.6380	−0.0061	0.0111
(20, 15)	1.6675	3.1053	2.6176	0.6174	−0.0274	0.0166
(15, 25)	1.6831	2.9464	2.3125	0.5741	−0.0700	0.0106
(25, 15)	1.5960	1.7858	2.6101	0.6034	−0.0419	0.0071
(20, 25)	1.6159	3.1946	2.7131	0.6198	−0.0250	0.0060
(25, 20)	1.4687	3.1153	2.8283	0.6556	0.0103	0.0034

Table 6: MLE when $k = 2$, $\alpha = 1.5$, $\beta_1 = 2.5$ and $\beta_2 = 2.75$.

(n, m)	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15, 15)	1.0637	1.6394	2.7108	2.3042	0.5491	−0.0950	0.0106
(20, 20)	0.9842	1.6196	2.6559	2.4108	0.5779	−0.0670	0.0110
(25, 25)	0.9916	1.8235	2.6076	2.4758	0.5824	−0.0629	0.0159
(15, 20)	0.9449	1.3620	2.5189	2.3047	0.5885	−0.0557	0.0037
(20, 15)	1.0560	1.9955	2.8272	2.4110	0.5297	−0.1152	0.0185
(15, 25)	0.9265	1.4211	3.7023	2.3022	0.5857	−0.0584	0.0051
(25, 15)	0.9867	1.5628	2.2423	2.4753	0.5911	−0.0542	0.0056
(20, 25)	1.0057	1.4799	2.4455	2.4105	0.5931	−0.0517	0.0056
(25, 20)	0.9153	1.5412	2.7855	2.4752	0.5947	−0.0506	0.0045

Table 7: Bootstrap method when $k = 2$, $\alpha = 1.5$, $\beta_1 = 2.5$, $\beta_2 = 2.75$ and $\lambda = 1$.

(n, m)	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15,15)	2.21351	0.65190	1.30161	0.348023	−0.254428	0.0671290
(20,20)	2.42581	0.46529	1.21980	0.315847	−0.297756	0.0922301
(25,25)	1.95369	3.91502	2.95105	0.601363	−0.018931	0.0036032
(15,20)	1.14788	5.47549	3.46179	0.748267	0.145816	0.0250333
(20,15)	1.50094	6.10851	3.89440	0.709387	0.095784	0.0205170
(15,25)	1.42394	0.66237	1.25841	0.439300	−0.163151	0.0443621
(25,15)	4.55750	1.89003	2.24141	0.362190	−0.258104	0.0943967
(20,25)	1.09092	1.39501	1.75356	0.603435	−0.010168	0.0032842
(25,20)	1.53441	1.77145	2.36214	0.589008	−0.031286	0.0134582

Table 8: Bootstrap method when MLE when $k = 2$, $\alpha = 1.5$, $\beta_1 = 2.5$ and $\beta_2 = 2.75$.

(n, m)	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{R}$	Bias($\hat{R}$)	MSE($\hat{R}$)
(15,15)	1.5127	2.0224	2.2197	2.3071	0.5339	−0.1102	0.0150
(20,20)	1.0199	1.5270	1.5802	2.4121	0.6068	−0.0380	0.0047
(25,25)	1.1685	2.2125	2.3427	2.4754	0.5302	−0.1151	0.0150
(15,20)	2.2332	1.8661	2.5324	2.3066	0.5534	−0.0908	0.0110
(20,15)	1.3911	0.6630	1.7005	2.4124	0.7806	0.1357	0.0209
(15,25)	1.3795	3.0205	3.4349	2.3031	0.4660	−0.1781	0.0424
(25,15)	0.6082	1.1873	2.5821	2.4751	0.6773	0.0321	0.0020
(20,25)	0.7843	1.9834	2.8723	2.4099	0.5562	−0.0887	0.0103
(25,20)	1.1818	2.1626	2.0703	2.4756	0.5314	−0.1139	0.0142

6. NUMERICAL EXAMPLE

In this section an numerical example is illustrated and the results of different methods are compared. to do this, the data has been generated using $k = 2$, $m = n = 15$, $\alpha = 1.50$, $\beta_1 = 2.50$, $\beta_2 = 2.75$ and $\lambda = 1$. The data has been truncated after four decimal places and it has been presented below. The Y values are

0.1656	1.4907	0.1297	0.1890	1.0442	0.2366	2.0775	2.0741
1.6354	0.3315	1.4178	1.0370	4.0119	1.3847	1.9806	

and the corresponding X values are

3.5641	3.5056	4.9680	2.4494	2.6494	2.7850	3.3939	5.0067
4.8371	2.3331	3.4162	3.7709	3.4634	1.8660	1.7731	

Now, we obtain the MLE estimates of α, β_1, β_2 and R as, $\hat{\alpha} = 2.234$, $\hat{\beta}_1 = 2.5$, $\hat{\beta}_2 = 10.43$, $R = 0.6441$ and therefore $\hat{R} = 0.7542$. Also, using Equation (4.3) the Bayes estimation becomes $\hat{R}_{\mathrm{B}} = 0.7623$.

In case (ii), when λ is unknown, the Y values are

1.5746	0.1059	0.5531	0.1378	0.2374	2.1082	1.5347	0.6255
3.3972	0.1119	0.8613	0.7467	1.8130	1.9542	0.3958	

and the corresponding X values are

3.6642	3.5416	4.1511	4.3893	4.5871	3.0850	4.2729	4.1823
2.7502	2.5972	3.6886	6.5070	3.2589	1.6457	0.7974	

Then $\hat{\alpha} = 0.7731$, $\hat{\beta}_1 = 2.5$, $\hat{\beta}_2 = 2.75$, $\hat{\lambda} = 0.5405$, $R = 0.6441$ and $\hat{R} = 0.7783$. Also, the Bayes estimation becomes $\hat{R}_{\mathrm{B}} = 0.7763$ using Equation (4.2).

For the bootstrap method when λ is known, the Y values are

0.2550	1.3994	0.9810	1.8751	1.6076	2.7293	2.6022	0.6569
1.5485	0.4147	0.1028	1.7211	0.9942	0.9493	2.7400	

and the corresponding X values are

4.0273	4.0531	5.2043	4.8492	3.9213	2.8151	2.9842	5.4328
2.1106	3.6646	2.7675	7.1520	4.4030	1.4194	1.3471	

Then $\hat{\alpha} = 1.7297$, $\hat{\beta}_1 = 2.5$, $\hat{\beta}_2 = 6.206, R = 0.6441$ and $\hat{R} = 0.7566$.

In the bootstrap method when λ is unknown, the Y values are

1.9301	3.3788	0.6447	1.4552	0.8611	2.1686	1.8280	0.3618
2.3616	4.9962	1.0273	2.5419	1.2103	0.3400	0.4183	

and the corresponding X values are

3.4369	4.5594	4.9697	4.7634	3.2003	3.7920	2.4787	2.5690
2.6606	4.2689	3.6796	2.8361	3.6791	0.6259	0.3760	

Then $\hat{\alpha} = 1.7886$, $\hat{\beta}_1 = 2.5$, $\hat{\beta}_2 = 2.75$, $\hat{\lambda} = 0.7535$, $R = 0.6441$ and $\hat{R} = 0.6029$.

7. CONCLUSION

In this paper, we have studied the estimation of $P(Y < X)$ for the GED. We assume that the sample from each population contains k-outlier. Two cases scale parameter is known or unknown are considered in this context. The MLE and Bayes estimator of R are obtained in each case.

When the common scale parameter is unknown, it is observed that the maximum likelihood estimator works quite well. Based on the simulation results, when the sample size is very small, we recommend to use the parametric bootstrap percentile method. The similar results was obtained in the case of the common scale parameter is known.

REFERENCES

[1] Awad, A.M.; Azzam, M.M. and Hamdan, M.A. (1981). Some inference results on $\Pr(X<Y)$ in the bivariate exponential model, *Communications in Statistics – Theory and Methods*, **10**(24), 2515–2525.

[2] Church, J.D. and Harris, B. (1970). The estimation of reliability from stress-strength relationships, *Technometrics*, **12**(1), 49–54.

[3] Congdon, P. (2001). *Bayesian Statistical Modelling*, Wiley, New York.

[4] Constantine, K.; Tse, S.-K. and Karson, M. (1986). Estimation of $P(Y<X)$ in the gamma case, *Communications in Statistics – Simulation and Computation*, **15**(2), 365–388.

[5] Dixit, U.J. (1989). Estimation of parameters of the Gamma distribution in the presence of outliers, *Communications in Statistics – Theory and Methods*, **18**(3), 3071–3085.

[6] Downton, F. (1973). The estimation of $\Pr(Y<X)$ in the normal case. *Technometrics*, **15**(3), 551–558.

[7] Efron, B. (1982). The jackknife, the bootstrap and other resampling plans, *CBMS-NSF Regional Conference Series in Applied Mathematics*, **38**, SIAM, Philadelphia, PA.

[8] Ghanizadeh, A. (2011). Estimation of $P(Y<X)$ in the Rayleigh distribution in the presence of k outliers, *Journal of Applied Sciences*, **11**(17), 3192–3197.

[9] Ghanizadeh, A.; Pazira, H. and Lotfi, R. (2011). Classical estimations of the exponentiated gamma distribution parameters with presence of k outliers, *Australian Journal of Basic and Applied Sciences*, **5**(3), 571–579.

[10] Gupta, R.D. and Gupta, R.C. (1990). Estimation of $\Pr(a'x > b'y)$ in the multivariate normal case, *Statistics*, **21**(1), 91–97.

[11] Gupta, R.D. and Kundu, D. (1999). Generalized exponential distributions, *Australian and New Zealand Journal of Statistics*, **41**(2), 173–188.

[12] Gupta, R.D. and Kundu, D. (2001a). Generalized exponential distribution: different method of estimations, *Journal of Statistical Computation and Simulation*, **69**(4), 315–337.

[13] Gupta, R.D. and Kundu, D. (2001b). Exponentiated exponential family: an alternative to gamma and Weibull distributions, *Biometrical Journal*, **43**(1), 117–130.

[14] Gupta, R.D. and Kundu, D. (2002). Generalized exponential distributions: statistical inferences, *Journal of Statistical Theory and Applications*, **1**(1), 101–118.

[15] Jeevanand, E.S. and Nair, N.U. (1994). Estimating $P[X>Y]$ from exponential samples containing spurious observations, *Communications in Statistics – Theory and Methods*, **23**(9), 2629–2642.

[16] Kelley, G.D.; Kelley, J.A. and Schucany, W.R. (1976). Efficient estimation of $P(Y<X)$ in the exponential case, *Technometrics*, **18**(3), 359–360.

[17] Kim, C. and Chung, Y. (2006). Bayesian estimation of $P(Y<X)$ from Burr-type X model containing spurious observations, *Statistical Papers*, **47**(4), 643–651.

[18] Kundu, D. and Gupta, R.D. (2005). Estimation of $P[Y<X]$ for generalized exponential distribution, *Metrika*, **61**(3), 291–308.

[19] Li, C. and Hao, H. (2016). Likelihood and Bayesian estimation in stress strength model from generalized exponential distribution containing outliers, *IAENG International Journal of Applied Mathematics*, **46**(2), 155–159.

[20] Mudholkar, G.S. and Srivastava, D.K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*, **42**(2), 299–302.

[21] Nasiri, P. (2016). Estimation parameter of $R = P(Y < X)$ for Lomax distribution with presence of outliers, *International Mathematical Forum*, **11**(5), 239–248.

[22] Nasiri, P. and Jabbari Nooghabi, M. (2009). Estimation of $P(Y < X)$ for generalized exponential distribution in presence of outlier, *Iranian Journal of Numerical Analysis and Optimization*, **2**(1), 69–80.

[23] Nasiri, P. and Pazira, H. (2010a). Bayesian approach on the generalized exponential distribution in the presence of outliers, *Journal of Statistical Theory and Practice*, **4**(3), 453–475.

[24] Nasiri, P. and Pazira, H. (2010b). Estimation of $P(Y < X)$ for exponential distribution in the presence of k outliers, *Journal of Applied Sciences*, **10**, 2176–2189.

[25] Nasiri, P.; Shadrokh, A. and Pazira, H. (2012). Estimation of parameters of the exponential geometric distribution with presence of outliers generated from uniform distribution, *Australian Journal of Basic and Applied Sciences*, **6**(1), 198–206.

[26] Owen, D.B.; Craswell, K.J. and Hanson, D.L. (1964). Nonparametric upper confidence bounds for $\Pr\{Y < X\}$ and confidence limits for $\Pr\{Y < X\}$ when X and Y are normal, *Journal of the American Statistical Association*, **59**(307), 906–924.

[27] Pazira, H. and Nasiri, P. (2010a). Estimations on the generalized exponential distribution using grouped data, *Journal of Modern Applied Statistical Methods*, **9**(1), 235–247.

[28] Pazira, H. and Nasiri, P. (2010b). *Estimation of $P(Y < X)$ for generalized exponential distribution in the presence of outliers when scale parameter is known.* In "Proceedings of the 10^{th} Iranian Statistical Conference", University of Tabriz, Iran, 335–344.

[29] Raqab, M.Z. (2002). Inferences for generalized exponential distribution based on record statistics, *Journal of Statistical Planning and Inference*, **104**(2), 339–350.

[30] Raqab, M.Z. and Ahsanullah, M. (2001). Estimation of the location and scale parameters of generalized exponential distribution based on order statistics, *Journal of Statistical Computation and Simulation*, **69**(2), 109–123.

[31] Sathe, Y.S. and Dixit, U.J. (2001). Estimation of $P[X \leq Y]$ in the negative binomial distribution, *Journal of Statistical Planning and Inference*, **93**(1–2), 83–92.

[32] Sathe, Y.S. and Shah, S.P. (1981). On estimating $P(X > Y)$ for the exponential distribution, *Communications in Statistics – Theory and Methods*, **10**(1), 39–47.

[33] Surles, J.G. and Padgett, W.J. (2001). Inference for reliability and stress-strength for a scaled Burr type X distribution, *Lifetime Data Analysis*, **7**(2), 187–200.

[34] Woodward, W.A. and Kelley, G.D. (1977). Minimum variance unbiased estimation of $P[Y < X]$ in the normal case, *Technometrics*, **19**(1), 95–98.

[35] Yarmohammadi, M. and Pazira, H. (2010). Classical and Bayesian estimations on the generalized exponential distribution using censored data, *International Journal of Mathematical Analysis*, **4**(29), 1417–1431.

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 15–32

DYNAMIC RELIABILITY MODELING WITH MEDICAL APPLICATIONS

Authors: M. Mahdizadeh *
– Department of Statistics, Hakim Sabzevari University, P.O. Box 397, Sabzevar, Iran
mahdizadeh.m@hsu.ac.ir; mahdizadeh.m@live.com

M. Bolbolian Ghalibaf
– Department of Statistics, Hakim Sabzevari University, P.O. Box 397, Sabzevar, Iran
m.bolbolian@hsu.ac.ir

Received: October 2018 Revised: October 2019 Accepted: November 2019

Abstract:

- This article explores a dynamic reliability index for dependent stress and strength variables. Toward this end, a copula approach is utilized to model the association between the two variables. Under Farlie–Gumbel–Morgenstern copula and its generalization, expressions of the reliability measure are derived. Numerical results are used to assess effect of the marginal distributions and the reference copula parameters on the reliability index. Application of the proposed method in the context of medicine is also presented.

Keywords:

- *copula; dependence structure; stress-strength model.*

AMS Subject Classification:

- 62N05, 62F86.

*Corresponding author.

1. INTRODUCTION

Let X and Y be two continuous random variables. A large body of literature has grown around statistical inference for $R = P(X > Y)$. This enthusiasm roots in applicability of this quantity in diverse areas. In the so-called stress-strength model in engineering, R measures the reliability of a component, where X and Y represent the strength of the component, and the stress that it is undergoing, respectively. For example, Weerahandi and Johnson [22] considered a rocket-motor experiment in which X represents the chamber burst strength, and Y represents the operating pressure. In medicine, R may be interpreted as a measure of treatment's effectiveness if X and Y are the response variables from treatment and control groups, respectively (Ventura and Racugno [20]). It is also related to receiver operating characteristic (ROC) curve, which is a useful tool in analysis of the discriminatory accuracy of a diagnostic test or marker in distinguishing between diseased and non-diseased individuals. Bamber [2] showed that the area under the ROC curve equals R. Wolfe and Hogg [23] considered R as a general measure for the difference between two populations.

The estimation of R has received considerable attention in the statistical literature. A comprehensive account of this topic appears in Kotz *et al.* [11]. To facilitate mathematical development, most of the pertinent articles assume that X and Y are independent. In many real situations, however, the two variables are correlated. In the following, three examples in the context of engineering, education and economics are presented (see Domma and Giordano [4]):

- Let X and Y be the lifetimes of two electronic devices, stimulated by a single source. Then R is the probability that one survives after the other one.
- Some universities in Japan use an admission test based on Japanese (X) and English (E) knowledge. In order to get admission, a candidate must qualify $X + E > c$, where c is a pre-determined cut-off score. If we set $Y = c - E$, then the admission probability is given by R.
- Let X and Y be household consumption and income, respectively. If consumption exceeds income, then household will face financial stress. Thus, R is a measure of household financial fragility.

The reliability estimation has been studied for some bivariate distributions, including bivariate normal (Gupta and Subramanian [10]), bivariate beta (Nadarajah [15]), bivariate exponential (Nadarajah and Kotz [16]), and bivariate log-normal (Gupta *et al.* [9]), among others. A limitation shared by these articles is that the marginal distributions are of the same type. Moreover, a specific form of dependence between margins is allowed. Bivariate normal distribution is a nice example clarifying these points. Here, the marginal distributions are normal, and their association is linear. To overcome the above shortcomings, Domma and Giordano [4] built on a copula to model the association between the two variables.

Let the random variables X and Y be the lifetimes of two systems. If both systems are operating at time $t > 0$, then their residual lifetimes are given by $X_t = (X - t \,|\, X > t)$ and $Y_t = (Y - t \,|\, Y > t)$. Zardasht and Asadi [24] proposed $R(t) = P(X_t > Y_t)$ as a time-dependent criterion to compare the two residual lifetimes. They studied properties of this measure, and developed a nonparametric estimator for $R(t)$ based on two independent random samples.

Mahdizadeh and Zamanzade [12, 13, 14] are examples of recent works on inference about $R(t)$. In light of the above argument for R, we think that $R(t)$ is also applicable in settings where X_t and Y_t are not independent. For example, in the third example provided above, $R(t)$ can be considered as a measure of household financial fragility, given that the consumption and income exceed a lower bound t. This article employs a copula approach to account for dependence in evaluating $R(t)$. Our approach is similar to that adopted by Domma and Giordano [4].

Section 2 presents some basic properties of copulas. Section 3 provides expressions of $R(t)$ for some parametric family of copulas, and margins. Section 4 contains numerical results evaluating the effect of the marginal distributions and the reference copula parameters on the reliability index. In Section 5, the proposed method is applied to a data set. Final conclusions appear in Section 6. Figures are collected in an Appendix.

2. THE COPULA APPROACH

If $\mathcal{I}$ is the interval [0,1], then a bivariate copula can be represented as $C\colon \mathcal{I}\times\mathcal{I} \to \mathcal{I}$, where C fulfils the following properties:

- For all $u, v \in \mathcal{I}$, $C(u,0) = 0$, $C(0,v) = 0$, $C(u,1) = u$, and $C(1,v) = v$.
- For all $u_1, u_2, v_1, v_2 \in \mathcal{I}$, with $u_1 \le u_2$ and $v_1 \le v_2$,

$$C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ \ge\ 0\,.$$

A famous theorem by Sklar [19] provides the connection between bivariate copulas and bivariate distribution functions. It states that for any two continuous random variables X and Y with joint distribution function H, there exists a unique copula C such that

$$H(x,y) = C\big(F(x), G(y)\big)\,, \qquad \forall\, x, y \in \mathbb{R}\,,$$

where F and G are the marginal distributions of X and Y, respectively. Let f and g be the corresponding marginal density functions. Then, the joint density function is

$$h(x,y) \,=\, c\big(F(x), G(y)\big)\, f(x)\, g(y)\,, \tag{2.1}$$

where $c\big(F(x), G(y)\big) = \dfrac{\partial^2 C\big(F(x), G(y)\big)}{\partial F(x)\, \partial G(y)}$ is called the copula density.

In general, $C \in C_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ is a vector of parameters that determines the degree of dependence between the two random variables. Also, $F \in F_{\boldsymbol{\gamma}}$ and $G \in G_{\boldsymbol{\nu}}$, where $\boldsymbol{\gamma}$ and $\boldsymbol{\nu}$ are vectors of parameters associated with the marginal distributions. For simplicity in notation, all such parameters are assumed implicitly.

A salient feature of copulas is that they allow us to model the dependence structure between random variables independently of the marginal distributions. Owing to this flexibility, the copula approach has drawn much interest in recent years. It has been successfully applied in a variety of scientific fields. Some applications are provided in the following. In biomedical research, Escarela and Carrière [5] employed copula in studying competing risks.

In the actuarial context, Frees and Wang [6] modeled dependent mortality and losses using copulas. In the engineering context, Genest and Favre [7] utilized copulas in hydrological modeling.

3. COMPUTATION OF $R(t)$

We first provide a representation for $R(t)$ which is helpful in our mathematical development. It is easily seen that

$$R(t) = \frac{R_1(t)}{R_2(t)}\,, \tag{3.1}$$

where $R_1(t) = P(X > Y > t)$ and $R_2(t) = P(X > t, Y > t)$. Using $h(x, y)$ in (2.1), components of $R(t)$ can be written as

$$R_1(t) \;=\; \int_t^\infty\!\!\int_t^x c\big(F(x), G(y)\big)\, f(x)\, g(y)\, \mathrm{d}y\, \mathrm{d}x\,, \tag{3.2}$$

and

$$R_2(t) \;=\; \int_t^\infty\!\!\int_t^\infty c\big(F(x), G(y)\big)\, f(x)\, g(y)\, \mathrm{d}y\, \mathrm{d}x\,. \tag{3.3}$$

In the following, the marginal distributions and the copulas used in computing (3.1) are introduced.

Burr [3] introduced a family of distributions that includes twelve distribution types. Two important cases are the Burr type III (BIII), and Burr type XII. The former distribution is more flexible in the sense that it covers wider ranges of skewness and kurtosis, often exhibited by real data. It has been applied in a multitude of data-modeling contexts. The interested reader is referred to Zimmer *et al.* [25] and Shao [18] for some applications in reliability and environmental studies, among others.

The cumulative distribution function (CDF) and the probability density function (PDF) of the BIII distribution are given by

$$F(x) = \big(1 + x^{-\delta}\big)^{-\alpha}, \qquad x > 0;\;\; \alpha, \delta > 0\,,$$

and

$$f(x) = \alpha\,\delta\, x^{-(\delta+1)}\big(1 + x^{-\delta}\big)^{-(\alpha+1)}, \qquad x > 0;\;\; \alpha, \delta > 0\,,$$

respectively. The random variable X with this distribution will be denoted by $X \sim \mathrm{BIII}(\alpha, \delta)$. In our reliability modeling, it is assumed that both stress and strength variables follow the BIII distribution. The positivity assumption for X and Y is not restrictive, because it is possible to use an increasing transformation to create positive random variables from arbitrary X and Y, while preserving the dependence structure. See Theorem 2.4.3 in Nelsen [17].

To model the association between the two variables, we consider two famous copulas: Farlie–Gumbel–Morgenstern (FGM), and generalized Farlie–Gumbel–Morgenstern (GFGM). These copulas enjoy the advantage of mathematical tractability. In particular, it turns out that under both families, $R_1(t)$ and $R_2(t)$ are decomposed into two components. The first one represents the numerator/denominator in (3.1) when X and Y are independent, and the second one indicates contribution of the association between the two variables in the value of the corresponding quantity. This property is not shared by all other copulas.

3.1. Using FGM copula

The FGM copula is one of the most popular parametric family of copulas that has been widely used due to its simple form. It is defined as

$$C\big(F(x), G(y)\big) \,=\, F(x)\, G(y) \left(1 + \theta\big[1 - F(x)\big]\,\big[1 - G(y)\big]\right), \qquad \theta \in [-1, 1]\,.$$

The corresponding copula density is given by

$$c\big(F(x), G(y)\big) \,=\, 1 + \theta\big[1 - 2F(x)\big]\,\big[1 - 2G(y)\big]\,, \qquad \theta \in [-1, 1]\,. \tag{3.4}$$

Substituting (3.4) in (3.2) and with some algebra, it follows that

$$R_1(t) \,=\, R_1^I(t) + \theta\, R_1^D(t)\,,$$

where

$$R_1^I(t) \,=\, \int_t^\infty G(x)\,\mathrm{d}F(x) \,-\, G(t)\big[1 - F(t)\big]\,, \tag{3.5}$$

and

$$R_1^D(t) \,=\, \int_t^\infty \big[1 - 2F(x)\big]\,\big[G(x) - G^2(x)\big]\,\mathrm{d}F(x) \,+\, \big[F(t) - F^2(t)\big]\,\big[G(t) - G^2(t)\big]\,. \tag{3.6}$$

Again, substituting (3.4) in (3.3) and some simplification yield

$$R_2(t) \,=\, R_2^I(t) + \theta\, R_2^D(t)\,,$$

where

$$R_2^I(t) \,=\, \big[1 - F(t)\big]\,\big[1 - G(t)\big]\,, \tag{3.7}$$

and

$$R_2^D(t) \,=\, F(t)\, G(t)\big[1 - F(t)\big]\,\big[1 - G(t)\big]\,. \tag{3.8}$$

If $X \sim \mathrm{BIII}(\alpha, \delta)$ and $Y \sim \mathrm{BIII}(\beta, \delta)$, then it is possible to obtain a closed-form expression for $R(t)$. For notational convenience, $S(t; \delta, k)$ is defined as

$$S(t; \delta, k) = \frac{1}{k}\left[1 - \big(1 + t^{-\delta}\big)^{-k}\right], \qquad t, k, \delta > 0\,. \tag{3.9}$$

Incorporating the PDF and CDF of the the BIII distribution in (3.5) and (3.6), we get

$$R_1^I(t) \,=\, \alpha\, S(t; \delta, \alpha + \beta) - \alpha\, S(t; \delta, \alpha)\left[1 - \beta\, S(t; \delta, \beta)\right],$$

and

$$\begin{aligned}R_1^D(t) \;=\; & \alpha\Big[S(t;\delta,\alpha{+}\beta) - S(t;\delta,\alpha{+}2\beta)\Big] \\ & - 2\alpha\Big[S(t;\delta,2\alpha{+}\beta) - S(t;\delta,2(\alpha{+}\beta))\Big] \\ & + \alpha\,\beta\,S(t;\delta,\alpha)\,S(t;\delta,\beta)\,\Big[1-\alpha\,S(t;\delta,\alpha)\Big]\Big[1-\beta\,S(t;\delta,\beta)\Big].\end{aligned}$$

Similarly, one can verify that (3.7) and (3.8) take the forms

$$R_2^I(t) \;=\; \alpha\,\beta\,S(t;\delta,\alpha)\,S(t;\delta,\beta)\,,$$

and

$$R_2^D(t) \;=\; \alpha\,\beta\,S(t;\delta,\alpha)\,S(t;\delta,\beta)\,\Big[1-\alpha\,S(t;\delta,\alpha)\Big]\Big[1-\beta\,S(t;\delta,\beta)\Big].$$

3.2. Using GFGM copula

Any copula depends on some parameters which determine the degree of dependence between the margins. Two common measures of the association are Spearman's ρ and Kendall's τ coefficients. Under the FGM copula, $\rho\in[-1/3,1/3]$ and $\tau\in[-2/9,2/9]$, meaning that a relatively weak dependence is allowed. As a result, several modifications of the original FGM copula have been proposed. In the following, we consider a generalization due to Bairamov *et al.* [1]. The GFGM copula is defined as

$$C\big(F(x),G(y)\big) \;=\; F(x)\,G(y)\,\Big(1+\theta\left[1-F(x)^{m_1}\right]^{p_1}\left[1-G(y)^{m_2}\right]^{p_2}\Big)\,,$$

where m_1, m_2, p_1, and p_2 are positive parameters, and $\theta\in[\theta_\ell,\theta_u]$ with

$$\theta_\ell \;=\; -\min\left\{1,\;\frac{1}{m_1 m_2}\left(\frac{1+m_1 p_1}{m_1(p_1-1)}\right)^{p_1-1}\left(\frac{1+m_2 p_2}{m_2(p_2-1)}\right)^{p_2-1}\right\},$$

and

$$\theta_u \;=\; \min\left\{\frac{1}{m_1}\left(\frac{1+m_1 p_1}{m_1(p_1-1)}\right)^{p_1-1},\;\frac{1}{m_2}\left(\frac{1+m_2 p_2}{m_2(p_2-1)}\right)^{p_2-1}\right\}.$$

The corresponding copula density is given by

$$\begin{aligned}c\big(F(x),G(y)\big) \;=\; & 1+\theta\Big[1-F(x)^{m_1}\Big]^{p_1-1}\Big[1-(1+m_1p_1)\,F(x)^{m_1}\Big] \\ & \times\Big[1-G(y)^{m_2}\Big]^{p_2-1}\Big[1-(1+m_2p_2)\,G(y)^{m_2}\Big].\end{aligned} \tag{3.10}$$

Clearly, by setting $m_1=m_2=p_1=p_2=1$ in the above equation, we arrive at (3.4).

Let p_1 and p_2 be two positive integers. Then using the binomial expansion in (3.10), it can be shown that

$$\begin{aligned}c\big(F(x),G(y)\big) \;=\; & 1+\theta\sum_{i=0}^{p_1-1}\sum_{j=0}^{p_2-1}\binom{p_1-1}{i}\binom{p_2-1}{j}(-1)^{i+j}\,F(x)^{m_1 i}\,G(y)^{m_2 j} \\ & \times\Big[1-(1+m_1p_1)\,F(x)^{m_1}\Big]\Big[1-(1+m_2p_2)\,G(y)^{m_2}\Big].\end{aligned} \tag{3.11}$$

This representation will be used in computing $R(t)$. Proceeding as in the previous sub-section, we get

$$R_1(t) = R_1^I(t) + \theta R_1^D(t)\,,$$

where $R_1^I(t)$ is given in (3.5), and

$$R_1^D(t) = \sum_{i=0}^{p_1-1}\binom{p_1-1}{i}(-1)^i\int_t^\infty F(x)^{m_1 i}\Big[1-(1+m_1p_1)\,F(x)^{m_1}\Big]\,J(x)\,\mathrm{d}F(x)\,, \tag{3.12}$$

with

$$\begin{aligned} J(x) = \sum_{j=0}^{p_2-1}\binom{p_2-1}{j}(-1)^j\Bigg(&\frac{1}{m_2j+1}\Big[G(x)^{m_2j+1}-G(t)^{m_2j+1}\Big]\\ &-\frac{1+m_2p_2}{m_2(j+1)+1}\Big[G(x)^{m_2(j+1)+1}-G(t)^{m_2(j+1)+1}\Big]\Bigg). \end{aligned}$$

Similarly, it is concluded that

$$R_2(t) = R_2^I(t) + \theta R_2^D(t)\,,$$

where $R_2^I(t)$ is given in (3.7), and

$$\begin{aligned} R_2^D(t) = &\sum_{i=0}^{p_1-1}\sum_{j=0}^{p_2-1}\binom{p_1-1}{i}\binom{p_2-1}{j}(-1)^{i+j}\\ &\times\left(\frac{1}{m_1i+1}\Big[1-F(t)^{m_1i+1}\Big]-\frac{1+m_1p_1}{m_1(i+1)+1}\Big[1-F(t)^{m_1(i+1)+1}\Big]\right)\\ &\times\left(\frac{1}{m_2j+1}\Big[1-G(t)^{m_2j+1}\Big]-\frac{1+m_2p_2}{m_2(j+1)+1}\Big[1-G(t)^{m_2(j+1)+1}\Big]\right). \end{aligned} \tag{3.13}$$

If $X\sim\text{BIII}(\alpha,\delta)$, $Y\sim\text{BIII}(\beta,\delta)$, and $S(t;\delta,k)$ is defined as in (3.9), then from (3.12) and (3.13) we have

$$\begin{aligned} R_1^D(t) = &\sum_{i=0}^{p_1-1}\sum_{j=0}^{p_2-1}\binom{p_1-1}{i}\binom{p_2-1}{j}(-1)^{i+j}\\ &\times\Bigg\{\frac{\alpha}{(m_2j+1)}\,S\Big(t;\,\delta,\,\alpha(m_1i+1)+\beta(m_2j+1)\Big)\\ &-\frac{\alpha\,(1+m_2p_2)}{m_2(j+1)+1}\,S\Big(t;\,\delta,\,\alpha(m_1i+1)+\beta\big(m_2(j+1)+1\big)\Big)\\ &-\frac{\alpha\,(1+m_1p_1)}{m_2j+1}\,S\Big(t;\,\delta,\,\alpha\big(m_1(i+1)+1\big)+\beta(m_2j+1)\Big)\\ &+\frac{\alpha\,(1+m_1p_1)\,(1+m_2p_2)}{m_2(j+1)+1}\,S\Big(t;\,\delta,\,\alpha\big(m_1(i+1)+1\big)+\beta\big(m_2(j+1)+1\big)\Big)\\ &+\Bigg((1+m_2p_2)\left[\frac{1}{m_2(j+1)+1}-\beta\,S\Big(t;\,\delta,\,\beta\big(m_2(j+1)+1\big)\Big)\right]\\ &-\left[\frac{1}{m_2j+1}-\beta\,S\Big(t;\,\delta,\,\beta(m_2j+1)\Big)\right]\Bigg)\\ &\times\alpha\left[S\Big(t;\,\delta,\,\alpha(m_1i+1)\Big)-(1+m_1p_1)\,S\Big(t;\,\delta,\,\alpha\big(m_1(i+1)+1\big)\Big)\right]\Bigg\}, \end{aligned}$$

and

$$
\begin{aligned}
R_2^D(t) = & \sum_{i=0}^{p_1-1} \sum_{j=0}^{p_2-1} \binom{p_1-1}{i} \binom{p_2-1}{j} (-1)^{i+j} \\
& \times \alpha \left[S\Big(t;\delta,\alpha(m_1 i+1)\Big) - (1+m_1 p_1)\, S\Big(t;\delta,\alpha\big(m_1(i+1)+1\big)\Big) \right] \\
& \times \beta \left[S\Big(t;\delta,\beta(m_2 j+1)\Big) - (1+m_2 p_2)\, S\Big(t;\delta,\beta\big(m_2(j+1)+1\big)\Big) \right].
\end{aligned}
$$

4. NUMERICAL RESULTS

We now evaluate $R(t)$ for some specific choices of the marginal distributions, and the reference copula parameters. Figures 1–5 show the curves of $R(t)$, where the involved parameters are given in the caption of each figure. The following configurations of

$$(m_1, m_2, p_1, p_2) \in \Big\{(1,1,1,1),\ (1,4,2,7),\ (1,4,1,10),\ (4,1,3,2),\ (5,5,2,1)\Big\}$$

are associated with Figures 1–5, respectively. In each case, sixteen combinations of (α, β, δ) and θ are considered whose values can be found in the caption of the figures. In particular, black/solid curves indicate the situation that X and Y are independent, i.e. $\theta = 0$ in (3.4) and (3.10).

Figure 1 is given to the FGM copula. For fixed t, $R(t)$ is a decreasing function of θ if $\alpha < \beta$. The situation is reversed if $\alpha > \beta$. For example, compare panels (a) and (c). These properties are easily concluded in the special case of $t = 0$, as mentioned by Domma and Giordano [4]. The plot presented in panel (d) is very interesting. In fact, it can be shown that if the marginal distributions are identical ($\alpha = \beta$), then $R(t) = 0.5$ for all t, regardless of θ. Finally, one can see that $R(t)$ is a monotone function of t, given a fixed θ.

Figures 2–5 correspond to the GFGM copula. Depending on values of the involved parameters, the reliability measure takes a variety of functional forms. A marked difference from Figure 1 is that for fixed θ, $R(t)$ may not be a monotone function of t. This is observed in panel (d) of Figure 3, for example. If the margins are the same and $\theta = 0$ (see Figures 2 and 4), then it can be proved that $R(t) = 0.5$ for all t. We note that under the FGM copula, this property holds for arbitrary θ.

It should be emphasized that if the dependence between X and Y is not incorporated in computing the reliability, the resulting value could be higher/lower than the true one. Compare black/solid curve with the others in each panel of Figures 1–5. This highlights importance of the copula approach as it is an efficient way to capture dependence structure between random variables in developing inference procedures.

5. APPLICATION

In this section, application of the copula-based approach in reliability modeling is provided based on China Health and Nutrition Survey (CHNS) data. The CHNS is an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill, and the National Institute for Nutrition and Health at the Chinese Center for Disease Control and Prevention. It is designed to examine the effects of the health, nutrition, and family planning policies and programs implemented by national and local governments and to see how the social and economic transformation of Chinese society is affecting the health and nutritional status of its population.

Recent studies support the importance of the lipid-transporting apolipoproteins, such as ApoA and ApoB which transport high-density lipoprotein (HDL, good) cholesterol and low-density lipoprotein (LDL, bad) cholesterol particles, respectively. A healthy individual probably has larger ApoA value than ApoB, and thereby less risk for cardiovascular disease. As alternatives to the traditional LDL and HDL biomarkers, these apolipoproteins have some advantages (Walldius and Jungner [21]). Let R be the probability of ApoA being greater than ApoB, where both were from the same individual, i.e., $R = P(\text{ApoA} > \text{ApoB})$. If this probability is significantly larger than 0.5, then ApoA is stochastically larger than ApoB for the population, meaning that this population is relatively at lower risk of cardiovascular disease. Suppose from the previous studies, the researcher knows a lower bound t for values of the biomarkers in the population. Then, one can utilize the index

$$R(t) = \frac{P\big(\text{ApoA} > \text{ApoB} > t\big)}{P\big(\text{ApoA} > t,\, \text{ApoB} > t\big)}\,.$$

The CHNS data set[1] contains values of ApoA and ApoB biomarkers for 10,187 Chinese children and adults (aged ≥ 7) in year 2009. For the purpose of illustration, we estimated $R(t)$ based on the first 1,000 pairs of data. In doing so, we used the GFGM copula and assumed that the margins are $X \sim \text{BIII}(\alpha, \delta)$ and $Y \sim \text{BIII}(\beta, \delta)$, where X and Y denote ApoA and ApoB, respectively. In particular, the copula parameters were chosen as $m_1 = m_2 = 3$ and $p_1 = p_2 = 2$. This set of values allows for nearly the maximum degree of dependence between the margins under the GFGM copula. Moreover, it simplifies the model through setting $m_1 = m_2$ and $p_1 = p_2$. The last parameter of the copula can be estimated from the expression of Kendall's τ. Domma and Giordano [4] showed that Kendall's τ for the GFGM copula is given by

$$\tau = \frac{8\,\theta\, p_1 p_2}{(2 + m_1 p_1)\,(2 + m_2 p_2)}\; B\left(\frac{2}{m_1}, p_1\right) B\left(\frac{2}{m_2}, p_2\right),$$

where $B(\cdot,\cdot)$ is the beta function. By replacing τ in the above equation with its value from the sample, and setting $m_1 = m_2 = 3$ and $p_1 = p_2 = 2$, an estimate of θ is obtained as 0.176. It is to be noted that 0.176 falls into admissible range of θ in the GFGM copula with the aforesaid choices of m_i's and p_i's, i.e. $[-0.605, 0.778]$.

Before using the results of Sub-section 3.2, it is needed to formally assess fit of the above-mentioned copula to the data. Toward this end, we employed the test statistic $S_n^{(B)}$, introduced by Genest *et al.* [8], based on Rosenblatt's transform. The P-value associated

[1] It is available at http://www.cpc.unc.edu/projects/china/data/datasets.

with this test is determined through parametric bootstrap, where the details can be found in Appendix D of Genest *et al.* [8]. In doing so, parameters α, β and δ were estimated from data by 1.286, 0.522 and 7.398, respectively. Based on 1,000 bootstrap replications, an approximate P-value for the test was computed as 0.316. So the null hypothesis that the selected copula fits the data is not rejected at 0.05 significance level. Figure 6 shows the PDF constructed using this specific GFGM copula with the Burr III marginal distributions.

Plugging in the above set of parameters into the expression of $R(t)$ in Sub-section 3.2 yields an estimate of the dynamic reliability. The corresponding graph is depicted in Figure 7 with blue/dashed curve. A similar graph may be plotted by ignoring the dependence between the two variables, i.e. replacing 0.176 with 0 in the computations. The result is displayed by black/solid curve in Figure 7. It is worth commenting that failing to incorporate the dependence structure leads to inaccuracy in estimating $R(t)$.

6. CONCLUSION

In the classical stress-strength model, the interest centers on $R = P(X > Y)$ for a unit, where X and Y are the strength of the unit and the environmental stress, respectively. This model has attracted much interest in the statistical literature. There are abundant applications in the areas of reliability, quality control, psychology, medicine and clinical trials. Recently, R has been extended to a dynamic form $R(t) = P(X_t > Y_t)$, where X_t and Y_t are residual lifetimes of two systems. Although the latter measure was motivated by a problem in reliability theory, it is potentially applicable in many other situations. This article puts forward a copula approach to account for dependence in evaluating $R(t)$. Some explicit expressions for $R(t)$ are provided when the margins follow the BIII distribution, and the reference copula is either the FGM or GFGM. The proposed method is explored by means of numerical results and real data analysis.

It would be interesting to use other copulas, which allow for higher correlation between the stress and strength variables, in dynamic reliability modeling. This will be considered in a separate study.

APPENDIX

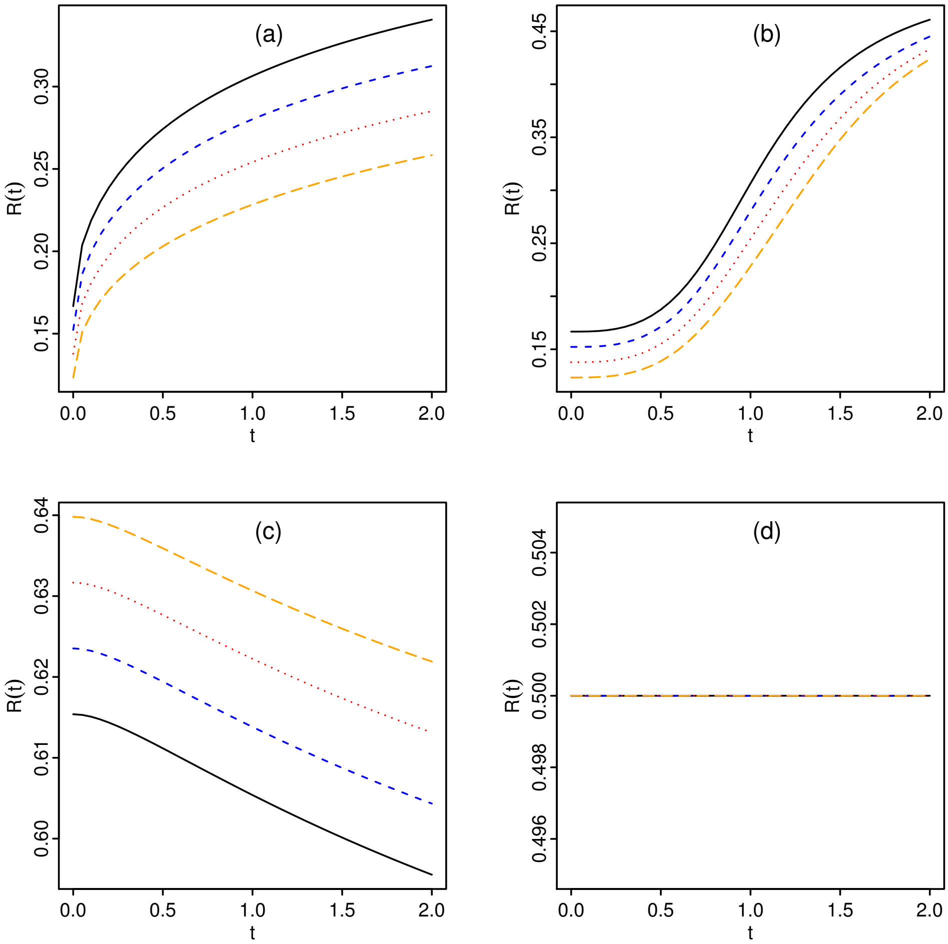

Figure 1: Plot of $R(t)$ based on the FGM copula and the Burr III marginal distributions with: (a) $(\alpha, \beta, \delta) = (1, 5, 0.5)$, (b) $(\alpha, \beta, \delta) = (1, 5, 3)$, (c) $(\alpha, \beta, \delta) = (8, 5, 0.5)$, and (d) $(\alpha, \beta, \delta) = (8, 8, 3)$. Black/solid, blue/dashed, red/dotted, and orange/long-dashed curves relate to $\theta = 0, 0.333, 0.667, 1$, respectively.

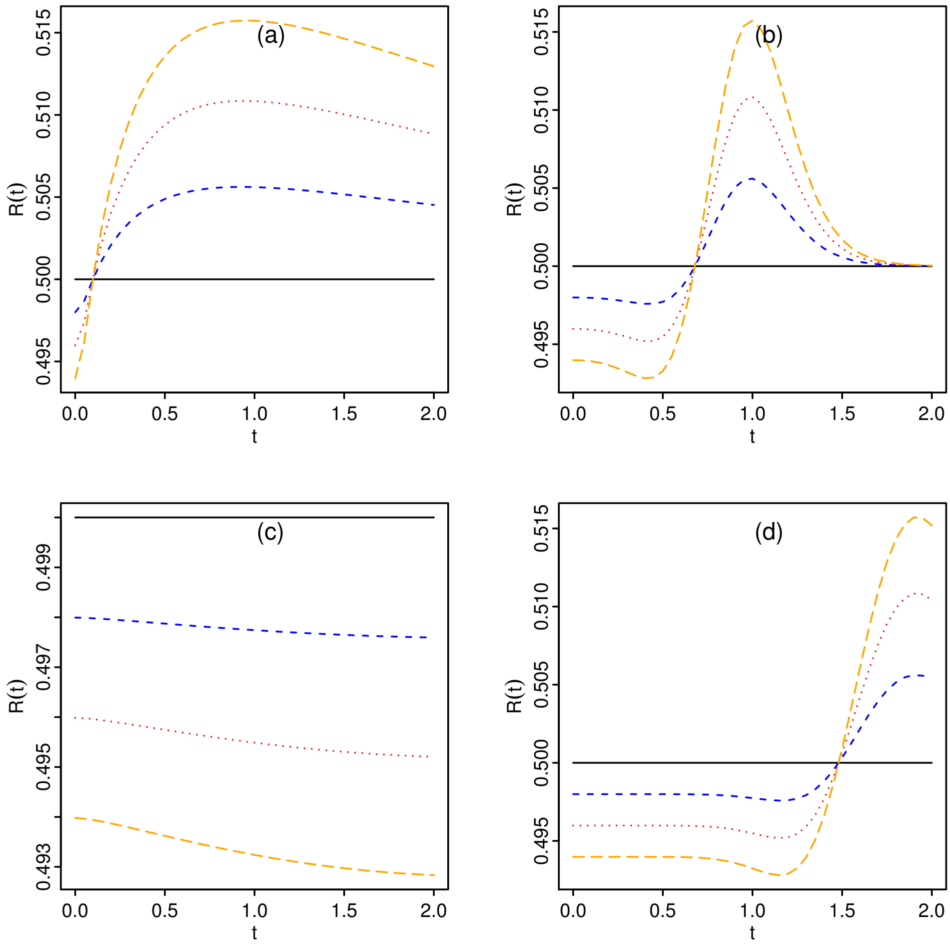

Figure 2: Plot of $R(t)$ based on the GFGM copula with $(m_1, m_2, p_1, p_2) = (1, 4, 2, 7)$, and the Burr III marginal distributions with: (a) $(\alpha, \beta, \delta) = (0.75, 0.75, 0.5)$, (b) $(\alpha, \beta, \delta) = (0.75, 0.75, 3)$, (c) $(\alpha, \beta, \delta) = (4, 4, 0.5)$, and (d) $(\alpha, \beta, \delta) = (4, 4, 3)$. Black/solid, blue/dashed, red/dotted, and orange/long-dashed curves relate to $\theta = 0, 0.259, 0.519, 0.778$, respectively.

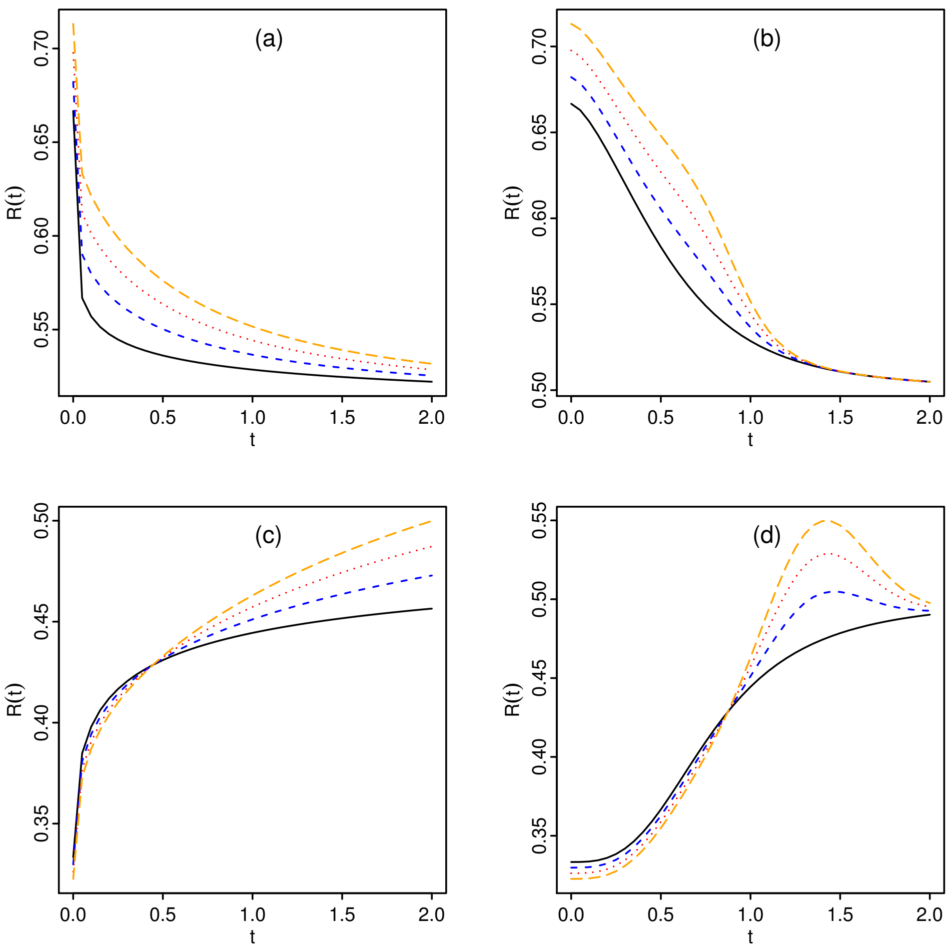

Figure 3: Plot of $R(t)$ based on the GFGM copula with $(m_1, m_2, p_1, p_2) = (1, 4, 1, 10)$, and the Burr III marginal distributions with: (a) $(\alpha, \beta, \delta) = (1, 0.5, 0.5)$, (b) $(\alpha, \beta, \delta) = (1, 0.5, 3)$, (c) $(\alpha, \beta, \delta) = (1, 2, 0.5)$, and (d) $(\alpha, \beta, \delta) = (1, 2, 3)$. Black/solid, blue/dashed, red/dotted, and orange/long-dashed curves relate to $\theta = 0, 0.269, 0.537, 0.806$, respectively.

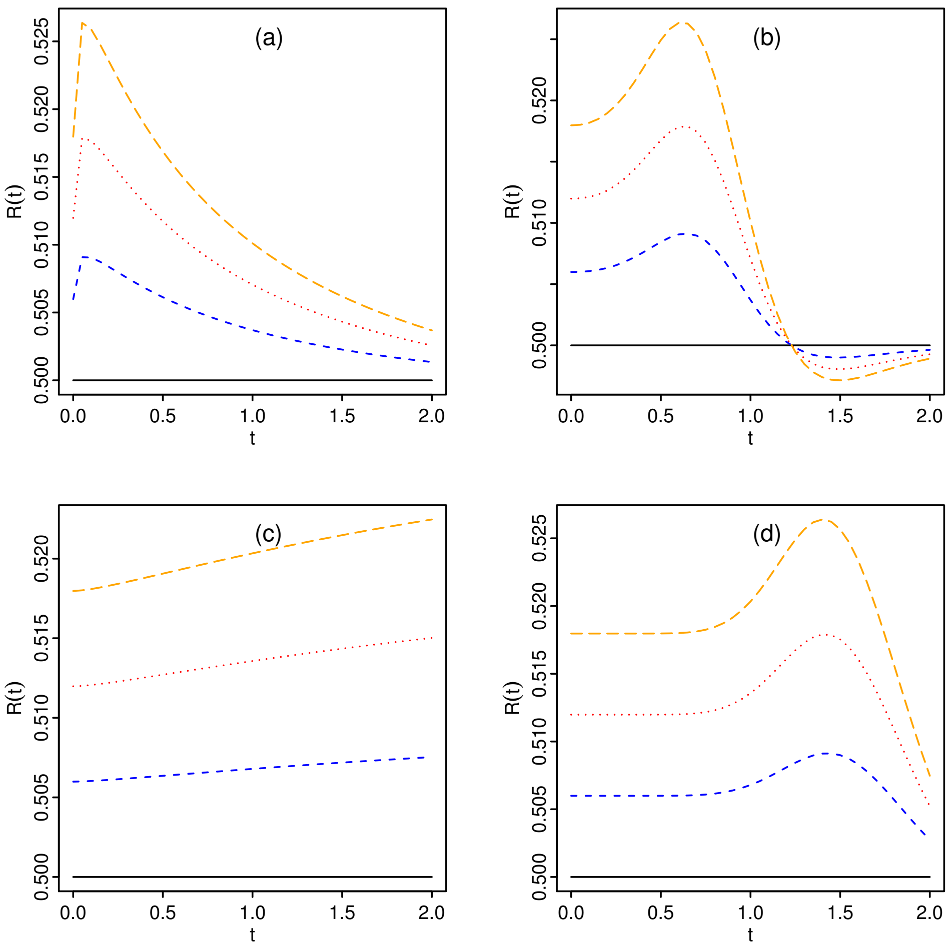

Figure 4: Plot of $R(t)$ based on the GFGM copula with $(m_1, m_2, p_1, p_2) = (4, 1, 3, 2)$, and the Burr III marginal distributions with: (a) $(\alpha, \beta, \delta) = (0.75, 0.75, 0.5)$, (b) $(\alpha, \beta, \delta) = (0.75, 0.75, 3)$, (c) $(\alpha, \beta, \delta) = (4, 4, 0.5)$, and (d) $(\alpha, \beta, \delta) = (4, 4, 3)$. Black/solid, blue/dashed, red/dotted, and orange/long-dashed curves relate to $\theta = 0, 0.22, 0.44, 0.66$, respectively.

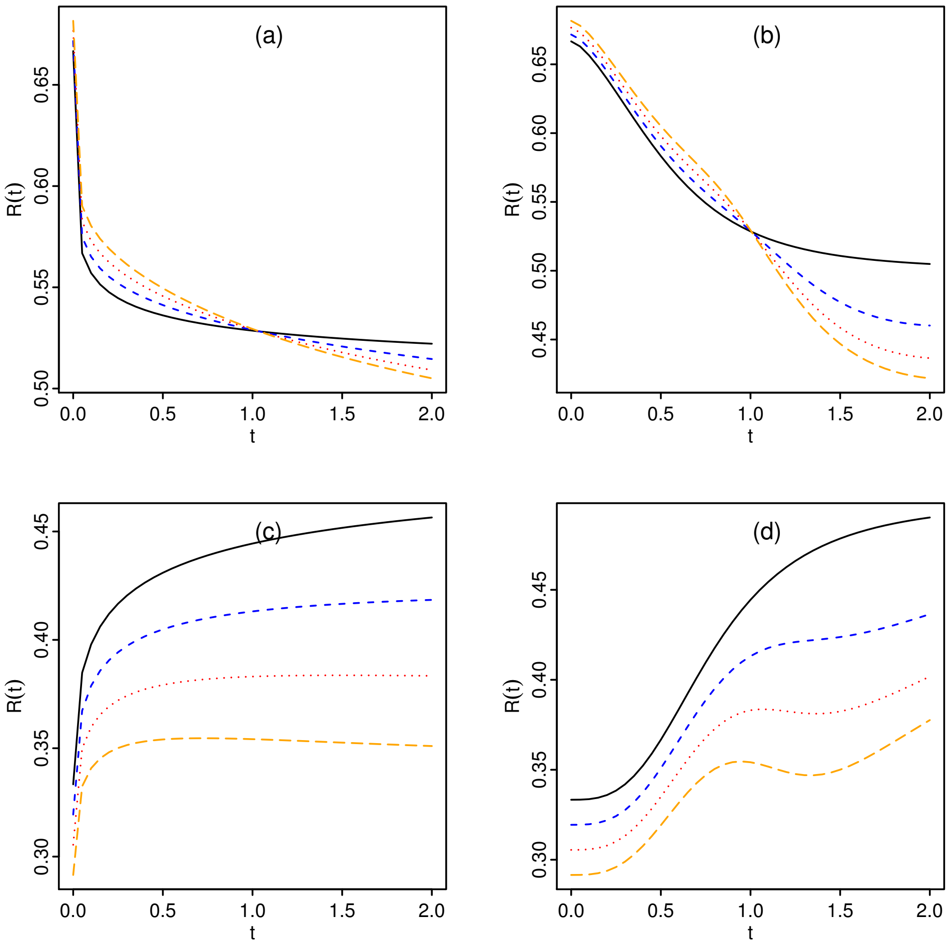

Figure 5: Plot of $R(t)$ based on the GFGM copula with $(m_1, m_2, p_1, p_2) = (5, 5, 2, 1)$, and the Burr III marginal distributions with: (a) $(\alpha, \beta, \delta) = (1, 0.5, 0.5)$, (b) $(\alpha, \beta, \delta) = (1, 0.5, 3)$, (c) $(\alpha, \beta, \delta) = (1, 2, 0.5)$, and (d) $(\alpha, \beta, \delta) = (1, 2, 3)$. Black/solid, blue/dashed, red/dotted, and orange/long-dashed curves relate to $\theta = 0, 0.067, 0.133, 0.200$, respectively.

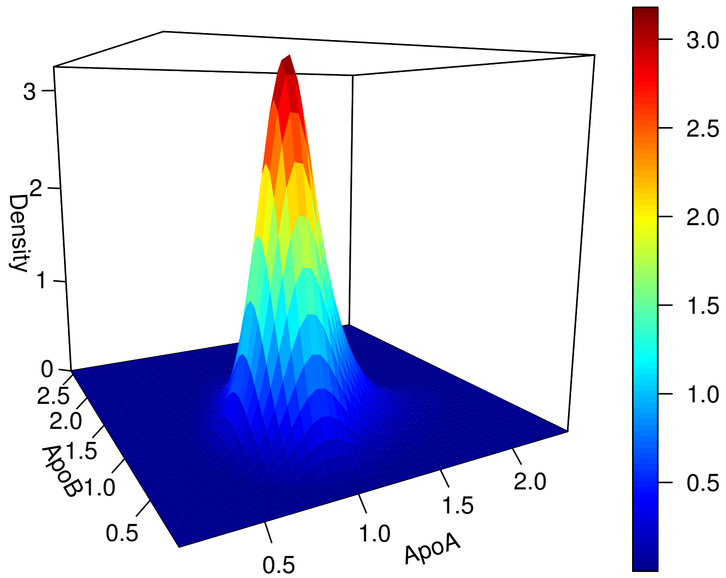

Figure 6: Plot of the PDF constructed using the GFGM copula with $(m_1, m_2, p_1, p_2) = (3, 3, 2, 2)$, and the Burr III marginal distributions with $(\alpha, \beta, \delta) = (1.286, 0.522, 7.398)$.

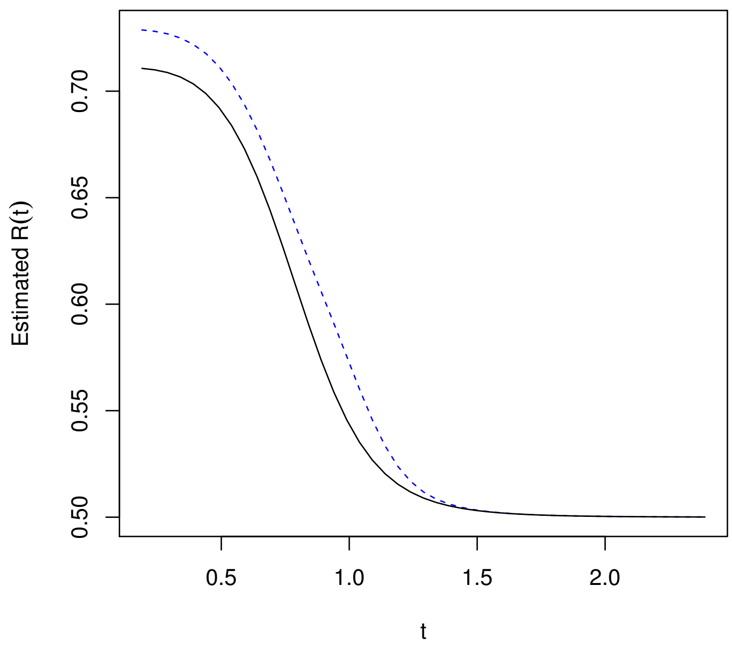

Figure 7: Plot of $R(t)$ estimated from the CHNS data set based on the GFGM copula with $(m_1, m_2, p_1, p_2) = (3, 3, 2, 2)$, and the Burr III marginal distributions with $(\alpha, \beta, \delta) = (1.286, 0.522, 7.398)$. Black/solid and blue/dashed curves relate to $\theta = 0$ and $\theta = 0.176$, respectively.

ACKNOWLEDGMENTS

The authors wish to thank the Associate Editor and the reviewer for their constructive comments that helped us to improve this article significantly.

REFERENCES

[1] Bairamov, I.G.; Kotz, S. and Bekci, M. (2001). New generalized Farlie–Gumbel–Morgenstern distributions and concomitants of order statistics, *Journal of Applied Statistics*, **28**, 521–536.

[2] Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph, *Journal of Mathematical Psychology*, **12**, 387–415.

[3] Burr, I.W. (1942). Cumulative frequency functions, *The Annals of Mathematical Statistics*, **13**, 215–232.

[4] Domma, F. and Giordano, S. (2013). A copula-based approach to account for dependence in stress-strength models, *Statistical Papers*, **54**, 807–826.

[5] Escarela, G. and Carrière, J.F. (2003). Fitting competing risks with an assumed copula, *Statistical Methods in Medical Research*, **12**, 333–349.

[6] Frees, E.W. and Wang, P. (2005). Credibility using copulas, *North American Actuarial Journal*, **9**, 31–48.

[7] Genest, C. and Favre, A.C. (2007). Everything you always wanted to know about copula modeling but were afraid to ask, *Journal of Hydrologic Engineering*, **12**, 347–368.

[8] Genest, C.; Rémillard, B. and Beaudoin, D. (2009). Goodness-of-fit tests for copulas: a review and a power study, *Insurance: Mathematics and Economics*, **44**, 199–213.

[9] Gupta, R.C.; Ghitany, M.E. and Al-Mutairi, D.K. (2013). Estimation of reliability from a bivariate log-normal data, *Journal of Statistical Computation and Simulation*, **83**, 1068–1081.

[10] Gupta, R.C. and Subramanian, S. (1998). Estimation of reliability in a bivariate normal distribution with equal coefficients of variation, *Communications in Statistics: Simulation and Computation*, **27**, 675–698.

[11] Kotz, S.; Lumelskii, Y. and Pensky, M. (2003). *The stress-strength model and its generalizations. Theory and applications*, World Scientific, Singapore.

[12] Mahdizadeh, M. and Zamanzade, E. (2018). A new reliability measure in ranked set sampling, *Statistical Papers*, **59**, 861–891.

[13] Mahdizadeh, M. and Zamanzade, E. (2018). Smooth estimation of a reliability function in ranked set sampling, *Statistics*, **52**, 750–768.

[14] Mahdizadeh, M. and Zamanzade, E. (2019). Dynamic reliability estimation in a rank-based design, *Probability and Mathematical Statistics*, **39**, 1–18.

[15] Nadarajah, S. (2005). Reliability for some bivariate beta distributions, *Mathematical Problems in Engineering*, **2005**, Art. ID 638187, 11 pages.

[16] Nadarajah, S. and Kotz, S. (2006). Reliability for some bivariate exponential distributions, *Mathematical Problems in Engineering*, **2006**, Art. ID 41652, 14 pages.

[17] Nelsen, R.B. (2006). *An Introduction to Copulas*, Springer, New York.

[18] Shao, Q. (2000). Estimation for hazardous concentrations based on NOEC toxicity data: an alternative approach, *Environmetrics*, **11**, 583–595.

[19] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges, *Publications de l'Institut de Statistique de l'Université de Paris*, **8**, 229–231.

[20] Ventura, L. and Racugno, W. (2011). Recent advances on Bayesian inference for $P(X<Y)$, *Bayesian Analysis*, **6**, 411–428.

[21] Walldius, G. and Jungner, I. (2006). The apoB/apoA-I ratio: a strong, new risk factor for cardiovascular disease and a target for lipid-lowering therapy: a review of the evidence, *Journal of Internal Medicine*, **259**, 493–519.

[22] Weerahandi, S. and Johnson, R.A. (1992). Testing reliability in stress-strength model when X and Y are normally distributed, *Technometrics*, **34**, 83–91.

[23] Wolfe, D.A. and Hogg, R.V. (1971). On constructing statistics and reporting data, *The American Statistician*, **25**, 27–30.

[24] Zardasht, V. and Asadi, M. (2010). Evaluation of $P(X_t > Y_t)$ when both X_t and Y_t are residual lifetimes of two systems, *Statistica Neerlandica*, **64**, 460–481.

[25] Zimmer, W.J.; Keats, J.B. and Wang, F.K. (1998). The Burr XII distribution in reliability analysis, *Journal of Quality & Technology*, **30**, 386–394.

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 33–50

GENERALIZED MATRIX t DISTRIBUTION BASED ON NEW MATRIX GAMMA DISTRIBUTION

Authors: Anis Iranmanesh *
– Department of Mathematics and Statistics, Mashhad Branch, Islamic Azad University, Mashhad, Iran
iranmanesh0030@mshdiau.ac.ir

Daya K. Nagar
– Instituto de Matemáticas, Universidad de Antioquia, Medellín, Colombia
dayaknagar@yahoo.com

Sara Shokri
– Department of Mathematics and Statistics, Mashhad Branch, Islamic Azad University, Mashhad, Iran
sara.shokri30@gmail.com

Saralees Nadarajah
– School of Mathematics, University of Manchester, United Kingdom
Saralees.Nadarajah@manchester.ac.uk

Received: January 2019 Revised: October 2019 Accepted: November 2019

Abstract:

- In this paper, a generalized matrix variate gamma distribution, which includes a trace function in the kernel of the density, is introduced. Some important statistical properties including Laplace transform, distributions of functions, expected value of the determinant and expected value of zonal polynomial of the generalized gamma matrix are derived. Further, by using the distribution of the inverse of this newly defined generalized gamma matrix as the prior for the characteristic matrix of a matrix variate normal distribution, a new generalized matrix t type family of distributions is generated. Some important statistical characteristics of this family are also exhibited.

Keywords:

- *Bayes estimator; generalized matrix gamma distribution; generalized matrix t distribution; zonal polynomial.*

AMS Subject Classification:

- 62E05, 62H99.

*Corresponding author.

1. INTRODUCTION

In 1964, Lukacs and Laha defined the matrix variate gamma (MG) distribution. In multivariate statistical analysis, the MG distribution has been the subject of considerable interest, study, and applications for many years. For example, the Wishart distribution, which is the distribution of the sample variance covariance matrix when sampling from a multivariate normal distribution, is a special case of the MG distribution. Applications of the MG distribution have included: damping modeling (Adhikari [1]); models for stochastic upscaling for inelastic material behavior from limited experimental data (Das and Ghanem [7], [8]); models for fusion yield [15]; models for uncertainty quantification (Pascual and Adhikari [31]); characterizing the distribution of anisotropic micro-structural environments with diffusion-weighted imaging (Scherrer *et al.* [32]); models for magnetic tractography (Chamberland *et al.* [4]); models for diffusion compartment imaging (Scherrer *et al.* [33]); models for image classification (Luo *et al.* [24]); models for accurate signal reconstruction (Jian *et al.* [22], Bates *et al.* [3]). Two recent applications of the Wishart distribution can be found in Arashi *et al.* [2] and Ferreira *et al.* [13].

However, generalizations of the MG distribution have been neglected and there is no account on this matter in the literature. The only extension that we are aware of is the inverted matrix variate gamma distribution due to Iranmanesh *et al.* [17]: if $\boldsymbol{X}$ has the MG distribution then $\boldsymbol{X}^{-1}$ has the inverted matrix variate gamma distribution. A generalization of the MG distribution must contain the MG distribution as a particular case. See also Iranmanesh *et al.* [18] and references there in for more details.

The goal of this paper is to give the first generalization to the MG distribution, where its kernel includes zonal polynomials (Takemura [34]). The generalization proposed has two shape parameters. One of the shape parameters acts on the determinant of the data while the other acts on the trace of the data. The MG distribution has only one shape parameter acting on the determinant of the data. The proposed generalization can be more flexible for data modeling:

- **i)** if both trace and determinant are significant (that is, the empirical distribution of the data has significant patterns involving both the trace and determinant);
- **ii)** if trace is significant but determinant is not (that is, the empirical distribution of the data has significant patterns involving only the trace);
- **iii)** if trace is more significant than determinant is (that is, the empirical distribution of the data has more significant patterns involving the trace).

For our purpose, we first provide the reader with some preliminary definitions and lemmas. Most of the following definitions and results can be found in Gupta and Nagar [14], Muirhead [27], and Mathai [26].

2. PRELIMINARIES

In this section we state certain well known definitions and results. These results will be used in subsequent sections.

Let $\boldsymbol{A} = (a_{ij})$ be a $p \times p$ matrix. Then, $\boldsymbol{A}'$ denotes the transpose of $\boldsymbol{A}$; $\operatorname{tr}(\boldsymbol{A}) = a_{11} + \cdots + a_{pp}$; $\operatorname{etr}(\boldsymbol{A}) = \exp(\operatorname{tr}(\boldsymbol{A}))$; $\det(\boldsymbol{A})$ = determinant of $\boldsymbol{A}$; norm of $\boldsymbol{A} = \|\boldsymbol{A}\|$ = maximum of absolute values of eigenvalues of the matrix $\boldsymbol{A}$; $\boldsymbol{A}^{1/2}$ denotes a symmetric positive definite square root of $\boldsymbol{A}$; $\boldsymbol{A} > \boldsymbol{0}$ means that $\boldsymbol{A}$ is symmetric positive definite and $\boldsymbol{0} < \boldsymbol{A} < \boldsymbol{I}_p$ means that the matrices $\boldsymbol{A}$ and $\boldsymbol{I}_p - \boldsymbol{A}$ are symmetric positive definite. The multivariate gamma function which is frequently used in multivariate statistical analysis is defined by

$$
\begin{aligned}
\Gamma_p(a) &= \int_{\boldsymbol{X}>\boldsymbol{0}} \operatorname{etr}(-\boldsymbol{X}) \det(\boldsymbol{X})^{a-(p+1)/2}\, d\boldsymbol{X} \\
&= \pi^{p(p-1)/4} \prod_{i=1}^{p} \Gamma\left(a - \frac{i-1}{2}\right), \qquad \operatorname{Re}(a) > \frac{p-1}{2}\,. \qquad (2.1)
\end{aligned}
$$

Let $C_\kappa(\boldsymbol{X})$ be the zonal polynomial of $p \times p$ complex symmetric matrix $\boldsymbol{X}$ corresponding to the ordered partition $\kappa = (k_1, \ldots, k_p)\,, k_1 \geq \cdots \geq k_p \geq 0$, $k_1 + \cdots + k_p = k$ and $\sum_\kappa$ denotes summation over all partitions κ of k. The generalized hypergeometric coefficient $(a)_\kappa$ used above is defined by

$$
(a)_\kappa = \prod_{i=1}^{p} \left(a - \frac{i-1}{2}\right)_{k_i},
$$

where $(a)_r = a(a+1)\cdots(a+r-1)$, $r = 1, 2, \ldots$ with $(a)_0 = 1$.

Lemma 2.1. *Let $\boldsymbol{Z}$ be a complex symmetric $p \times p$ matrix with $\operatorname{Re}(\boldsymbol{Z}) > \boldsymbol{0}$, and let $\boldsymbol{Y}$ be a symmetric $p \times p$ matrix. Then, for $\operatorname{Re}(a) > (p-1)/2$, we have*

$$
\int_{\boldsymbol{X}>\boldsymbol{0}} \operatorname{etr}(-\boldsymbol{X}\boldsymbol{Z}) (\det \boldsymbol{X})^{a-(p+1)/2} C_\kappa(\boldsymbol{X}\boldsymbol{Y})\, d\boldsymbol{X} = (a)_\kappa\, \Gamma_p(a)\, (\det \boldsymbol{Z})^{-a} C_\kappa\left(\boldsymbol{Y}\boldsymbol{Z}^{-1}\right). \qquad (2.2)
$$

Lemma 2.2. *Let $\boldsymbol{Z}$ be a complex symmetric $p \times p$ matrix with $\operatorname{Re}(\boldsymbol{Z}) > \boldsymbol{0}$, and let $\boldsymbol{Y}$ be a symmetric $p \times p$ matrix. Then, for $\operatorname{Re} > (p-1)/2$, we have*

$$
\int_{\boldsymbol{X}>\boldsymbol{0}} \operatorname{etr}(-\boldsymbol{X}\boldsymbol{Z}) (\det \boldsymbol{X})^{a-(p+1)/2} \left[\operatorname{tr}(\boldsymbol{X}\boldsymbol{Y})\right]^k d\boldsymbol{X} = \Gamma_p(a)\, (\det \boldsymbol{Z})^{-a} \sum_\kappa (a)_\kappa C_\kappa(\boldsymbol{Y}\boldsymbol{Z}^{-1})\,. \qquad (2.3)
$$

For $\boldsymbol{Z} = \boldsymbol{Y}$ in (2.3), we get

$$
\begin{aligned}
\int_{\boldsymbol{X}>\boldsymbol{0}} \operatorname{etr}(-\boldsymbol{X}\boldsymbol{Y}) (\det \boldsymbol{X})^{a-(p+1)/2} \left[\operatorname{tr}(\boldsymbol{X}\boldsymbol{Y})\right]^k d\boldsymbol{X} &= \Gamma_p(a)\, (\det \boldsymbol{Y})^{-a} \sum_\kappa (a)_\kappa C_\kappa(\boldsymbol{I}_p) \\
&= \Gamma_p(a)\, (ap)_k\, (\det \boldsymbol{Y})^{-a}\,. \qquad (2.4)
\end{aligned}
$$

The above result was derived by Khatri [23].

Davis [9, 10] introduced a class of polynomials $C_\phi^{\kappa,\lambda}(\boldsymbol{X},\boldsymbol{Y})$ of $p\times p$ symmetric matrix arguments $\boldsymbol{X}$ and $\boldsymbol{Y}$, which are invariant under the transformation $(\boldsymbol{X},\boldsymbol{Y}) \rightarrow (\boldsymbol{HXH'},\boldsymbol{HYH'})$, $\boldsymbol{H}\in O(p)$. For properties and applications of invariant polynomials we refer to Davis [9, 10], Chikuse [5] and Nagar and Gupta [28]. Let κ, λ, ϕ and ρ be ordered partitions of non-negative integers k, ℓ, $f=k+\ell$ and r, respectively, into not more than p parts. Then

$$C_\phi^{\kappa,\lambda}(\boldsymbol{X},\boldsymbol{X}) = \theta_\phi^{\kappa,\lambda} C_\phi(\boldsymbol{X}), \qquad \theta_\phi^{\kappa,\lambda} = \frac{C_\phi^{\kappa,\lambda}(\boldsymbol{I}_p,\boldsymbol{I}_p)}{C_\phi(\boldsymbol{I}_p)}, \tag{2.5}$$

$$C_\phi^{\kappa,\lambda}(\boldsymbol{X},I_p) = \theta_\phi^{\kappa,\lambda} \frac{C_\phi(\boldsymbol{I}_p)\,C_\kappa(\boldsymbol{X})}{C_\kappa(\boldsymbol{I}_p)}, \qquad C_\phi^{\kappa,\lambda}(I_p,\boldsymbol{Y}) = \theta_\phi^{\kappa,\lambda} \frac{C_\phi(\boldsymbol{I}_p)\,C_\lambda(\boldsymbol{X})}{C_\lambda(\boldsymbol{I}_p)}, \tag{2.6}$$

$$C_\kappa^{\kappa,0}(\boldsymbol{X},\boldsymbol{Y}) \equiv C_\kappa(\boldsymbol{X}), \qquad C_\lambda^{0,\lambda}(\boldsymbol{X},\boldsymbol{Y}) \equiv C_\lambda(\boldsymbol{Y})$$

and

$$C_\kappa(\boldsymbol{X})\,C_\lambda(\boldsymbol{Y}) = \sum_{\phi\in\kappa\cdot\lambda} \theta_\phi^{\kappa,\lambda} C_\phi^{\kappa,\lambda}(\boldsymbol{X},\boldsymbol{Y}), \tag{2.7}$$

where $\phi\in\kappa\cdot\lambda$ signifies that irreducible representation of $Gl(p,R)$ indexed by 2ϕ, occurs in the decomposition of the Kronecker product $2\kappa\otimes 2\lambda$ of the irreducible representations indexed by 2κ and 2λ. Further,

$$\begin{aligned}\int_{\boldsymbol{R}>\boldsymbol{0}} \operatorname{etr}(-\boldsymbol{CR})\,\det(\boldsymbol{R})^{t-(p+1)/2}\,C_\phi^{\kappa,\lambda}\left(\boldsymbol{ARA'},\boldsymbol{BRB'}\right)\,d\boldsymbol{R} &=\\ = \Gamma_p(t,\phi)\,\det(\boldsymbol{C})^{-t}\,C_\phi^{\kappa,\lambda}&\left(\boldsymbol{AC}^{-1}\boldsymbol{A'},\boldsymbol{BC}^{-1}\boldsymbol{B'}\right),\end{aligned} \tag{2.8}$$

$$\begin{aligned}\int_{\boldsymbol{0}}^{\boldsymbol{I}_p} \det(\boldsymbol{R})^{t-(p+1)/2}\,\det(\boldsymbol{I}_p-\boldsymbol{R})^{u-(p+1)/2}\,C_\phi^{\kappa,\lambda}(\boldsymbol{R},\boldsymbol{I}_p-\boldsymbol{R})\,d\boldsymbol{R} &=\\ = \frac{\Gamma_p(t,\kappa)\,\Gamma_p(u,\lambda)}{\Gamma_p(t+u,\phi)}\,&\theta_\phi^{\kappa,\lambda}\,C_\phi(\boldsymbol{I}_p)\end{aligned} \tag{2.9}$$

and

$$\begin{aligned}\int_{\boldsymbol{0}}^{\boldsymbol{I}_p} \det(\boldsymbol{R})^{t-(p+1)/2}\,\det(\boldsymbol{I}_p-\boldsymbol{R})^{u-(p+1)/2}\,C_\phi^{\kappa,\lambda}(\boldsymbol{AR},\boldsymbol{BR})\,d\boldsymbol{R} &=\\ = \frac{\Gamma_p(t,\phi)\,\Gamma_p(u)}{\Gamma_p(t+u,\phi)}\,&C_\phi^{\kappa,\lambda}(\boldsymbol{A},\boldsymbol{B}).\end{aligned} \tag{2.10}$$

In expressions (2.8), (2.9) and (2.10), $\Gamma_p(a,\rho)$ is defined by

$$\Gamma_p(a,\rho) = (a)_\rho\,\Gamma_p(a). \tag{2.11}$$

Note that $\Gamma_p(a,0)=\Gamma_p(a)$, which is the multivariate gamma function.

Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{X}$ and $\boldsymbol{Y}$ be $p\times p$ symmetric matrices. Then

$$\int_{\boldsymbol{H}\in O(p)} C_\phi^{\kappa,\lambda}\big(\boldsymbol{AH}'\boldsymbol{XH},\boldsymbol{BH}'\boldsymbol{YH}\big)\,[d\boldsymbol{H}] \;=\; \frac{C_\phi^{\kappa,\lambda}(\boldsymbol{A},\boldsymbol{B})\,C_\phi^{\kappa,\lambda}(\boldsymbol{X},\boldsymbol{Y})}{\theta_\phi^{\kappa,\lambda}\,C_\phi(\boldsymbol{I}_p)}\,, \tag{2.12}$$

where $[d\boldsymbol{H}]$ is the unit invariant Haar measure. The above result is a generalization of Davis [10, Eq. 5.4] and is due to Díaz-García [11]. Finally, using (2.9) and (2.12), it is straightforward to see that

$$\begin{aligned}\int_{\boldsymbol{0}}^{\boldsymbol{I}_p} \det(\boldsymbol{R})^{t-(p+1)/2}\,\det(\boldsymbol{I}_p-\boldsymbol{R})^{u-(p+1)/2}\,C_\phi^{\kappa,\lambda}\big(\boldsymbol{AR},\boldsymbol{B}\,(\boldsymbol{I}_p-\boldsymbol{R})\big)\,d\boldsymbol{R} \;=&\\ =\;\frac{\Gamma_p(t,\kappa)\,\Gamma_p(u,\lambda)}{\Gamma_p(t+u,\phi)}\,C_\phi^{\kappa,\lambda}(\boldsymbol{A},\boldsymbol{B})\,.&\end{aligned} \tag{2.13}$$

Definition 2.1. The $n\times p$ random matrix $\boldsymbol{X}$ is said to have a matrix variate normal distribution with $n\times p$ mean matrix $\boldsymbol{M}$ and $np\times np$ covariance matrix $\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma}$, denoted by $\boldsymbol{X}\sim N_{n,p}\,(\boldsymbol{M},\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$, if its probability density function (p.d.f) is given by (Gupta and Nagar [14])

$$(2\pi)^{-np/2}\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Sigma})^{-n/2}\exp\left\{-\frac{1}{2}\,\mathrm{tr}\big[\boldsymbol{\Omega}^{-1}(\boldsymbol{X}-\boldsymbol{M})\,\boldsymbol{\Sigma}^{-1}(\boldsymbol{X}-\boldsymbol{M})'\big]\right\},\qquad \boldsymbol{X}\in\mathbb{R}^{n\times p},\quad \boldsymbol{M}\in\mathbb{R}^{n\times p},$$

where $\boldsymbol{\Sigma}(p\times p)>\boldsymbol{0}$ and $\boldsymbol{\Omega}(n\times n)>\boldsymbol{0}$.

If $\boldsymbol{X}\sim N_{n,p}\,(\boldsymbol{M},\boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$, then the characteristic function of $\boldsymbol{X}$ is

$$\begin{aligned}\phi_{\boldsymbol{X}}(\boldsymbol{Z}) &= E\big[\exp\big(\mathrm{tr}\,(\iota\,\boldsymbol{Z}'\boldsymbol{X})\big)\big]\\ &= \exp\left[\mathrm{tr}\left(\iota\,\boldsymbol{Z}'\boldsymbol{M}-\frac{1}{2}\,\boldsymbol{Z}'\boldsymbol{\Omega}\boldsymbol{Z}\boldsymbol{\Sigma}\right)\right],\qquad \boldsymbol{Z}\in\mathbb{R}^{n\times p},\quad \iota=\sqrt{-1}\,.\end{aligned}$$

The present paper has been organized in the following sections. In Section 3, a new generalized matrix gamma (GMG) distribution has been defined. Some important properties of this newly defined distribution are given in Section 4. In Section 5, using the conditioning approach for the matrix variate normal distribution, a new matrix t type family of distributions is introduced. Some important statistical characteristics of this family are studied in Section 6. A Bayesian application is given in Section 7. The paper is concluded in Section 8.

3. GENERALIZED MATRIX GAMMA DISTRIBUTION

Recently, Nagar *et al.* [30] defined a generalized matrix variate gamma distribution by generalizing the multivariate gamma function. We also refer to Nagar *et al.* [29] for further generalizations. In this paper, by incorporating an additional factor in the p.d.f, we give a generalization of the matrix variate gamma distribution (Das and Dey [6], Iranmanesh *et al.* [17]).

In the following we provide the reader with the definition of the generalized matrix variate gamma distribution.

Definition 3.1. A random symmetric matrix $\boldsymbol{X}$ of order p is said to have a generalized matrix gamma (GMG) distribution with parameters α, β, k, $\boldsymbol{\Sigma}$ and $\boldsymbol{U}$, denoted by $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})$, if its p.d.f is given by

$$C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})\ \mathrm{etr}\left(-\frac{1}{\beta}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right)\det(\boldsymbol{X})^{\alpha-(p+1)/2}\big[\mathrm{tr}(\boldsymbol{X}\boldsymbol{U})\big]^k, \qquad \boldsymbol{X} > \boldsymbol{0}, \tag{3.1}$$

where $\alpha > (p-1)/2$, $\beta > 0$, $\boldsymbol{\Sigma} > \boldsymbol{0}$, $\boldsymbol{U} > \boldsymbol{0}$, $k \in \mathbb{N}_0$ and $C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$ is the normalizing constant.

By integrating the p.d.f of $\boldsymbol{X}$ over its support set, the normalizing constant $C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$ can be evaluated as

$$\begin{aligned}\big[C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})\big]^{-1} &= \int_{\boldsymbol{X}>\boldsymbol{0}} \mathrm{etr}\left(-\frac{1}{\beta}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\right)\det(\boldsymbol{X})^{\alpha-(p+1)/2}\big[\mathrm{tr}(\boldsymbol{X}\boldsymbol{U})\big]^k\,d\boldsymbol{X}\\ &= \beta^{p\alpha+k}\,\Gamma_p(\alpha)\,\det(\boldsymbol{\Sigma})^{\alpha}\sum_{\kappa}(\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Sigma}),\end{aligned} \tag{3.2}$$

where the last line has been obtained by using (2.3).

The distribution given by the p.d.f (3.1) is a generalization of the matrix variate gamma distribution (Das and Dey [6], Iranmanesh *et al.* [17]). For $\boldsymbol{U} = \boldsymbol{\Sigma}^{-1}$, the p.d.f in (3.1) simplifies to

$$\frac{\mathrm{etr}\big(-\boldsymbol{\Sigma}^{-1}\boldsymbol{X}/\beta\big)\det(\boldsymbol{X})^{\alpha-(p+1)/2}\big[\mathrm{tr}\big(\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\big)\big]^k}{\beta^{\alpha p+k}(\alpha p)_k\,\Gamma_p(\alpha)\,\det(\boldsymbol{\Sigma})^{\alpha}}, \qquad \boldsymbol{X} > \boldsymbol{0}. \tag{3.3}$$

Further, for $\boldsymbol{U} = \boldsymbol{0}$ or $k = 0$ the p.d.f (3.1) reduces to the matrix variate gamma p.d.f given by

$$\frac{\mathrm{etr}\big(-\boldsymbol{\Sigma}^{-1}\boldsymbol{X}/\beta\big)\det(\boldsymbol{X})^{\alpha-(p+1)/2}}{\beta^{\alpha p}\,\Gamma_p(\alpha)\,\det(\boldsymbol{\Sigma})^{\alpha}}, \qquad \boldsymbol{X} > \boldsymbol{0}. \tag{3.4}$$

By suitably choosing β we can derive a number of special cases of (3.3). If we choose $\alpha = n/2$ and $\beta = 2$, then $\boldsymbol{X}$ has a generalized Wishart distribution with p.d.f

$$\frac{\mathrm{etr}\big(-\boldsymbol{\Sigma}^{-1}\boldsymbol{X}/2\big)\det(\boldsymbol{X})^{n/2-(p+1)/2}\big[\mathrm{tr}\big(\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\big)\big]^k}{2^{np/2+k}\,\Gamma_p(n/2)\,(np/2)_k\,\det(\boldsymbol{\Sigma})^{n/2}}, \qquad \boldsymbol{X} > \boldsymbol{0}. \tag{3.5}$$

Note that n is a positive integer, generally considered as the sample size. If we choose $\boldsymbol{\Sigma} = \boldsymbol{I}_p$, $\beta = 2$ and $p = 1$ in (3.3), then the scalar variable X follows a chi-square distribution with $n + 2k$ degrees of freedom. Further, if we take $p = 1$ and $\beta = 1$ in (3.3), then the scalar variable X follows a univariate gamma distribution with shape parameter $\alpha + k$ and scale parameter σ. Finally, for $\boldsymbol{\Sigma} = \boldsymbol{I}_p$ and $p = 1$, the scalar variable X follows a univariate gamma distribution with shape parameter $\alpha + k$ and scale parameter β.

Definition 3.2. If $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$ then $\boldsymbol{X}^{-1}$ is said to have an inverted generalized matrix gamma (IGMG) distribution with parameters α, β, k, $\boldsymbol{\Sigma}^{-1}$ and $\boldsymbol{U}$, denoted by $\boldsymbol{X}^{-1} \sim \mathrm{IGMG}_p\big(\alpha,\beta,k,\boldsymbol{\Sigma}^{-1},\boldsymbol{U}\big)$.

In the following theorem, the p.d.f of the IGMG distribution is derived.

Proposition 3.1. *Let* $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})$*. Then,* $\boldsymbol{Y} = \boldsymbol{X}^{-1} \sim \mathrm{IGMG}_p(\alpha, \beta, k, \boldsymbol{\Sigma}^{-1}, \boldsymbol{U})$ *has the p.d.f given by*

$$C\big(\alpha, \beta, k, \boldsymbol{\Sigma}^{-1}, \boldsymbol{U}\big)\ \mathrm{etr}\left(-\frac{1}{\beta}\,\boldsymbol{\Sigma}\boldsymbol{Y}^{-1}\right) \det(\boldsymbol{Y})^{-\alpha-(p+1)/2} \left[\mathrm{tr}\big(\boldsymbol{Y}^{-1}\boldsymbol{U}\big)\right]^k, \qquad \boldsymbol{Y} > \boldsymbol{0}, \tag{3.6}$$

where $\alpha > (p-1)/2$, $\beta > 0$, $\boldsymbol{\Sigma} > \boldsymbol{0}$, $\boldsymbol{U} > \boldsymbol{0}$, $k \in \mathbb{N}_0$ *and* $C\left(\alpha, \beta, k, \boldsymbol{\Sigma}^{-1}, \boldsymbol{U}\right)$ *is the normalizing constant.*

Proof: The proof follows from the fact that the Jacobian of the transformation $\boldsymbol{Y} = \boldsymbol{X}^{-1}$ is given by $J(\boldsymbol{X} \to \boldsymbol{Y}) = \det(\boldsymbol{Y})^{-(p+1)}$. □

By taking $\boldsymbol{U} = \boldsymbol{\Sigma}$, $\alpha = n/2$ and $\beta = 2$ in (3.6), the inverted generalized Wishart p.d.f can be obtained as

$$\frac{\mathrm{etr}\big(-\boldsymbol{\Sigma}\boldsymbol{Y}^{-1}/2\big)\ \det(\boldsymbol{Y})^{-n/2-(p+1)/2} \left[\mathrm{tr}\big(\boldsymbol{\Sigma}\boldsymbol{Y}^{-1}\big)\right]^k}{2^{np/2+k}\ \Gamma_p(n/2)\,(np/2)_k\ \det(\boldsymbol{\Sigma})^{-n/2}}, \qquad \boldsymbol{Y} > \boldsymbol{0}. \tag{3.7}$$

4. PROPERTIES OF GMG AND IGMG DISTRIBUTIONS

In this section, various properties of the GMG and IGMG distributions are derived.

Proposition 4.1. *Let* $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})$*. Then, the Laplace transform of* $\boldsymbol{X}$ *is*

$$\varphi_{\boldsymbol{X}}(\boldsymbol{T}) \;=\; \det(\boldsymbol{I}_p + \beta\boldsymbol{\Sigma}\boldsymbol{T})^{-\alpha}\, \frac{\sum_\kappa (\alpha)_\kappa\, C_\kappa\Big(\boldsymbol{U}\big(\beta\boldsymbol{T} + \boldsymbol{\Sigma}^{-1}\big)^{-1}\Big)}{\sum_\kappa (\alpha)_\kappa\, C_\kappa(\boldsymbol{U}\boldsymbol{\Sigma})}, \tag{4.1}$$

where $\boldsymbol{T}$ *is a complex symmetric matrix of order* p *with* $\mathrm{Re}(\boldsymbol{T}) > \boldsymbol{0}$.

Proof: By definition, we have

$$\begin{aligned}
\varphi_{\boldsymbol{X}}(\boldsymbol{T}) \;&=\; E\big[\exp\big(-\,\mathrm{tr}(\boldsymbol{T}\boldsymbol{X})\big)\big] \\
&=\; C(\alpha, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U}) \int_{\boldsymbol{X}>\boldsymbol{0}} \mathrm{etr}\left[-\boldsymbol{X}\left(\boldsymbol{T} + \frac{1}{\beta}\,\boldsymbol{\Sigma}^{-1}\right)\right] \det(\boldsymbol{X})^{\alpha-(p+1)/2} \left[\mathrm{tr}(\boldsymbol{X}\boldsymbol{U})\right]^k d\boldsymbol{X}\,.
\end{aligned}$$

Now, evaluating the above integral by using (3.2) and simplifying the resulting expression, we get the desired result. □

Corollary 4.0.1. *Let* $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$*. Then the characteristic function of* $\boldsymbol{X}$ *is*

$$\psi_{\boldsymbol{X}}(\boldsymbol{T}) = \det(\boldsymbol{I}_p - \iota\,\beta\boldsymbol{\Sigma}\boldsymbol{T})^{-\alpha}\, \frac{\sum_\kappa (\alpha)_\kappa\, C_\kappa\Big(\boldsymbol{U}\boldsymbol{\Sigma}\,(\boldsymbol{I}_p - \iota\,\beta\boldsymbol{T}\boldsymbol{\Sigma})^{-1}\Big)}{\sum_\kappa (\alpha)_\kappa\, C_\kappa(\boldsymbol{U}\boldsymbol{\Sigma})}\,, \tag{4.2}$$

where $\iota = \sqrt{-1}$, $\boldsymbol{T}$ *is a symmetric positive definite matrix of order* p *with* $\boldsymbol{T} = ((1+\delta_{ij})\,t_{ij}/2)$ *and* δ_{ij} *is the Kronecker delta.*

Proposition 4.2. *If* $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$*, then for a* $p\times p$ *non-singular constant matrix* $\boldsymbol{A}$*, we have*

$$\boldsymbol{A}\boldsymbol{X}\boldsymbol{A}' \sim \mathrm{GMG}_p\Big(\alpha,\beta,k,\boldsymbol{A}\boldsymbol{\Sigma}\boldsymbol{A}', \big(\boldsymbol{A}^{-1}\big)'\boldsymbol{U}\boldsymbol{A}^{-1}\Big)\,.$$

Proposition 4.3. *Let* $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$*. Then*

$$E\big[\det(\boldsymbol{X})^h\big] = \det(\beta\boldsymbol{\Sigma})^h\, \frac{\Gamma_p(\alpha+h)}{\Gamma_p(\alpha)}\, \frac{\sum_\kappa (\alpha+h)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Sigma})}{\sum_\kappa (\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Sigma})}\,.$$

Proof: By definition

$$\begin{aligned} E\big[\det(\boldsymbol{X})^h\big] &= C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U}) \int_{\boldsymbol{X}>\boldsymbol{0}} \mathrm{etr}\Big(-\frac{1}{\beta}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\Big) \det(\boldsymbol{X})^{\alpha+h-(p+1)/2}\, \big[\mathrm{tr}(\boldsymbol{X}\boldsymbol{U})\big]^k\, d\boldsymbol{X} \\ &= \frac{C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})}{C(\alpha+h,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})}\,. \end{aligned}$$

Now, simplification of the above expression yields the desired result. □

Proposition 4.4. *If* $\boldsymbol{X} \sim \mathrm{GMG}_p\big(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{\Sigma}^{-1}\big)$*. Then*

$$E\big[\det(\boldsymbol{X})^h\big] = \det(\beta\boldsymbol{\Sigma})^h\, \frac{\Gamma_p(\alpha+h)}{\Gamma_p(\alpha)}\, \frac{(\alpha p + hp)_k}{(\alpha p)_k}\,.$$

In order to find the expectation of the trace of a GMG random matrix, it is useful to find the expectation of zonal polynomials, which is derived below.

Theorem 4.1. *Let* $\boldsymbol{X} \sim \mathrm{GMG}_p(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})$ *and* $\boldsymbol{B}$ *be a constant symmetric matrix of order* p*. Then*

$$E\big[C_\tau(\boldsymbol{X}\boldsymbol{B})\big] = C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})\,\beta^{p\alpha+t+k}\det(\boldsymbol{\Sigma})^\alpha \sum_\kappa \sum_{\phi\in\kappa\cdot\tau} \theta_\phi^{\kappa,\tau}\,\Gamma_p(\alpha,\phi)\, C_\phi^{\kappa,\tau}(\boldsymbol{U}\boldsymbol{\Sigma},\boldsymbol{B}\boldsymbol{\Sigma})\,.$$

Proof: By definition, we have

$$\begin{aligned} E\big[C_\tau(\boldsymbol{X}\boldsymbol{B})\big] &= C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U}) \\ &\quad\times \int_{\boldsymbol{X}>\boldsymbol{0}} \mathrm{etr}\Big(-\frac{1}{\beta}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\Big) \det(\boldsymbol{X})^{\alpha-(p+1)/2}\, \big[\mathrm{tr}(\boldsymbol{X}\boldsymbol{U})\big]^k\, C_\tau(\boldsymbol{X}\boldsymbol{B})\, d\boldsymbol{X}\,. \end{aligned}$$

Now, writing

$$\begin{aligned}\left[\mathrm{tr}(\boldsymbol{XU})\right]^k C_\tau(\boldsymbol{XB}) &= \sum_\kappa C_\kappa(\boldsymbol{XU})\, C_\tau(\boldsymbol{XB})\\ &= \sum_\kappa \sum_{\phi\in\kappa\cdot\tau} \theta_\phi^{\kappa,\tau}\, C_\phi^{\kappa,\tau}(\boldsymbol{XU},\boldsymbol{XB})\,,\end{aligned}$$

where we have used (2.7), and integrating $\boldsymbol{X}$ by using (2.8), we obtain

$$\begin{aligned}E\big[C_\tau(\boldsymbol{XB})\big] &= C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U}) \sum_\kappa \sum_{\phi\in\kappa\cdot\tau} \theta_\phi^{\kappa,\tau}\\ &\quad\times \int_{\boldsymbol{X}>\boldsymbol{0}} \mathrm{etr}\Big(-\frac{1}{\beta}\,\boldsymbol{\Sigma}^{-1}\boldsymbol{X}\Big)\, \det(\boldsymbol{X})^{\alpha-(p+1)/2}\, C_\phi^{\kappa,\tau}(\boldsymbol{XU},\boldsymbol{XB})\; d\boldsymbol{X}\\ &= C(\alpha,\beta,k,\boldsymbol{\Sigma},\boldsymbol{U})\, \det(\beta\boldsymbol{\Sigma})^\alpha \sum_\kappa \sum_{\phi\in\kappa\cdot\tau} \theta_\phi^{\kappa,\tau}\, \Gamma_p(\alpha,\phi)\, C_\phi^{\kappa,\tau}(\beta\boldsymbol{U\Sigma},\beta\boldsymbol{B\Sigma})\,.\end{aligned}$$

Now, the result follows from the fact that

$$C_\phi^{\kappa,\tau}(\beta\boldsymbol{U\Sigma},\beta\boldsymbol{B\Sigma}) = \beta^{k+t}\, C_\phi^{\kappa,\tau}(\boldsymbol{U\Sigma},\boldsymbol{B\Sigma})\,. \qquad \square$$

Theorem 4.2. *Let $\boldsymbol{Y}\sim \mathrm{IGMG}_p(\alpha,\beta,k,\boldsymbol{\Psi},\boldsymbol{U})$. Then, the Laplace transform of $\boldsymbol{Y}$ is given by*

$$\varphi_{\boldsymbol{Y}}(\boldsymbol{T}) = C\big(\alpha,\beta,k,\boldsymbol{\Psi}^{-1},\boldsymbol{U}\big)\, \det(\boldsymbol{T})^\alpha \left[\frac{d^k}{dz^k}\, B_\alpha\Big(\boldsymbol{T}\big(\beta^{-1}\boldsymbol{\Psi}-z\boldsymbol{U}\big)\Big)\right]_{z=0}, \tag{4.3}$$

where $\boldsymbol{T}$ is a complex symmetric matrix of order p with $\mathrm{Re}(\boldsymbol{T})>\boldsymbol{0}$ and $B_\delta(\cdot)$ is the Bessel function of matrix argument (Herz [16]) given by

$$B_\delta(\boldsymbol{WZ}) = \det(\boldsymbol{W})^{-\delta} \int_{\boldsymbol{S}>\boldsymbol{0}} \det(\boldsymbol{S})^{\delta-(p+1)/2}\, \mathrm{etr}\big(-\boldsymbol{SZ}-\boldsymbol{S}^{-1}\boldsymbol{W}\big)\; d\boldsymbol{S}\,. \tag{4.4}$$

Proof: The Laplace transform of $\boldsymbol{Y}$, denoted by $\varphi_{\boldsymbol{Y}}(\boldsymbol{T})$ can be derived as

$$\begin{aligned}\varphi_{\boldsymbol{Y}}(\boldsymbol{T}) &= C\big(\alpha,\beta,k,\boldsymbol{\Psi}^{-1},\boldsymbol{U}\big)\\ &\quad\times \int_{\boldsymbol{Y}>\boldsymbol{0}} \mathrm{etr}(-\boldsymbol{TY})\, \mathrm{etr}\Big(-\frac{1}{\beta}\,\boldsymbol{\Psi Y}^{-1}\Big)\, \det(\boldsymbol{Y})^{-\alpha-(p+1)/2} \left[\mathrm{tr}\big(\boldsymbol{Y}^{-1}\boldsymbol{U}\big)\right]^k d\boldsymbol{Y}\,.\end{aligned} \tag{4.5}$$

Note that we can write

$$\left[\mathrm{tr}\big(\boldsymbol{Y}^{-1}\boldsymbol{U}\big)\right]^k = \left[\frac{d^k}{dz^k}\, \exp\left[z\,\mathrm{tr}\big(\boldsymbol{Y}^{-1}\boldsymbol{U}\big)\right]\right]_{z=0}. \tag{4.6}$$

Now, substituting (4.6) in (4.5), we have

$$\begin{aligned}\varphi_{\boldsymbol{Y}}(\boldsymbol{T}) &= C\big(\alpha,\beta,k,\boldsymbol{\Psi}^{-1},\boldsymbol{U}\big)\\ &\quad\times \left[\frac{d^k}{dz^k} \int_{\boldsymbol{Y}>\boldsymbol{0}} \mathrm{etr}(-\boldsymbol{TY})\, \mathrm{etr}\left[-\big(\beta^{-1}\boldsymbol{\Psi}-z\boldsymbol{U}\big)\,\boldsymbol{Y}^{-1}\right] \det(\boldsymbol{Y})^{-\alpha-(p+1)/2}\; d\boldsymbol{Y}\right]_{z=0}\\ &= C\big(\alpha,\beta,k,\boldsymbol{\Psi}^{-1},\boldsymbol{U}\big)\\ &\quad\times \left[\frac{d^k}{dz^k} \int_{\boldsymbol{Y}>\boldsymbol{0}} \mathrm{etr}\left[-\boldsymbol{TY}^{-1}-\big(\beta^{-1}\boldsymbol{\Psi}-z\boldsymbol{U}\big)\,\boldsymbol{Y}\right] \det(\boldsymbol{Y})^{\alpha-(p+1)/2}\; d\boldsymbol{Y}\right]_{z=0}.\end{aligned} \tag{4.7}$$

Finally, using (4.4) in (4.7), we get the desired result. $\square$

Proposition 4.5. *Let $\boldsymbol{Y} \sim \mathrm{IGMG}_p(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$. Then*

$$E\big[\det(\boldsymbol{Y})^h\big] = \frac{\det(\boldsymbol{\Psi})^h\,\Gamma_p(\alpha-h)}{\beta^{ph}\,\Gamma_p(\alpha)}\;\frac{\sum_\kappa(\alpha-h)_\kappa\,C_\kappa\big(\boldsymbol{U}\boldsymbol{\Psi}^{-1}\big)}{\sum_\kappa(\alpha)_\kappa\,C_\kappa(\boldsymbol{U}\boldsymbol{\Psi}^{-1})}\,, \qquad \mathrm{Re}(\alpha-h) > \frac{p-1}{2}\,.$$

Proof: By definition,

$$\begin{aligned} E\big[\det(\boldsymbol{Y})^h\big] &= C\big(\alpha, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\big) \\ &\quad \times \int_{\boldsymbol{Y}>\boldsymbol{0}} \mathrm{etr}\left(-\beta^{-1}\boldsymbol{\Psi}\boldsymbol{Y}^{-1}\right)\,\det(\boldsymbol{Y})^{-(\alpha-h)-(p+1)/2}\left[\mathrm{tr}\left(\boldsymbol{Y}^{-1}\boldsymbol{U}\right)\right]^k\,d\boldsymbol{Y} \\ &= \frac{C\big(\alpha, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\big)}{C\left(\alpha-h, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\right)}\,, \qquad \mathrm{Re}(\alpha-h) > \frac{p-1}{2}\,. \end{aligned}$$

Now, the desired result is obtained by simplifying the above expression. □

Proposition 4.6. *Let $\boldsymbol{Y} \sim \mathrm{IGMG}_p(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$ and $\boldsymbol{A}$ be a constant symmetric matrix of order p. Then $\boldsymbol{A}\boldsymbol{Y}\boldsymbol{A}' \sim \mathrm{IGMG}_p(\alpha, \beta, k, \boldsymbol{A}\boldsymbol{\Psi}\boldsymbol{A}', \boldsymbol{A}\boldsymbol{U}\boldsymbol{A}')$.*

Proof: The Jacobian of the transformation $\boldsymbol{Z} = \boldsymbol{A}\boldsymbol{Y}\boldsymbol{A}'$ is $J(\boldsymbol{Y}\!\rightarrow \boldsymbol{Z}) = \det(\boldsymbol{A})^{-(p+1)}$. Substituting appropriately in the p.d.f of $\boldsymbol{Y}$, we get the desired result. □

Theorem 4.3. *Let the $p\times p$ random symmetric matrices $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ be independent, $\boldsymbol{X}_1 \sim \mathrm{GMG}_p(\alpha_1, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})$ and $\boldsymbol{X}_2 \sim \mathrm{GMG}_p(\alpha_2, \beta, l, \boldsymbol{\Sigma}, \boldsymbol{U})$. Define $\boldsymbol{R} = (\boldsymbol{X}_1+\boldsymbol{X}_2)^{-1/2} \times \boldsymbol{X}_1\,(\boldsymbol{X}_1+\boldsymbol{X}_2)^{-1/2}$ and $\boldsymbol{S} = \boldsymbol{X}_1+\boldsymbol{X}_2$. The p.d.f of $\boldsymbol{S}$ is given by*

$$\begin{aligned} &C(\alpha_1, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})\;C(\alpha_2, \beta, l, \boldsymbol{\Sigma}, \boldsymbol{U})\;\mathrm{etr}\Big[-(\beta\boldsymbol{\Sigma})^{-1}\boldsymbol{S}\Big]\;\det(\boldsymbol{S})^{\alpha_1+\alpha_2-(p+1)/2}\times \\ &\qquad\times\sum_\kappa\sum_\lambda\sum_{\phi\in\kappa\cdot\lambda}\theta_\phi^{\kappa,\lambda}\,\frac{\Gamma_p(\alpha_1,\kappa)\;\Gamma_p(\alpha_1,\lambda)}{\Gamma_p(\alpha_1+\alpha_2,\phi)}\;C_\phi^{\kappa,\lambda}(\boldsymbol{S}\boldsymbol{U}, \boldsymbol{S}\boldsymbol{U})\;, \qquad \boldsymbol{S}>\boldsymbol{0}\,. \end{aligned}$$

Further, for $\boldsymbol{U} = \boldsymbol{I}_p$, the p.d.f of $\boldsymbol{R}$ is given by

$$\begin{aligned} &C(\alpha_1, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{I}_p)\;C(\alpha_2, \beta, l, \boldsymbol{\Sigma}, \boldsymbol{I}_p)\;\det(\beta\boldsymbol{\Sigma})^{\alpha_1+\alpha_2}\;\det(\boldsymbol{R})^{\alpha_1-(p+1)/2}\;\det(\boldsymbol{I}_p-\boldsymbol{R})^{\alpha_2-(p+1)/2}\times \\ &\qquad\times\sum_\kappa\sum_\lambda\sum_{\phi\in\kappa\cdot\lambda}\theta_\phi^{\kappa,\lambda}\,\Gamma_p(\alpha_1+\alpha_2,\phi)\;C_\phi^{\kappa,\lambda}\big(\beta\boldsymbol{\Sigma}\boldsymbol{R}, \beta\boldsymbol{\Sigma}\,(\boldsymbol{I}_p-\boldsymbol{R})\big), \quad \boldsymbol{0}<\boldsymbol{R}<\boldsymbol{I}_p\,. \end{aligned}$$

Proof: The joint p.d.f of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ is given by

$$\begin{aligned} &C(\alpha_1, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})\;C(\alpha_2, \beta, l, \boldsymbol{\Sigma}, \boldsymbol{U})\;\mathrm{etr}\Big[-(\beta\boldsymbol{\Sigma})^{-1}\,(\boldsymbol{X}_1+\boldsymbol{X}_2)\Big]\times \\ &\qquad\times\;\det(\boldsymbol{X}_1)^{\alpha_1-(p+1)/2}\,\det(\boldsymbol{X}_2)^{\alpha_2-(p+1)/2}\left[\mathrm{tr}(\boldsymbol{X}_1\boldsymbol{U})\right]^k\left[\mathrm{tr}(\boldsymbol{X}_2\boldsymbol{U})\right]^l, \qquad \boldsymbol{X}_1>\boldsymbol{0},\;\boldsymbol{X}_2>\boldsymbol{0}\,. \end{aligned}$$

Transforming $\boldsymbol{R} = (\boldsymbol{X}_1+\boldsymbol{X}_2)^{-1/2}\,\boldsymbol{X}_1\,(\boldsymbol{X}_1+\boldsymbol{X}_2)^{-1/2}$ and $\boldsymbol{S} = \boldsymbol{X}_1+\boldsymbol{X}_2$ with the Jacobian $J\,(\boldsymbol{X}_1, \boldsymbol{X}_2\!\rightarrow \boldsymbol{R}, \boldsymbol{S}) = \det(\boldsymbol{S})^{(p+1)/2}$ in the joint p.d.f of $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$, the joint p.d.f of $\boldsymbol{R}$ and $\boldsymbol{S}$ can be derived as

$$\begin{aligned} &C(\alpha_1, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})\;C(\alpha_2, \beta, l, \boldsymbol{\Sigma}, \boldsymbol{U})\;\mathrm{etr}\Big[-(\beta\boldsymbol{\Sigma})^{-1}\boldsymbol{S}\Big]\;\det(\boldsymbol{S})^{\alpha_1+\alpha_2-(p+1)/2}\,\det(\boldsymbol{R})^{\alpha_1-(p+1)/2}\times \\ &\qquad\times\;\det(\boldsymbol{I}_p-\boldsymbol{R})^{\alpha_2-(p+1)/2}\left[\mathrm{tr}\big(\boldsymbol{S}^{1/2}\boldsymbol{R}\boldsymbol{S}^{1/2}\boldsymbol{U}\big)\right]^k\left[\mathrm{tr}\big(\boldsymbol{S}^{1/2}\,(\boldsymbol{I}_p-\boldsymbol{R})\,\boldsymbol{S}^{1/2}\boldsymbol{U}\big)\right]^l, \end{aligned} \tag{4.8}$$

where $\boldsymbol{S} > \boldsymbol{0}$ and $\boldsymbol{0} < \boldsymbol{R} < \boldsymbol{I}_p$. Now, writing

$$\left[\operatorname{tr}\left(\boldsymbol{S}^{1/2}\boldsymbol{R}\boldsymbol{S}^{1/2}\boldsymbol{U}\right)\right]^k \left[\operatorname{tr}\left(\boldsymbol{S}^{1/2}\left(\boldsymbol{I}_p - \boldsymbol{R}\right)\boldsymbol{S}^{1/2}\boldsymbol{U}\right)\right]^l = \\ = \sum_{\kappa}\sum_{\lambda}\sum_{\phi\in\kappa\cdot\lambda} \theta_\phi^{\kappa,\lambda}\, C_\phi^{\kappa,\lambda}\left(\boldsymbol{S}^{1/2}\boldsymbol{U}\boldsymbol{S}^{1/2}\boldsymbol{R},\, \boldsymbol{S}^{1/2}\boldsymbol{U}\boldsymbol{S}^{1/2}\left(\boldsymbol{I}_p - \boldsymbol{R}\right)\right)$$

in (4.8), the joint p.d.f of $\boldsymbol{R}$ and $\boldsymbol{S}$ can be re-written as

$$\begin{aligned} &C(\alpha_1, \beta, k, \boldsymbol{\Sigma}, \boldsymbol{U})\, C(\alpha_2, \beta, l, \boldsymbol{\Sigma}, \boldsymbol{U}) \operatorname{etr}\left[-(\beta\boldsymbol{\Sigma})^{-1}\boldsymbol{S}\right] \det(\boldsymbol{S})^{\alpha_1+\alpha_2-(p+1)/2} \times \\ &\qquad \times \det(\boldsymbol{R})^{\alpha_1-(p+1)/2} \det(\boldsymbol{I}_p - \boldsymbol{R})^{\alpha_2-(p+1)/2} \\ &\qquad \times \sum_{\kappa}\sum_{\lambda}\sum_{\phi\in\kappa\cdot\lambda} \theta_\phi^{\kappa,\lambda}\, C_\phi^{\kappa,\lambda}\left(\boldsymbol{S}^{1/2}\boldsymbol{U}\boldsymbol{S}^{1/2}\boldsymbol{R},\, \boldsymbol{S}^{1/2}\boldsymbol{U}\boldsymbol{S}^{1/2}\left(\boldsymbol{I}_p - \boldsymbol{R}\right)\right), \end{aligned} \tag{4.9}$$

where $\boldsymbol{S} > \boldsymbol{0}$ and $\boldsymbol{0} < \boldsymbol{R} < \boldsymbol{I}_p$. Finally, integrating the above expression with respect to $\boldsymbol{R}$ by using (2.13), we get the p.d.f of $\boldsymbol{S}$. Further, substituting $\boldsymbol{U} = \boldsymbol{I}_p$ in the above expression and integrating $\boldsymbol{S}$ by using (2.8), we get the p.d.f of $\boldsymbol{R}$. □

5. FAMILY OF GENERALIZED MATRIX VARIATE *t*-DISTRIBUTIONS

In this section, a new family of matrix variate t distributions is introduced. This distribution will be useful in Bayesian analysis.

Definition 5.1. The $n \times p$ random matrix $\boldsymbol{T}$ is said to have a generalized matrix variate t distribution (GMT) with parameters $\boldsymbol{M} \in \mathbb{R}^{n\times p}$, $\boldsymbol{\Psi}(p\times p) > \boldsymbol{0}$, $\boldsymbol{\Omega}(n\times n) > \boldsymbol{0}$, $\boldsymbol{U}(p\times p) > \boldsymbol{0}$, $\alpha > (p-1)/2$, $\beta > 0$, $\kappa = (k_1, ..., k_p)$, $k_1 \geq \cdots \geq k_p \geq 0$, if its p.d.f is given by

$$\begin{aligned} &\frac{\det(\boldsymbol{\Omega})^{-p/2} \det(\boldsymbol{\Psi})^{-n/2}\, \Gamma_p(\alpha + n/2)}{\Gamma_p(\alpha) \sum_\kappa (\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Psi}^{-1})} \left(\frac{\beta}{2\pi}\right)^{np/2} \times \\ &\qquad \times \det\left(\boldsymbol{I}_n + \frac{\beta}{2}\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\,\boldsymbol{\Psi}^{-1}(\boldsymbol{T}-\boldsymbol{M})'\right)^{-(\alpha+n/2)} \\ &\qquad \times \sum_\kappa \left(\alpha + \frac{n}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Psi} + \frac{\beta}{2}(\boldsymbol{T}-\boldsymbol{M})'\,\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\right)^{-1}\right), \qquad \boldsymbol{T} \in \mathbb{R}^{n\times p}. \end{aligned} \tag{5.1}$$

We shall use the notation $\boldsymbol{T} \sim \mathrm{GMT}_{n,p}(\alpha, \beta, k, \boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Psi}, \boldsymbol{U})$.

For $\beta = 2$, $\alpha = (m + p - 1)/2$ and $k = 0$, the GMT distribution simplifies to the matrix variate t distribution (Gupta and Nagar [14]). Further, for $k = 0$, the GMT simplifies to the generalized matrix variate t distribution defined by Iranmanesh *et al.* [19].

Theorem 5.1. *Let* $\boldsymbol{X}|\boldsymbol{\Sigma} \sim N_{n,p}(\boldsymbol{0}, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$ *and* $\boldsymbol{\Sigma} \sim \mathrm{IGMG}_p(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$. *Then,* $\boldsymbol{X} \sim \mathrm{GMT}_{n,p}(\alpha, \beta, k, \boldsymbol{0}, \boldsymbol{\Omega}, \boldsymbol{\Psi}, \boldsymbol{U})$.

Proof: Let $g(\boldsymbol{X}|\boldsymbol{\Sigma})$ be the conditional p.d.f of $\boldsymbol{X}$ given $\boldsymbol{\Sigma}$. Further, let $h(\boldsymbol{\Sigma})$ be the marginal p.d.f of $\boldsymbol{\Sigma}$. Then, using conditional method, we find the marginal p.d.f of $\boldsymbol{X}$ as

$$f(\boldsymbol{X}) = \int_{\boldsymbol{\Sigma}>\boldsymbol{0}} g(\boldsymbol{X}|\boldsymbol{\Sigma})\, h(\boldsymbol{\Sigma})\, d\boldsymbol{\Sigma}.$$

Now, substituting for $g(\boldsymbol{X}|\boldsymbol{\Sigma})$ and $h(\boldsymbol{\Sigma})$ above, we get the marginal p.d.f of $\boldsymbol{X}$ as

$$\begin{aligned} f(\boldsymbol{X}) &= (2\pi)^{-np/2} \det(\boldsymbol{\Omega})^{-p/2} C\left(\alpha, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\right) \\ &\quad \times \int_{\boldsymbol{\Sigma}>\boldsymbol{0}} \operatorname{etr}\left[-\frac{1}{\beta}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)\boldsymbol{\Sigma}^{-1}\right] \det(\boldsymbol{\Sigma})^{-\alpha-(n+p+1)/2} \left[\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{U}\right)\right]^k d\boldsymbol{\Sigma}. \end{aligned}$$

Further, substituting $\boldsymbol{\Sigma}^{-1} = \boldsymbol{Z}$ with the Jacobian $J(\boldsymbol{\Sigma} \to \boldsymbol{Z}) = \det(\boldsymbol{Z})^{-(p+1)}$ in the above integral and using (3.2), we get

$$f(\boldsymbol{X}) = (2\pi)^{-np/2} \det(\boldsymbol{\Omega})^{-p/2} \frac{C\left(\alpha, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\right)}{C\left(\alpha + n/2, \beta, k, \left(\boldsymbol{\Psi} + \beta\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}/2\right)^{-1}, \boldsymbol{U}\right)}.$$

Finally, simplifying the above expression, we get

$$\begin{aligned} &\frac{\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Psi})^{-n/2}\Gamma_p(\alpha+n/2)}{\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa\left(\boldsymbol{U}\boldsymbol{\Psi}^{-1}\right)} \left(\frac{\beta}{2\pi}\right)^{np/2} \det\left(\boldsymbol{I}_n + \frac{\beta}{2}\boldsymbol{\Omega}^{-1}\boldsymbol{X}\boldsymbol{\Psi}^{-1}\boldsymbol{X}'\right)^{-(\alpha+n/2)} \times \\ &\qquad \times \sum_\kappa \left(\alpha + \frac{n}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}\right), \quad \boldsymbol{X} \in \mathbb{R}^{n\times p}, \end{aligned}$$

which is the desired result. □

Next, in Corollary 5.1.1, Corollary 5.1.2 and Theorem 5.2, we give three different variations of the above theorem.

Corollary 5.1.1. *Let* $\boldsymbol{Y}|\boldsymbol{\Sigma} \sim N_{p,n}(\boldsymbol{0}, \boldsymbol{\Sigma}\otimes\boldsymbol{\Omega})$ *and* $\boldsymbol{\Sigma} \sim \mathrm{IGMG}_p(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$*. Then, the marginal p.d.f of* $\boldsymbol{Y}$ *is given by*

$$\begin{aligned} &\frac{\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Psi})^{-n/2}\Gamma_p(\alpha+n/2)}{\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa\left(\boldsymbol{U}\boldsymbol{\Psi}^{-1}\right)} \left(\frac{\beta}{2\pi}\right)^{np/2} \det\left(\boldsymbol{I}_p + \frac{\beta}{2}\boldsymbol{\Psi}^{-1}\boldsymbol{Y}\boldsymbol{\Omega}^{-1}\boldsymbol{Y}'\right)^{-(\alpha+n/2)} \times \\ &\qquad \times \sum_\kappa \left(\alpha + \frac{n}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{Y}\boldsymbol{\Omega}^{-1}\boldsymbol{Y}'\right)^{-1}\right), \quad \boldsymbol{Y} \in \mathbb{R}^{p\times n}. \end{aligned}$$

Proof: Take $\boldsymbol{Y} = \boldsymbol{X}'$ in Theorem 5.1. □

Corollary 5.1.2. *Let* $\boldsymbol{X}|\boldsymbol{\Omega} \sim N_{n,p}(0, \boldsymbol{\Omega}\otimes\boldsymbol{\Sigma})$ *and* $\boldsymbol{\Omega} \sim \mathrm{IGMG}_n(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$*. Then, the marginal p.d.f of* $\boldsymbol{X}$ *is*

$$\begin{aligned} &\frac{\det(\boldsymbol{\Sigma})^{-n/2}\det(\boldsymbol{\Psi})^{-p/2}\Gamma_n(\alpha+p/2)}{\Gamma_n(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa\left(\boldsymbol{U}\boldsymbol{\Psi}^{-1}\right)} \left(\frac{\beta}{2\pi}\right)^{np/2} \det\left(\boldsymbol{I}_n + \frac{\beta}{2}\boldsymbol{\Psi}^{-1}\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}'\right)^{-(\alpha+p/2)} \times \\ &\qquad \times \sum_\kappa \left(\alpha + \frac{p}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}'\right)^{-1}\right), \quad \boldsymbol{X} \in \mathbb{R}^{n\times p}. \end{aligned}$$

Proof: This result can be obtained from Corollary 5.1.1. □

Theorem 5.2. *Let* $\boldsymbol{Y}|\boldsymbol{\Omega} \sim N_{p,n}(0, \boldsymbol{\Sigma} \otimes \boldsymbol{\Omega})$ *and* $\boldsymbol{\Omega} \sim \mathrm{IGMG}_n(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$*. Then,* $\boldsymbol{Y} \sim \mathrm{GMT}_{p,n}(\alpha, \beta, k, \boldsymbol{0}, \boldsymbol{\Sigma}, \boldsymbol{\Psi}, \boldsymbol{U})$*.*

Proof: Let $g(\boldsymbol{Y}|\boldsymbol{\Omega})$ be the conditional p.d.f of $\boldsymbol{Y}$ given $\boldsymbol{\Omega}$. Further, let $h(\boldsymbol{\Omega})$ be the marginal p.d.f of $\boldsymbol{\Omega}$. Then, using conditional method, we find the marginal p.d.f of $\boldsymbol{Y}$ as

$$f_{\boldsymbol{Y}}(\boldsymbol{Y}) = \int_{\boldsymbol{\Omega}>\boldsymbol{0}} g(\boldsymbol{Y}|\boldsymbol{\Omega}) h(\boldsymbol{\Omega})\, d\boldsymbol{\Omega}.$$

Now, substituting for $g(\boldsymbol{Y}|\boldsymbol{\Omega})$ and $h(\boldsymbol{\Omega})$ above, we get the marginal p.d.f of $\boldsymbol{Y}$ as

$$\begin{aligned} f_{\boldsymbol{Y}}(\boldsymbol{Y}) &= (2\pi)^{-np/2} \det(\boldsymbol{\Sigma})^{-n/2} C\left(\alpha, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\right) \\ &\quad \times \int_{\boldsymbol{\Omega}>\boldsymbol{0}} \operatorname{etr}\left[-\frac{1}{\beta}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}\right)\boldsymbol{\Omega}^{-1}\right] \det(\boldsymbol{\Omega})^{-\alpha-(n+p+1)/2} \left[\operatorname{tr}\left(\boldsymbol{\Omega}^{-1}\boldsymbol{U}\right)\right]^k\, d\boldsymbol{\Omega}. \end{aligned}$$

Further, substituting $\boldsymbol{\Omega}^{-1} = \boldsymbol{Z}$ with the Jacobian $J(\boldsymbol{\Omega} \to \boldsymbol{Z}) = \det(\boldsymbol{Z})^{-(p+1)}$ in the above integral and using (3.2), we get

$$f_{\boldsymbol{Y}}(\boldsymbol{Y}) = (2\pi)^{-np/2} \det(\boldsymbol{\Sigma})^{-n/2} \frac{C\left(\alpha, \beta, k, \boldsymbol{\Psi}^{-1}, \boldsymbol{U}\right)}{C\left(\alpha + p/2, \beta, k, \left(\boldsymbol{\Psi} + \beta\boldsymbol{Y}'\boldsymbol{\Sigma}^{-1}\boldsymbol{Y}/2\right)^{-1}, \boldsymbol{U}\right)}.$$

Finally, simplifying the above expression, we get the desired result. □

6. SOME PROPERTIES OF THE GMT FAMILY OF DISTRIBUTIONS

In this section, various properties of the GMT distribution are derived.

Proposition 6.1. *Let* $\boldsymbol{T} \sim \mathrm{GMT}_{n,p}(\alpha, \beta, k, \boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Psi}, \boldsymbol{U})$*. Let* $\boldsymbol{A}(n \times n)$ *and* $\boldsymbol{B}(p \times p)$ *be constant nonsingular matrices. Then,* $\boldsymbol{ATB} \sim \mathrm{GMT}_{n,p}(\alpha, \beta, k, \boldsymbol{AMB}, \boldsymbol{A\Omega A}', \boldsymbol{B}'\boldsymbol{\Psi B}, \boldsymbol{B}'\boldsymbol{UB})$*.*

Proof: Transforming $\boldsymbol{W} = \boldsymbol{ATB}$, with the Jacobian $J(\boldsymbol{T} \to \boldsymbol{W}) = \det(\boldsymbol{A})^{-p} \det(\boldsymbol{B})^{-n}$, in the p.d.f (5.1) of $\boldsymbol{T}$, and simplifying the resulting expression, we get the result. □

Corollary 6.0.1. *If* $\boldsymbol{T} \sim \mathrm{GMT}_{n,p}(\alpha, \beta, k, \boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Psi}, \boldsymbol{U})$*, then*

$$\boldsymbol{\Omega}^{-1/2}\boldsymbol{TB} \sim \mathrm{GMT}_{n,p}\left(\alpha, \beta, k, \boldsymbol{\Omega}^{-1/2}\boldsymbol{MB}, \boldsymbol{I}_n, \boldsymbol{B}'\boldsymbol{\Psi B}, \boldsymbol{B}'\boldsymbol{UB}\right),$$

$$\boldsymbol{AT\Psi}^{-1/2} \sim \mathrm{GMT}_{n,p}\left(\alpha, \beta, k, \boldsymbol{AM\Psi}^{-1/2}, \boldsymbol{A\Omega A}', \boldsymbol{I}_p, \boldsymbol{\Psi}^{-1/2}\boldsymbol{U\Psi}^{-1/2}\right)$$

and

$$\boldsymbol{\Omega}^{-1/2}\boldsymbol{T\Psi}^{-1/2} \sim \mathrm{GMT}_{n,p}\left(\alpha, \beta, k, \boldsymbol{\Omega}^{-1/2}\boldsymbol{M\Psi}^{-1/2}, \boldsymbol{I}_n, \boldsymbol{I}_p, \boldsymbol{\Psi}^{-1/2}\boldsymbol{U\Psi}^{-1/2}\right).$$

Proposition 6.2. *If $\boldsymbol{T} \sim \mathrm{GMT}_{n,p}(\alpha, \beta, k, \boldsymbol{M}, \boldsymbol{\Omega}, \boldsymbol{\Psi}, \boldsymbol{U})$, then for $n \geq p$, the p.d.f of $\boldsymbol{Z} = (\boldsymbol{T} - \boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T} - \boldsymbol{M})$ is given by*

$$\frac{\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Psi})^{-n/2}\Gamma_p(\alpha+n/2)}{\Gamma_p(n/2)\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Psi}^{-1})}\left(\frac{\beta}{2}\right)^{np/2}\det(\boldsymbol{Z})^{(n-p-1)/2}\times$$

$$\times\det\left(\boldsymbol{I}_p + \frac{\beta}{2}\boldsymbol{\Psi}^{-1}\boldsymbol{Z}\right)^{-(\alpha+n/2)}\sum_\kappa\left(\alpha+\frac{n}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Psi}+\frac{\beta}{2}\boldsymbol{Z}\right)^{-1}\right), \quad \boldsymbol{Z} > \boldsymbol{0}. \tag{6.1}$$

Proof: The p.d.f of $\boldsymbol{Z}$ is given by

$$\frac{\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Psi})^{-n/2}\Gamma_p(\alpha+n/2)}{\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Psi}^{-1})}\left(\frac{\beta}{2\pi}\right)^{np/2}\sum_\kappa\left(\alpha+\frac{n}{2}\right)_\kappa\times$$

$$\times\int_{(\boldsymbol{T}-\boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})=\boldsymbol{Z}}\det\left(\boldsymbol{I}_p+\frac{\beta}{2}\boldsymbol{\Psi}^{-1}(\boldsymbol{T}-\boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\right)^{-(\alpha+n/2)}$$

$$\times C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Psi}+\frac{\beta}{2}(\boldsymbol{T}-\boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\right)^{-1}\right)d\boldsymbol{Z}, \quad \boldsymbol{Z} > \boldsymbol{0}.$$

Now, evaluating the above integral by using Theorem 1.4.10 of Gupta and Nagar [14], we get the result. □

The following result is a generalization of the work of Dickey [12].

Theorem 6.1. *Let $\boldsymbol{X} \sim N_{n,p}(0, \boldsymbol{\Omega}\otimes\boldsymbol{I}_p)$, independent of $\boldsymbol{S} \sim \mathrm{GMG}_p\left(\alpha, \beta, k, \boldsymbol{\Lambda}^{-1}, \boldsymbol{U}\right)$. Define $\boldsymbol{T} = \boldsymbol{X}\boldsymbol{S}^{-1/2} + \boldsymbol{M}$, where $\boldsymbol{M}$ is an $n \times p$ constant matrix and $\boldsymbol{S}^{1/2}\left(\boldsymbol{S}^{1/2}\right)' = \boldsymbol{S}$. Then, the p.d.f of $\boldsymbol{T}$ is given by*

$$\frac{\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Lambda})^{-n/2}\Gamma_p(\alpha+n/2)}{\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Lambda}^{-1})}\left(\frac{\beta}{2\pi}\right)^{np/2}\times$$

$$\times\det\left(\boldsymbol{I}_p+\frac{\beta}{2}\boldsymbol{\Lambda}^{-1}(\boldsymbol{T}-\boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\right)^{-(\alpha+n/2)}$$

$$\times\sum_\kappa\left(\alpha+\frac{n}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Lambda}+\frac{\beta}{2}(\boldsymbol{T}-\boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\right)^{-1}\right), \quad \boldsymbol{T}\in\mathbb{R}^{n\times p}.$$

Proof: The joint p.d.f of $\boldsymbol{X}$ and $\boldsymbol{S}$ is given by

$$\frac{(2\pi)^{-np/2}\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Lambda})^{\alpha}}{\beta^{p\alpha+k}\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Lambda}^{-1})}\det(\boldsymbol{S})^{\alpha-(p+1)/2}\left[\mathrm{tr}(\boldsymbol{S}\boldsymbol{U})\right]^k\times$$

$$\times\exp\left[-\mathrm{tr}\left(\frac{1}{\beta}\boldsymbol{\Lambda}\boldsymbol{S}+\frac{1}{2}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)\right], \quad \boldsymbol{S}>\boldsymbol{0}, \quad \boldsymbol{X}\in\mathbb{R}^{n\times p}.$$

Now, let $\boldsymbol{T} = \boldsymbol{X}\boldsymbol{S}^{-1/2}+\boldsymbol{M}$. The Jacobian of this transformation is $J(\boldsymbol{X}\to\boldsymbol{T}) = \det(\boldsymbol{S})^{n/2}$. Substituting for $\boldsymbol{X}$ in terms of $\boldsymbol{T}$ in the joint p.d.f of $\boldsymbol{X}$ and $\boldsymbol{S}$, and multiplying the resulting expression by $J(\boldsymbol{X}\to\boldsymbol{T})$, we get the joint p.d.f of $\boldsymbol{T}$ and $\boldsymbol{S}$ as

$$\frac{(2\pi)^{-np/2}\det(\boldsymbol{\Omega})^{-p/2}\det(\boldsymbol{\Lambda})^{\alpha}}{\beta^{p\alpha+k}\Gamma_p(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa(\boldsymbol{U}\boldsymbol{\Lambda}^{-1})}\det(\boldsymbol{S})^{\alpha+n/2-(p+1)/2}\left[\mathrm{tr}(\boldsymbol{S}\boldsymbol{U})\right]^k\times$$

$$\times\,\mathrm{etr}\left[-\frac{1}{\beta}\left(\boldsymbol{\Lambda}+\frac{\beta}{2}(\boldsymbol{T}-\boldsymbol{M})'\boldsymbol{\Omega}^{-1}(\boldsymbol{T}-\boldsymbol{M})\right)\boldsymbol{S}\right], \quad \boldsymbol{S}>\boldsymbol{0}, \quad \boldsymbol{T}\in\mathbb{R}^{n\times p}.$$

Now, integrating out $\boldsymbol{S}$ by using (3.2) and simplifying the resulting expression the p.d.f of $\boldsymbol{T}$ is obtained. □

Theorem 6.2. *Let $\boldsymbol{X} \sim N_{n,p}\left(0, \boldsymbol{I}_n \otimes \boldsymbol{\Sigma}\right)$, independent of $\boldsymbol{S} \sim \mathrm{GMG}_n\left(\alpha, \beta, k, \boldsymbol{\Lambda}^{-1}, \boldsymbol{U}\right)$. Define $\boldsymbol{T} = \left(\boldsymbol{S}^{-1/2}\right)' \boldsymbol{X} + \boldsymbol{M}$, where $\boldsymbol{M}$ is an $n \times p$ constant matrix and $\boldsymbol{S}^{1/2}\left(\boldsymbol{S}^{1/2}\right)' = \boldsymbol{S}$. Then, the p.d.f of $\boldsymbol{T}$ is*

$$\frac{\det(\boldsymbol{\Sigma})^{-n/2}\det(\boldsymbol{\Lambda})^{-p/2}\Gamma_n(\alpha+p/2)}{\Gamma_n(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa\left(\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\right)}\left(\frac{\beta}{2\pi}\right)^{np/2}\times$$
$$\times\det\left(\boldsymbol{I}_n+\frac{\beta}{2}\boldsymbol{\Lambda}^{-1}(\boldsymbol{T}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{T}-\boldsymbol{M})'\right)^{-(\alpha+p/2)}$$
$$\times\sum_\kappa\left(\alpha+\frac{p}{2}\right)_\kappa C_\kappa\left(\boldsymbol{U}\left(\boldsymbol{\Lambda}+\frac{\beta}{2}(\boldsymbol{T}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{T}-\boldsymbol{M})'\right)^{-1}\right),\qquad \boldsymbol{T}\in\mathbb{R}^{n\times p}.$$

Proof: The joint p.d.f of $\boldsymbol{X}$ and $\boldsymbol{S}$ is given by

$$\frac{(2\pi)^{-np/2}\det(\boldsymbol{\Sigma})^{-n/2}\det(\boldsymbol{\Lambda})^{\alpha}}{\beta^{n\alpha+k}\Gamma_n(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa\left(\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\right)}\det(\boldsymbol{S})^{\alpha-(n+1)/2}\left[\mathrm{tr}(\boldsymbol{S}\boldsymbol{U})\right]^k\times$$
$$\times\exp\left[-\mathrm{tr}\left(\frac{1}{\beta}\boldsymbol{\Lambda}\boldsymbol{S}+\frac{1}{2}\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}'\right)\right],\quad \boldsymbol{S}>\boldsymbol{0},\quad \boldsymbol{X}\in\mathbb{R}^{n\times p}.$$

Now, let $\boldsymbol{T} = \left(\boldsymbol{S}^{-1/2}\right)'\boldsymbol{X}+\boldsymbol{M}$. The Jacobian of the transformation is $J(\boldsymbol{X}\to\boldsymbol{T}) = \det(\boldsymbol{S})^{p/2}$. Substituting for $\boldsymbol{X}$ in terms of $\boldsymbol{T}$ in the joint p.d.f of $\boldsymbol{X}$ and $\boldsymbol{S}$, and multiplying the resulting expression by $J(\boldsymbol{X}\to\boldsymbol{T})$, we get the joint p.d.f of $\boldsymbol{T}$ and $\boldsymbol{S}$ as

$$\frac{(2\pi)^{-np/2}\det(\boldsymbol{\Sigma})^{-n/2}\det(\boldsymbol{\Lambda})^{\alpha}}{\beta^{n\alpha+k}\Gamma_n(\alpha)\sum_\kappa(\alpha)_\kappa C_\kappa\left(\boldsymbol{U}\boldsymbol{\Lambda}^{-1}\right)}\det(\boldsymbol{S})^{\alpha+p/2-(n+1)/2}\left[\mathrm{tr}(\boldsymbol{S}\boldsymbol{U})\right]^k\times$$
$$\times\,\mathrm{etr}\left[-\left(\frac{1}{\beta}\boldsymbol{\Lambda}+\frac{1}{2}(\boldsymbol{T}-\boldsymbol{M})\boldsymbol{\Sigma}^{-1}(\boldsymbol{T}-\boldsymbol{M})'\right)\boldsymbol{S}\right],\quad \boldsymbol{S}>\boldsymbol{0},\quad \boldsymbol{X}\in\mathbb{R}^{n\times p}.$$

Now, integrating out $\boldsymbol{S}$ by using (3.2) and simplifying the resulting expression, the p.d.f of $\boldsymbol{T}$ is obtained. □

7. APPLICATIONS IN BAYESIAN ANALYSIS

As in Iranmanesh *et al.* [17], consider the Kullback-Leibler divergence loss (KLDL) function $\log\left(\frac{\pi(\boldsymbol{A}|\boldsymbol{D})}{\pi(\boldsymbol{\Sigma}|\boldsymbol{D})}\right)$ with the posterior expected loss function

$$\rho(\boldsymbol{\Sigma},\boldsymbol{A}) = E\left[\log\left(\frac{\pi(\boldsymbol{A}|\boldsymbol{D})}{\pi(\boldsymbol{\Sigma}|\boldsymbol{D})}\right)\right].$$

One may use the inverted generalized matrix gamma distribution as a prior distribution in Bayesian context. It is straightforward to prove that posterior distributions are IGMG. They are stated in Propositions 7.1 and 7.2 without proof.

Proposition 7.1. *Let $\boldsymbol{X}|\boldsymbol{\Sigma} \sim N_{n,p}(0, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$. Further suppose that the prior distribution of $\boldsymbol{\Sigma}$ is IGMG with parameters $(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$. Then, the posterior distribution of $\boldsymbol{\Sigma}$ is IGMG with parameters $\left(\alpha + n/2, \beta, k, \left(\boldsymbol{\Psi} + \beta \boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}/2\right)^{-1}, \boldsymbol{U}\right)$. That is, the posterior p.d.f of $\boldsymbol{\Sigma}$ is*

$$\pi(\boldsymbol{\Sigma}|\boldsymbol{X}) = C\left(\alpha + \frac{n}{2}, \beta, k, \left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)^{-1}, \boldsymbol{U}\right)$$
$$\times \operatorname{etr}\left[-\frac{1}{\beta}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X}\right)\boldsymbol{\Sigma}^{-1}\right] \det(\boldsymbol{\Sigma})^{-\alpha-(n+p+1)/2}\left[\operatorname{tr}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{U}\right)\right]^k, \quad \boldsymbol{\Sigma} > \boldsymbol{0}.$$

Proposition 7.2. *Let $\boldsymbol{X}|\boldsymbol{\Sigma} \sim N_{n,p}(0, \boldsymbol{\Omega} \otimes \boldsymbol{\Sigma})$. Further suppose that the prior distribution of $\boldsymbol{\Omega}$ is IGMG with parameters $(\alpha, \beta, k, \boldsymbol{\Psi}, \boldsymbol{U})$. Then, the posterior distribution of $\boldsymbol{\Omega}$ is IGMG with parameters $\left(\alpha + p/2, \beta, k, \left(\boldsymbol{\Psi} + \beta \boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}'/2\right)^{-1}, \boldsymbol{U}\right)$. That is, the posterior p.d.f of $\boldsymbol{\Omega}$ is*

$$\pi(\boldsymbol{\Omega}|\boldsymbol{X}) = C\left(\alpha + \frac{p}{2}, \beta, k, \left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}'\right)^{-1}, \boldsymbol{U}\right)$$
$$\times \operatorname{etr}\left[-\frac{1}{\beta}\left(\boldsymbol{\Psi} + \frac{\beta}{2}\boldsymbol{X}\boldsymbol{\Sigma}^{-1}\boldsymbol{X}'\right)\boldsymbol{\Omega}^{-1}\right] \det(\boldsymbol{\Omega})^{-\alpha-(n+p+1)/2}\left[\operatorname{tr}\left(\boldsymbol{\Omega}^{-1}\boldsymbol{U}\right)\right]^k, \quad \boldsymbol{\Omega} > \boldsymbol{0}.$$

By definition, the Bayes estimator of $\boldsymbol{\Sigma}$, under the KLDL function, is given by $\widehat{\boldsymbol{\Sigma}} = \arg\max_{\boldsymbol{\Sigma}} \pi(\boldsymbol{\Sigma}|\boldsymbol{X})$. Iranmanesh *et al.* [19] have shown that

$$\widehat{\boldsymbol{\Sigma}} = [\alpha + n/2 + (p+1)/2]^{-1}\left(\frac{1}{2}\boldsymbol{X}'\boldsymbol{\Omega}^{-1}\boldsymbol{X} + \frac{1}{\beta}\boldsymbol{\Psi}\right)$$

for the special case $k = 0$.

8. CONCLUSION

In this paper, a generalized matrix variate gamma distribution has been introduced. The corresponding inverted matrix variate gamma distribution has also been derived. By making use of this newly defined matrix variate distribution as the prior for the characteristic matrix of a matrix variate normal distribution, using conditioning approach, a family of generalized matrix variate t distributions has also been defined.

A future work is to consider estimation of the newly introduced matrix variate distributions. One issue is that the new distributions are over parameterized; that is, there is parameter redundancy. This can be accounted for numerically by constrained maximization of the log likelihood. For example, if the data follow the overparameterized p.d.f $ab\exp(-abx)$ then the log likelihood can be maximized using the constraint $ab = c$. Usually, partial derivatives of the log likelihood are not required for evaluating maximum likelihood estimates numerically.

ACKNOWLEDGMENTS

The authors would like to thank the Editor and the referee for careful reading and comments which greatly improved the paper.

REFERENCES

[1] Adhikari, S. (2007). *Characterization of uncertainty in damping modeling*. In "Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference", Waikiki, Hawaii.

[2] Arashi, M.; Bekker, A. and van Niekerk, J. (2017). Weighted-type Wishart distributions with application, *REVSTAT – Statistical Journal*, **15**, 205–222.

[3] Bates, A.; Daducci, A. and Caruyer, E. (2019). *Multi-dimensional diffusion MRI sampling scheme: B-tensor design and accurate signal reconstruction*. In "Proceedings of ISMRM 2019 – 27th Annual Meeting Exhibition", Montreal, Canada, pp. 1–4.

[4] Chamberland, M.; Scherrer, B.; Prabhu, S.P.; Madsen, J.; Fortin, D.; Whittingstall, K.; Descoteaux, M. and Warfield, S.K. (2017). Active delineation of Meyer's loop using oriented priors through MAGNEtic tractography (MAGNET), *Human Brain Mapping*, **38**, 509–527.

[5] Chikuse, Y. (1981). Distributions of some matrix variates and latent roots in multivariate Behrens-Fisher discriminant analysis, *Annals of Statistics*, **9**, 401–407.

[6] Das, S. and Dey, D.K. (2010). On Bayesian inference for generalized multivariate gamma distribution, *Statistics and Probability Letters*, **80**, 1492–1499.

[7] Das, S. and Ghanem, R. (2009). A bounded random matrix approach for stochastic upscaling, *Multiscale Modeling and Simulation*, **8**, 296–325.

[8] Das, S. and Ghanem, R. (2011). *Stochastic upscaling for inelastic material behavior from limited experimental data*. In "Computational Methods for Microstructure-Property Relationships" (S. Ghosh and D. Dimiduk, Eds.), Springer, Boston, MA, pp. 443–468.

[9] Davis, A.W. (1979). Invariant polynomials with two matrix arguments extending the zonal polynomials: applications to multivariate distribution theory, *Annals of the Institute of Statistical Mathematics*, **31**(4), 465–485.

[10] Davis, A.W. (1980). *Invariant polynomials with two matrix arguments extending the zonal polynomials*. In "Multivariate Analysis – V" (P.R. Krishnaiah, Ed.), pp. 287–299.

[11] Díaz-García, J.A. (2011). Generalisations of some properties of invariant polynomials with matrix arguments, *Applicationes Mathematicae* (Warsaw), **38**, 469–475.

[12] Dickey, J.M. (1967). Matrix variate generalizations of the multivariate t and the inverted multivariate t distribution, *Annals of Mathematical Statistics*, **38**, 511–518.

[13] Ferreira, J.T.; Bekker, A. and Arashi, M. (2020). Advances in Wishart-type modeling of channel capacity, *REVSTAT – Statistical Journal*, **18**, 237–255.

[14] Gupta, A.K. and Nagar, D.K. (2000). *Matrix Variate Distributions*, Chapman and Hall/CRC, Boca Raton.

[15] Haubold, H.J. and Kumar, D. (2011). Fusion yield: Guderley model and Tsallis statistics, *Journal of Plasma Physics*, **77**, 1–14.

[16] Herz, C.S. (1955). Bessel functions of matrix argument, *Annals of Mathematics*, **61**, 474–523.

[17] Iranmanesh, A.; Arashi, M.; Nagar, D.K. and Tabatabaey, S.M.M. (2013). On inverted matrix variate gamma distribution, *Communications in Statistics – Theory and Methods*, **42**, 28–41.

[18] Iranmanesh, A.; Arashi, M. and Shokri, S. (2014). Generalized matrix t-distribution through generalized multivariate gamma distribution, *ATINER's Conference Paper Series*, No: STA 2014-1114.

[19] Iranmanesh, A.; Arashi, M. and Tabatabaey, S.M.M. (2010). On conditional applications of matrix variate normal distribution, *Iranian Journal of Mathematical Sciences and Informatics*, **5**, 33–43.

[20] James, A.T. (1964). Distributions of matrix variate and latent roots derived from normal samples, *Annals of Mathematical Statistics*, **35**, 475–501.

[21] Javier, W.R. and Gupta, A.K. (1985). On matric variate-t distribution, *Communications in Statistics – Theory and Methods*, **14**, 1413–1425.

[22] Jian, B.; Vemuri, B.C.; Ozarslan, E.; Carney, P. and Mareci, T. (2007). *A continuous mixture of tensors model for diffusion-weighted MR signal reconstruction.* In "Proceedings of the 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro", Arlington, VA, pp. 772–775.

[23] Khatri, C.G. (1966). On certain distribution problems based on positive definite quadratic functions in normal vectors, *Annals of Mathematical Statistics*, **37**, 468–479.

[24] Luo, L.; Yang, J.; Zhang, B.; Jiang, J. and Huang, H. (2018). Nonparametric Bayesian correlated group regression with applications to image classification, *IEEE Transactions on Neural Networks and Learning Systems*, **29**, 5330–5344.

[25] Marx, D.G. (1983). Quadratic forms of a matric-t variate, *Annals of the Institute of Statistical Mathematics*, **35**, 347–353.

[26] Mathai, A.M. (1997). *Jacobians of Matrix Transformations and Functions of Matrix Argument*, World Scientific, London.

[27] Muirhead, R.J. (1982). *Aspects of Multivariate Statistical Theory*, John Wiley and Sons, New York.

[28] Nagar, D.K. and Gupta, A.K. (2002). Matrix-variate Kummer-beta distribution, *Journal of the Australian Mathematical Society*, **73**, 11–25.

[29] Nagar, D.K.; Morán-Vásquez, R.A. and Gupta, A.K. (2015). Extended matrix variate hypergeometric functions and matrix variate distribution, *International Journal of Mathematics and Mathematical Science*, **2015**, Art. ID 190723, 15 pages.

[30] Nagar, D.K.; Roldan-Correa, A. and Gupta, A.K. (2013). Extended matrix variate gamma and beta functions, *Journal of Multivariate Analysis*, **122**, 53–69.

[31] Pascual, B. and Adhikari, S. (2012). Combined parametric–nonparametric uncertainty quantification using random matrix theory and polynomial chaos expansion, *Computers and Structures*, **112–113**, 364–379.

[32] Scherrer, B.; Schwartzman, A.; Taquet, M.; Prabhu, S.P.; Sahin, M.; Akhondi-Asl, A. and Warfield, S.K. (2013). *Characterizing the DIstribution of Anisotropic MicrOstructural eNvironments with Diffusion-weighted imaging (DIAMOND).* In "Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention" (MICCAI)(8151), Nagoya, Japan (K. Mori, I. Sakuma, Y. Sato, C. Barillot and N. Navab, Eds.), part III, pp. 518–526.

[33] Scherrer, B.; Taquet, M.; Schwartzman, A.; St-Onge, E.; Rensonnet, G.; Prabhu, S.P. and Warfield, S.K. (2017). *Decoupling axial and radial tissue heterogeneity in diffusion compartment imaging.* In "Information Processing in Medical Imaging", Lecture Notes in Computer Science (M. Niethammer *et al.*, Eds.), Springer, Cham, **10265**, pp. 440–452.

[34] Takemura, A. (1984). *Zonal Polynomials*, Institute of Mathematical Statistics Lecture Notes – Monograph Series, **4**, Institute of Mathematical Statistics, Hayward, CA.

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 51–83

INFERENCE ON STRESS-STRENGTH MODEL FOR A KUMARASWAMY DISTRIBUTION BASED ON HYBRID PROGRESSIVE CENSORED SAMPLE

Author: Akram Kohansal
– Department of Statistics, Imam Khomeini International University, Qazvin, Iran
kohansal@sci.ikiu.ac.ir

Received: March 2019 Revised: September 2019 Accepted: November 2019

Abstract:

- In this paper, we obtain the point and interval estimates of the stress-strength parameter under the hybrid progressive censored scheme, when stress and strength are considered as two independent random variables of Kumaraswamy. We solve the problem in three cases, as follows: First, assuming that stress and strength have different first shape parameters and the common second shape parameter, we obtain maximum likelihood estimation (MLE), approximation maximum likelihood estimation (AMLE) and two Bayesian approximation estimates due to the lack of explicit forms. Also, we construct the asymptotic and highest posterior density (HPD) intervals for R. Moreover, we consider the existence and uniqueness of the MLE. Second, assuming that common second shape parameter is identified, we derive the MLE and exact Bayes estimate of R. Third, assuming that all parameters are unknown and different, we achieve the statistical inference of R, namely MLE, AMLE and Bayesian inference of R. Furthermore, we apply the Monte Carlo simulations for comparing the performance of different methods. Finally, we analyze two data sets for illustrative purposes.

Keywords:

- *stress-strength model; hybrid progressive censored sample; Kumaraswamy distribution; Bayesian inference; Monte Carlo simulation.*

AMS Subject Classification:

- 62F10, 62F15, 62N05.

1. INTRODUCTION

One of the most interesting problems in reliability theory, is inference of the stress-strength parameter, $R = P(X < Y)$. The variables Y and X are known as strength and stress, respectively. In one system, if the applied stress is greater than its strength, as a result the system fails. In statistical science, more attention has been paid to the estimation of R since 1956, beginning with the work of Birnbaum [3]. From that time, estimating the R have been done from the frequentist and Bayesian viewpoints. Recently, some studies about the stress-strength model can be found in Rezaei *et al.* [21], Babayi *et al.* [2], Nadar *et al.* [18] and Kizilaslan and Nadar [8].

Although, in the complete sample case, many authors have been investigated the stress-strength models, they did not pay attention to the censored sample case. Whereas in really applicable situations, for many reasons like financial plane or limited time, the researchers confront censored data.

Among various censoring schemes, Type-I and Type-II are the two most fundamental schemes, which can be explained as follows. We finish the test in a pre-selected time and pre-chosen number of failures, in Type-I and Type-II schemes, respectively. Also, we finish the test at time $T^* = \min\{X_{m:n}, T\}$, where $X_{m:n}$ is the m-th failure times from n items and $T > 0$, in the hybrid scheme, which has been indicated by Epstein [5]. Also, In hybrid scheme, Singh and Goel [24] obtained reliability estimation of modified Weibull distribution. Because in the hybrid scheme, the removal of active units cannot be lost during the test, hybrid progressive (HP) scheme is introduced by Kundu and Joarder [14], which can be described as follows. Let N units be put on the test with censoring scheme $(R_1, ..., R_n)$ and pausing time $T^* = \min\{X_{n:n:N}, T\}$, where $X_{1:n:N} \leq ... \leq X_{n:n:N}$ be a progressive censoring scheme and $T > 0$ is a fixed time. It is obvious that if $X_{n:n:N} < T$ then we finish the test at time $X_{n:n:N}$ and $\{X_{1:n:N}, ..., X_{n:n:N}\}$ is the observed sample. Otherwise, if $X_{J:n:N} < T < X_{J+1:n:N}$ then we finish the test at time T and $\{X_{1:n:N}, ..., X_{J:n:N}\}$ is the observed sample. In symbol, we say that $\{X_{1:n:N}, ..., X_{J:n:N}\}$ is a HP censoring sample with scheme $\{N, n, T, J, R_1, ..., R_J\}$. Recently, some of the authors have studied the stress-strength model and censored data. For example, Shoaee and Khorram considered stress-strength reliability of a two-parameter bathtub-shaped lifetime distribution with respect to progressively censored samples, [22]. Also, they obtained some statistical inference of $R = P(Y < X)$ for Weibull distribution under Type-II progressively hybrid censored data, [23]. Kohansal [9] considered estimation of multicomponent stress-strength reliability for Kumaraswamy distribution under progressive censoring. Rasethuntsa and Nadar [20] studied stress-strength reliability of a non-identical-component-strengths system based on upper record values from the family of Kumaraswamy generalized distributions. Very recently, Maurya and Tripathi [17] derived the reliability estimation in a multicomponent stress strength model for Burr XII distribution under progressive censoring. In addition, Kohansal [10] obtained Bayesian and classical estimation of $R = P(X < Y)$ based on Burr type XII distribution under hybrid progressive censored samples. Kohansal and Rezakhah [12] considered the inference of $R = P(Y < X)$ for two-parameter Rayleigh distribution in terms of progressively censored samples. Ahmadi and Ghafouri [1] obtained the reliability estimation in a multicomponent stress-strength model under generalized half-normal distribution based on progressive Type-II censoring. Furthermore, Kohansal and Shoaee [13] derived Bayesian and classical estimation of reliability in a multicomponent stress-strength model under adaptive hybrid progressive censored data.

Finally, Kohansal and Nadarajah [11] estimated the stress-strength parameter based on Type-II hybrid progressive censored samples for a Kumaraswamy distribution. In this study, based on HP censoring scheme, the reliability parameter $R = P(X < Y)$ is estimated when X and Y are two independent random variables from the Kumaraswamy distribution (KuD). This paper has also some contribution in terms of inference. We consider the different point and interval estimations of R, and all of these estimates are considered in Bayesian and classical viewpoints. Also, we investigate the problem in three different cases, first at the time that X and Y have the unknown common one parameter, secondly when have known common one parameter, and third when they have different unknown parameters. Moreover, as the HP censoring is a general scheme, so we can obtain from it, some cases that are considered (up to now).

KuD with the first and second shape parameters α and λ, respectively, which is denoted by Ku(α, λ), has the probability density function (pdf), cumulative distribution function (cdf) and failure rate function as follows:

$$
\begin{aligned}
f(x) &= \alpha\lambda x^{\lambda-1}(1-x^{\lambda})^{\alpha-1}, && 0 < x < 1, \;\; \alpha, \lambda > 0, \\
F(x) &= 1-(1-x^{\lambda})^{\alpha}, && 0 < x < 1, \;\; \alpha, \lambda > 0, \\
H(x) &= \frac{\alpha\lambda x^{\lambda-1}}{1-x^{\lambda}}, && 0 < x < 1, \;\; \alpha, \lambda > 0,
\end{aligned}
$$

respectively. The probability density and failure rate functions of KuD are presented in Figure 1. KuD has an increasing failure rate function, so the KuD can be used for analyzing the real data sets if the empirical consideration suggests that the failure rate function of the prior distribution is increasing. Moreover, KuD is the very appropriate fit to many natural phenomena, which their outcomes have lower and upper bounds, such as the heights of individuals, scores obtained on a test, atmospheric temperatures, hydrological data, economic data, etc.

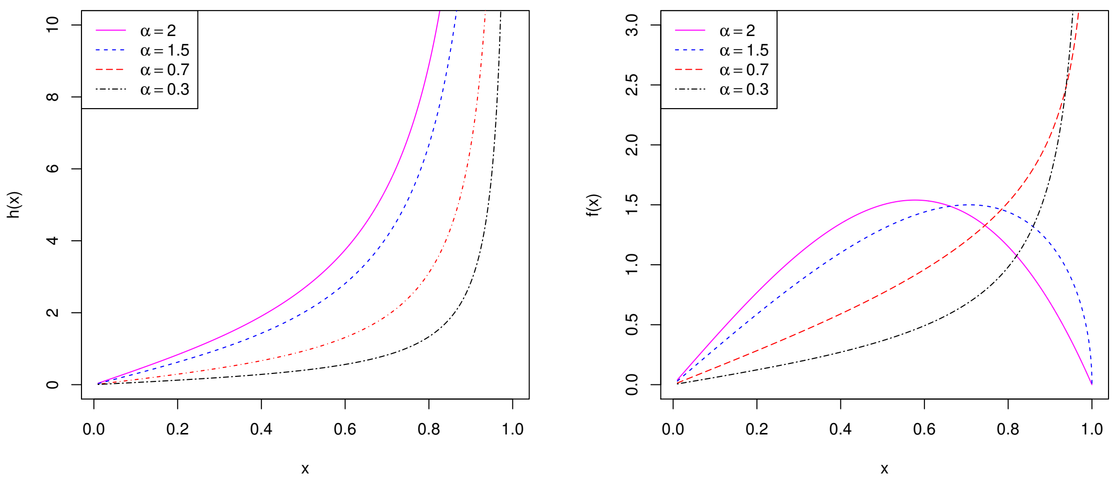

Figure 1: Shape of probability density (right) and failure rate (left) functions of KuD when $\lambda = 2$.

The other parts of this paper are arranged as follows: In Section 2, under the HP censoring scheme, assuming $X \sim$ Ku(α, λ) and $Y \sim$ Ku(β, λ), we obtain the point and interval estimates of $R = P(X < Y)$, from the frequentist and Bayesian viewpoints.

More specifically, in Section 2, the existence and uniqueness of MLEs are considered. Because the MLEs of unknown parameters and R cannot be earned in the closed forms, we obtain the AMLEs of parameters and R, which have the explicit forms. In addition, we develop the Bayes estimates of R, by applying Lindley's approximation and MCMC method due to the lack of explicit forms. Moreover, different confidence intervals such as asymptotic and HPD intervals of R are provided. In Section 3, by assuming that the common shape parameter is known, the MLE and exact Bayes estimate of R are earned. Because the assumption studied in Section 2 is quite strong, we consider the statistical inference of R in general case. Accordingly, in Section 4, under the HP censoring scheme, assuming $X \sim \mathrm{Ku}(\alpha, \lambda_1)$ and $Y \sim \mathrm{Ku}(\beta, \lambda_2)$, we provide the MLE, AMLE and Bayes estimate of R, respectively. In Section 5, we give the simulation results and data analysis, and following that we conclude the paper in Section 6.

2. INFERENCE ON R WITH UNKNOWN COMMON λ

2.1. MLE of R

The stress-strength parameter, when X and Y are two independent random variables from $\mathrm{Ku}(\alpha, \lambda)$ and $\mathrm{Ku}(\beta, \lambda)$, respectively, can be obtained simply as $R = P(X < Y) = \frac{\alpha}{\alpha+\beta}$. In this section, under the HP censoring scheme, we derive the MLE of R. Because R is a function of the unknown parameters, consequently at first we obtain the MLEs of α, β, and λ. If $\{X_1, ..., X_n\}$ and $\{Y_1, ..., Y_m\}$ be two HP censoring samples with censoring schemes $\{N, n, T_1, J_1, R_1, ..., R_{J_1}\}$ and $\{M, m, T_2, J_2, S_1, ..., S_{J_2}\}$, respectively, after that the likelihood function of the unknown parameters α, β and λ can be written as

$$\begin{aligned} L(\alpha, \beta, \lambda) \propto & \left[\prod_{i=1}^{J_1} f(x_i)[1 - F(x_i)]^{R_i}[1 - F(T_1)]^{R^*_{J_1}}\right] \\ & \times \left[\prod_{i=1}^{J_2} f(y_j)[1 - F(y_j)]^{S_j}[1 - F(T_2)]^{S^*_{J_2}}\right], \end{aligned}$$

where

$$R^*_{J_1} = N - J_1 - \sum_{i=1}^{J_1} R_i, \; S^*_{J_2} = M - J_2 - \sum_{j=1}^{J_2} S_j.$$

The proposed model, in association with the existing ones, has some differences and similarities. About the differences, we notice that it is a general model and some important models can be obtained from it. For example, by setting $T_1 = X_n$ and $T_2 = Y_m$, we derive the likelihood function for $R = P(X < Y)$ in the progressive censoring scheme. Also, by setting $T_1 = X_n, \; R_i = 0 \; (i = 1, ..., n-1), \; R_n = N - n$ and $T_2 = Y_m, \; S_j = 0 \; (j = 1, ..., m-1), \; S_m = M - m$, we obtain the likelihood function for $R = P(X < Y)$ in Type-II censoring scheme. Moreover, by setting $T_1 = X_n, \; R_i = 0 \; (i = 1, ..., n)$ and $T_2 = Y_m, \; S_j = 0 \; (j = 1, ..., m)$, we derive the likelihood function for $R = P(X < Y)$ in complete sample. About the similarities, we identify that most of the censoring schemes have complex computational needs.

The likelihood function, with respect to the observed data can be obtained as:

$$\begin{aligned} L(\text{data}|\alpha,\beta,\lambda) &\propto \alpha^{J_1}\beta^{J_2}\lambda^{J_1+J_2}\left(\prod_{i=1}^{J_1} x_i^{\lambda-1}(1-x_i^{\lambda})^{\alpha(R_i+1)-1}\right)(1-T_1^{\lambda})^{\alpha R^*_{J_1}} \\ &\times \left(\prod_{j=1}^{J_2} y_j^{\lambda-1}(1-y_j^{\lambda})^{\beta(S_j+1)-1}\right)(1-T_2^{\lambda})^{\beta S^*_{J_2}}. \end{aligned}$$

Therefore, the log-likelihood function, along with ignoring the constant value, is as:

$$\begin{aligned} \ell(\alpha,\beta,\lambda) &= J_1\log(\alpha)+\sum_{i=1}^{J_1}(\alpha(R_i+1)-1)\log(1-x_i^{\lambda})+\alpha R^*_{J_1}\log(1-T_1^{\lambda}) \\ &+ J_2\log(\beta)+\sum_{j=1}^{J_2}(\beta(S_j+1)-1)\log(1-y_j^{\lambda})+\beta S^*_{J_2}\log(1-T_2^{\lambda}) \\ &+ (\lambda-1)\sum_{i=1}^{J_1}\log(x_i)+(\lambda-1)\sum_{j=1}^{J_2}\log(y_j)+(J_1+J_2)\log(\lambda). \end{aligned} \tag{2.1}$$

Consequently, to earn the MLEs of α, β and λ, namely, $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\lambda}$, respectively, we should solve the following equations:

$$\frac{\partial\ell}{\partial\alpha} = \frac{J_1}{\alpha}+\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda})+R^*_{J_1}\log(1-T_1^{\lambda})=0, \tag{2.2}$$

$$\frac{\partial\ell}{\partial\beta} = \frac{J_2}{\beta}+\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda})+S^*_{J_2}\log(1-T_2^{\lambda})=0, \tag{2.3}$$

$$\begin{aligned} \frac{\partial\ell}{\partial\lambda} &= \frac{J_1+J_2}{\lambda}+\sum_{i=1}^{J_1}\log(x_i)-\sum_{i=1}^{J_1}\Big(\alpha(R_i+1)-1\Big)x_i^{\lambda}\frac{\log(x_i)}{1-x_i^{\lambda}}-\alpha R^*_{J_1}T_1^{\lambda}\frac{\log(T_1)}{1-T_1^{\lambda}} \\ &+\sum_{j=1}^{J_2}\log(y_j)-\sum_{j=1}^{J_2}\Big(\beta(S_j+1)-1\Big)y_j^{\lambda}\frac{\log(y_j)}{1-y_j^{\lambda}}-\beta S^*_{J_2}T_2^{\lambda}\frac{\log(T_2)}{1-T_2^{\lambda}}=0. \end{aligned} \tag{2.4}$$

From the equations (2.2) and (2.3), we have

$$\begin{aligned} \widehat{\alpha}(\lambda) &= -J_1\left\{\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda})+R^*_{J_1}\log(1-T_1^{\lambda})\right\}^{-1}, \\ \widehat{\beta}(\lambda) &= -J_2\left\{\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda})+S^*_{J_2}\log(1-T_2^{\lambda})\right\}^{-1}. \end{aligned}$$

Also, to derive $\widehat{\lambda}$, we apply one numerical method like Newton–Raphson on the equation (2.4). After obtaining the MLEs of α, β, and λ, by the use of the invariance property, the MLE of R can be derived as

$$\widehat{R}^{\text{MLE}} = \frac{\widehat{\alpha}}{\widehat{\alpha}+\widehat{\beta}}. \tag{2.5}$$

2.2. Existence and uniqueness of the MLEs

In this section, we consider the existence and uniqueness of the MLEs.

Theorem 2.1. *The MLEs of the parameters α and β, which were obtained by applying the following equations, are unique:*

$$\widehat{\alpha} = -J_1\left\{\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda}) + R_{J_1}^{*}\log(1-T_1^{\lambda})\right\}^{-1},$$
$$\widehat{\beta} = -J_2\left\{\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda}) + S_{J_2}^{*}\log(1-T_2^{\lambda})\right\}^{-1},$$

and $\widehat{\lambda}$ should be obtained by finding a solution for the following equation:

$$\begin{aligned} G(\lambda) \;=\; \frac{J_1+J_2}{\lambda} &+ \sum_{i=1}^{J_1}\log(x_i) - \sum_{i=1}^{J_1}\Big(\widehat{\alpha}(R_i+1)-1\Big)x_i^{\lambda}\frac{\log(x_i)}{1-x_i^{\lambda}} - \widehat{\alpha}R_{J_1}^{*}T_1^{\lambda}\frac{\log(T_1)}{1-T_1^{\lambda}} \\ &+ \sum_{j=1}^{J_2}\log(y_j) - \sum_{j=1}^{J_2}\Big(\widehat{\beta}(S_j+1)-1\Big)y_j^{\lambda}\frac{\log(y_j)}{1-y_j^{\lambda}} - \widehat{\beta}S_{J_2}^{*}T_2^{\lambda}\frac{\log(T_2)}{1-T_2^{\lambda}}. \end{aligned}$$

Proof: See Appendix A. □

2.3. AMLE of R

From Section 2.1, we observe that the MLEs of unknown parameters and R cannot be earned in the closed forms. As a result in this section, we obtain the AMLEs of the parameters, which have the explicit forms.

Lemma 2.1. *Let Z' and Z'' be Weibull and Extreme value distributions, in symbols $Z' \sim \mathrm{W}(\alpha,\theta)$ and $Z'' \sim \mathrm{EV}(\mu,\sigma)$, if they have the following cumulative distribution functions, respectively as:*

$$F_{Z'}(z) = 1 - e^{-\frac{x^{\alpha}}{\theta}}, \qquad z>0,\quad \alpha,\theta>0,$$
$$F_{Z''}(z) = 1 - e^{-e^{\frac{x-\mu}{\sigma}}}, \qquad z\in\mathbb{R},\quad \mu\in\mathbb{R},\sigma>0.$$

- **(i)** *If $Z \sim \mathrm{Ku}(\alpha,\lambda)$ and $Z' = (-\log(1-Z^{\lambda}))^{\frac{1}{\lambda}}$, then $Z' \sim \mathrm{W}(\lambda,\frac{1}{\alpha})$.*
- **(ii)** *If $Z' \sim \mathrm{W}(\lambda,\frac{1}{\alpha})$ and $Z'' = \log(Z')$, then $Z'' \sim \mathrm{EV}(\mu,\sigma)$, where $\mu = -\frac{1}{\lambda}\log(\alpha)$ and $\sigma = \frac{1}{\lambda}$.*

Proof: Obvious. □

Consider that $\{X_1, ..., X_n\}$ and $\{Y_1, ..., Y_m\}$ be two HP censoring samples with censoring schemes $\{N, n, T_1, J_1, R_1, ..., R_{J_1}\}$ and $\{M, m, T_2, J_2, S_1, ..., S_{J_2}\}$, respectively and

$$X_i' = (-\log(1 - X_i^\lambda))^{\frac{1}{\lambda}},\ U_i = \log(X_i') \ \text{ and } \ Y_j' = (-\log(1 - Y_j^\lambda))^{\frac{1}{\lambda}},\ V_j = \log(Y_j').$$

Applying Lemma 2.1, $U_i \sim \mathrm{EV}(\mu_1, \sigma)$ and $V_j \sim \mathrm{EV}(\mu_2, \sigma)$, where

$$\mu_1 = -\frac{1}{\lambda}\log(\alpha),\ \mu_2 = -\frac{1}{\lambda}\log(\beta),\ \text{and } \sigma = \frac{1}{\lambda}.$$

Therefore, in terms of the observed data $\{U_1, ..., U_n\}$ and $\{V_1, ..., V_m\}$, and by ignoring the constant value, the log-likelihood function is as follows:

$$\begin{aligned} \ell^*(\mu_1, \mu_2, \sigma) &= \sum_{i=1}^{J_1} t_i - \sum_{i=1}^{J_1}(R_i + 1)e^{t_i} - R_{J_1}^* e^{\delta_1} \\ &\quad + \sum_{j=1}^{J_2} z_j - \sum_{j=1}^{J_2}(S_j + 1)e^{z_j} - S_{J_2}^* e^{\delta_2} - (J_1 + J_2)\log(\sigma), \end{aligned} \tag{2.6}$$

where

$$\begin{aligned} t_i &= \frac{u_i - \mu_1}{\sigma},\ z_j = \frac{v_j - \mu_2}{\sigma},\ \delta_1 = \frac{a_1 - \mu_1}{\sigma},\ \delta_2 = \frac{a_2 - \mu_2}{\sigma}, \\ a_1 &= \log\left((-\log(1 - T_1^\lambda))^{\frac{1}{\lambda}}\right),\ a_2 = \log\left((-\log(1 - T_2^\lambda))^{\frac{1}{\lambda}}\right). \end{aligned}$$

Now by taking derivatives with respect to μ_1, μ_2 and σ from (2.6), we achieve the following equations:

$$\frac{\partial \ell^*}{\partial \mu_1} = -\frac{1}{\sigma}\left[J_1 - \sum_{i=1}^{J_1}(R_i + 1)e^{t_i} - R_{J_1}^* e^{\delta_1}\right] = 0, \tag{2.7}$$

$$\frac{\partial \ell^*}{\partial \mu_2} = -\frac{1}{\sigma}\left[J_2 - \sum_{j=1}^{J_2}(S_j + 1)e^{z_j} - S_{J_2}^* e^{\delta_2}\right] = 0, \tag{2.8}$$

$$\begin{aligned} \frac{\partial \ell^*}{\partial \sigma} &= -\frac{1}{\sigma}\Bigg[J_1 + J_2 + \sum_{i=1}^{J_1} t_i - \sum_{i=1}^{J_1}(R_i + 1)t_i e^{t_i} - R_{J_1}^* \delta_1 e^{\delta_1} \\ &\quad + \sum_{j=1}^{J_2} z_j - \sum_{j=1}^{J_2}(S_j + 1)z_j e^{z_j} - S_{J_2}^* \delta_2 e^{\delta_2}\Bigg] = 0. \end{aligned} \tag{2.9}$$

To obtain the AMLEs of μ_1, μ_2 and σ, let

$$q_i = 1 - \prod_{j=n-i+1}^{n} \frac{j + \sum\limits_{k=n-j+1}^{n} R_k}{j + 1 + \sum\limits_{k=n-j+1}^{n} R_k},\ i = 1, ..., n, \quad q_{J_1}^* = 1 - \frac{1}{2}(q_{J_1} + q_{J_1+1}),$$

$$\bar{q}_j = 1 - \prod_{i=m-j+1}^{m} \frac{i + \sum\limits_{k=m-i+1}^{m} S_k}{i + 1 + \sum\limits_{k=m-i+1}^{m} S_k},\ j = 1, ..., m, \quad \bar{q}_{J_2}^* = 1 - \frac{1}{2}(\bar{q}_{J_2} + \bar{q}_{J_2+1}).$$

Also, by expanding the functions e^{t_i}, e^{z_j}, e^{δ_1} and e^{δ_2} in Taylor series around the points

$$\begin{aligned} \nu_i &= \log\big(-\log(1 - q_i)\big), & \bar{\nu}_j &= \log\big(-\log(1 - \bar{q}_j)\big), \\ \nu_{J_1}^* &= \log\big(-\log(1 - q_{J_1}^*)\big), & \bar{\nu}_{J_2}^* &= \log\big(-\log(1 - \bar{q}_{J_2}^*)\big), \end{aligned}$$

respectively, and keeping the first order derivatives, we have

$$e^{t_i} = \alpha_i + \beta_i t_i, \quad e^{z_j} = \bar{\alpha}_j + \bar{\beta}_j z_j, \quad e^{\delta_1} = \alpha^*_{J_1} + \beta^*_{J_1}\delta_1, \quad e^{\delta_2} = \bar{\alpha}^*_{J_1} + \bar{\beta}^*_{J_2}\delta_2,$$

where

$$\alpha_i = e^{\nu_i}(1-\nu_i), \qquad \beta_i = e^{\nu_i}, \qquad \bar{\alpha}_j = e^{\bar{\nu}_j}(1-\bar{\nu}_j), \qquad \bar{\beta}_j = e^{\bar{\nu}_j},$$
$$\alpha^*_{J_1} = e^{\nu^*_{J_1}}(1-\nu^*_{J_1}), \ \beta^*_{J_1} = e^{\nu^*_{J_1}}, \ \bar{\alpha}^*_{J_2} = e^{\bar{\nu}^*_{J_2}}(1-\bar{\nu}^*_{J_2}), \ \bar{\beta}^*_{J_2} = e^{\bar{\nu}^*_{J_2}}.$$

Now, if we apply the linear approximations in equations (2.7)–(2.9) and solve them, then the AMLEs of μ_1, μ_2, and σ, say $\tilde{\mu}_1$, $\tilde{\mu}_2$ and $\tilde{\sigma}$, respectively, can be resulted from the following equation:

$$\tilde{\mu}_1 = A_1 - \tilde{\sigma}B_1, \qquad \tilde{\mu}_2 = A_2 - \tilde{\sigma}B_2,$$
$$\tilde{\sigma} = \frac{-(D_1+D_2) + \sqrt{(D_1+D_2)^2 + 4(C_1+C_2)(E_1+E_2)}}{2(C_1+C_2)},$$

where A_1, A_2, B_1, B_2, C_1, C_2, D_1, D_2, E_1, E_2 are given in details in Appendix B. After deriving $\tilde{\mu}_1$, $\tilde{\mu}_2$ and $\tilde{\sigma}$, the AMLEs of α, β, and λ, say $\tilde{\alpha}$, $\tilde{\beta}$ and $\tilde{\lambda}$, respectively, can be evaluated through

$$\tilde{\alpha} = e^{-\frac{\tilde{\mu}_1}{\tilde{\sigma}}}, \quad \tilde{\beta} = e^{-\frac{\tilde{\mu}_2}{\tilde{\sigma}}}, \quad \tilde{\lambda} = \frac{1}{\tilde{\sigma}}.$$

So, the AMLE of R, namely $\tilde{R}$, is

$$\tilde{R} = \frac{\tilde{\alpha}}{\tilde{\alpha}+\tilde{\beta}}. \tag{2.10}$$

2.4. Asymptotic confidence interval

In this section, we obtain the asymptotic confidence interval of R by the asymptotic distribution of $\widehat{R}$, which was obtained from the asymptotic distribution of $\widehat{\theta} = (\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})$. We denote the observed Fisher information matrix by $I(\theta) = [I_{ij}] = \left[-\frac{\partial^2 \ell}{\partial\theta_i\,\partial\theta_j}\right]$, $i,j = 1,2,3$. By differentiating from (2.1) for two times with respect to α, β, and λ, the inlines of $I(\theta)$ matrix can be obtained as:

$$I_{11} = \frac{J_1}{\alpha^2}, \qquad I_{22} = \frac{J_2}{\beta^2}, \qquad I_{12} = I_{21} = 0,$$
$$I_{13} = I_{31} = \sum_{i=1}^{J_1}(R_i+1)x_i^\lambda \frac{\log(x_i)}{1-x_i^\lambda} + R^*_{J_1}T_1^\lambda \frac{\log(T_1)}{1-T_1^\lambda},$$
$$I_{23} = I_{32} = \sum_{j=1}^{J_2}(S_j+1)y_j^\lambda \frac{\log(y_j)}{1-y_j^\lambda} + S^*_{J_2}T_2^\lambda \frac{\log(T_2)}{1-T_2^\lambda},$$
$$I_{33} = \frac{J_1+J_2}{\lambda^2} + \sum_{i=1}^{J_1}\Big(\alpha(R_i+1)-1\Big)x_i^\lambda\Big(\frac{\log(x_i)}{1-x_i^\lambda}\Big)^2 + \alpha R^*_{J_1}T_1^\lambda\Big(\frac{\log(T_1)}{1-T_1^\lambda}\Big)^2$$
$$+ \sum_{j=1}^{J_2}\Big(\beta(S_j+1)-1\Big)y_j^\lambda\Big(\frac{\log(y_j)}{1-y_j^\lambda}\Big)^2 + \beta S^*_{J_2}T_2^\lambda\Big(\frac{\log(T_2)}{1-T_2^\lambda}\Big)^2.$$

Theorem 2.2. *Let $\widehat{\alpha}, \widehat{\beta}$ and $\widehat{\lambda}$ be the MLEs of α, β, and λ, respectively. So*

$$[(\widehat{\alpha}-\alpha),\ (\widehat{\beta}-\beta),\ (\widehat{\lambda}-\lambda)]^{\mathsf{T}} \xrightarrow{D} N_3(0, \mathbf{I^{-1}}(\alpha,\beta,\lambda)),$$

where $\mathbf{I}(\alpha,\beta,\lambda)$ and $\mathbf{I^{-1}}(\alpha,\beta,\lambda)$ are symmetric matrices and

$$\mathbf{I}(\alpha,\beta,\lambda)=\begin{pmatrix} I_{11} & 0 & I_{13} \\ & I_{22} & I_{23} \\ & & I_{33}\end{pmatrix}, \quad \mathbf{I^{-1}}(\alpha,\beta,\lambda)=\frac{1}{|\mathbf{I}(\alpha,\beta,\lambda)|}\begin{pmatrix} b_{11} & b_{12} & b_{13} \\ & b_{22} & b_{23} \\ & & b_{33}\end{pmatrix},$$

in which $|\mathbf{I}(\alpha,\beta,\lambda)| = I_{11}I_{22}I_{33} - I_{11}I_{23}^2 - I_{13}^2 I_{22}$,

$$\begin{aligned} b_{11} &= I_{22}I_{33}-I_{23}^2, & b_{12} &= I_{13}I_{23}, & b_{13} &= -I_{13}I_{22}, \\ b_{22} &= I_{11}I_{33}-I_{13}^2, & b_{23} &= -I_{11}I_{23}, & b_{33} &= I_{11}I_{22}. \end{aligned}$$

Proof: From the asymptotic normality of the MLE, the theorem would be resulted. □

Theorem 2.3. *Let $\widehat{R}^{\text{MLE}}$ be the MLE of R. So,*

$$(\widehat{R}^{\text{MLE}} - R) \xrightarrow{D} N(0, B),$$

where

$$B = \frac{1}{|\mathbf{I}(\alpha,\beta,\lambda)|}\left[(\frac{\partial R}{\partial \alpha})^2 b_{11} + (\frac{\partial R}{\partial \beta})^2 b_{22} + 2(\frac{\partial R}{\partial \alpha})(\frac{\partial R}{\partial \beta}) b_{12}\right]. \tag{2.11}$$

Proof: Using Theorem 2.2 and applying the delta method, the asymptotic distribution of $\widehat{R} = \frac{\widehat{\alpha}}{\widehat{\alpha}+\widehat{\beta}}$ can be obtained as follows:

$$(\widehat{R}^{\text{MLE}} - R) \xrightarrow{D} N(0, B),$$

where $B = \mathbf{b}^{\mathsf{T}}\mathbf{I^{-1}}(\alpha,\beta,\lambda)\mathbf{b}$, with $\mathbf{b} = [\frac{\partial R}{\partial \alpha},\ \frac{\partial R}{\partial \beta},\ \frac{\partial R}{\partial \lambda}]^{\mathsf{T}} = [\frac{\partial R}{\partial \alpha},\ \frac{\partial R}{\partial \beta},\ 0]^{\mathsf{T}}$, in which

$$\frac{\partial R}{\partial \alpha} = \frac{\beta}{(\alpha+\beta)^2}, \qquad \frac{\partial R}{\partial \beta} = -\frac{\alpha}{(\alpha+\beta)^2}, \tag{2.12}$$

and $\mathbf{I^{-1}}(\alpha,\beta,\lambda)$ is defined in Theorem 2.2. Therefore, B can be represented as (2.11) and the theorem results. □

Using Theorem 2.3, the asymptotic confidence interval of R can be derived. It is notable that B should be estimated by the MLEs of α, β, and λ. So, a $100(1-\gamma)\%$ asymptotic confidence interval of R can be constructed as

$$(\widehat{R}^{\text{MLE}} - z_{1-\frac{\gamma}{2}}\sqrt{\widehat{B}}, \widehat{R}^{\text{MLE}} + z_{1-\frac{\gamma}{2}}\sqrt{\widehat{B}}),$$

where z_γ is 100γ-th percentile of $N(0,1)$.

2.5. Bayes estimation

In this section, under the squared error loss function, we infer the Bayesian estimation and corresponding credible interval of the stress-strength parameter, when $\alpha \sim \Gamma(a_1, b_1)$, $\beta \sim \Gamma(a_2, b_2)$ and $\lambda \sim \Gamma(a_3, b_3)$ are independent random variables. Accordingly, based on the observed censoring samples, the joint posterior density function of α, β and λ are achieved by:

$$\pi(\alpha, \beta, \lambda|\text{data}) = \frac{L(\text{data}|\alpha, \beta, \lambda)\pi_1(\alpha)\pi_2(\beta)\pi_3(\lambda)}{\int_0^\infty \int_0^\infty \int_0^\infty L(\text{data}|\alpha, \beta, \lambda)\pi_1(\alpha)\pi_2(\beta)\pi_3(\lambda)d\alpha d\beta d\lambda}, \tag{2.13}$$

where

$$\begin{aligned}
\pi_1(\alpha) &\propto \alpha^{a_1-1}e^{-b_1\alpha}, && \alpha > 0, \;\; a_1, b_1 > 0,\\
\pi_2(\beta) &\propto \beta^{a_2-1}e^{-b_2\beta}, && \beta > 0, \;\; a_2, b_2 > 0,\\
\pi_3(\lambda) &\propto \lambda^{a_3-1}e^{-b_3\lambda}, && \lambda > 0, \;\; a_3, b_3 > 0.
\end{aligned}$$

As we observe from (2.13), the Bayes estimates cannot be derived in the closed-form. Therefore, we approximate them by applying two following methods:

- Lindley's approximation,
- MCMC method.

2.5.1. Lindley's approximation

One of the most applicable numerical methods to approximate the Bayes estimate has been introduced by Lindley in [16]. This method can be described as follows. Let $U(\theta)$ be a function of the parameter value. The Bayes estimate of $U(\theta)$, under the squared error loss function, is

$$\mathbb{E}\big(u(\theta)|\text{data}\big) = \frac{\int u(\theta)e^{Q(\theta)}d\theta}{\int e^{Q(\theta)}d\theta},$$

where $Q(\theta) = \ell(\theta) + \rho(\theta)$, $\ell(\theta)$ and $\rho(\theta)$ are the logarithm of likelihood function and prior density of θ, respectively. Lindley has been approximated $E(u(\theta)|\text{data})$ as

$$\mathbb{E}\big(u(\theta)|\text{data}\big) = u + \frac{1}{2}\sum_i\sum_j(u_{ij} + 2u_i\rho_j)\sigma_{ij} + \frac{1}{2}\sum_i\sum_j\sum_k\sum_p \ell_{ijk}\sigma_{ij}\sigma_{kp}u_p\Big|_{\theta=\widehat{\theta}},$$

where $\theta = (\theta_1, ..., \theta_m)$, $i, j, k, p = 1, ..., m$, $\widehat{\theta}$ is the MLE of θ, $u = u(\theta)$, $u_i = \partial u/\partial\theta_i$, $u_{ij} = \partial^2 u/\partial\theta_i\partial\theta_j$, $\ell_{ijk} = \partial^3\ell/\partial\theta_i\partial\theta_j\partial\theta_k$, $\rho_j = \partial\rho/\partial\theta_j$, and $\sigma_{ij} = (i, j)$-th element in the inverse of matrix $[-\ell_{ij}]$ all calculated at the MLE of parameters.

When we face up to the case of three parameter $\theta = (\theta_1, \theta_2, \theta_3)$, Lindley's approximation conducts to

$$\begin{aligned}
\mathbb{E}(u(\theta)|\text{data}) \;=\;& u + (u_1d_1 + u_2d_2 + u_3d_3 + d_4 + d_5) + \frac{1}{2}\Big[A(u_1\sigma_{11} + u_2\sigma_{12} + u_3\sigma_{13})\\
&+\; B(u_1\sigma_{21} + u_2\sigma_{22} + u_3\sigma_{23}) + C(u_1\sigma_{31} + u_2\sigma_{32} + u_3\sigma_{33})\Big],
\end{aligned} \tag{2.14}$$

which their elements are presented in detail in Appendix C. Therefore, the Bayes estimate of R is

$$\begin{aligned}\widehat{R}^{\mathrm{Lin}}_{s,k} &= R + [u_1 d_1 + u_2 d_2 + d_4 + d_5] + \frac{1}{2}\Big[A(u_1\sigma_{11} + u_2\sigma_{12}) \\ &\quad + B(u_1\sigma_{21} + u_2\sigma_{22}) + C(u_1\sigma_{31} + u_2\sigma_{32})\Big]. \end{aligned} \tag{2.15}$$

It should be noted that all parameters are evaluated at $(\widehat{\alpha}, \widehat{\beta}, \widehat{\lambda})$, respectively.

As we observe, constructing the HPD credible interval is not possible by using the Lindley's approximation. So, we apply the Markov Chain Monte Carlo (MCMC) method to approximate the Bayes estimate and construct the corresponding HPD credible intervals.

2.5.2. MCMC method

After simplify equation (2.13), we get the posterior pdfs of α, β and λ as:

$$\begin{aligned}\alpha|\lambda, \text{data} &\sim \Gamma\Big(J_1 + a_1, b_1 - \sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda}) - R^*_{J_1}\log(1-T_1^{\lambda})\Big), \\ \beta|\lambda, \text{data} &\sim \Gamma\Big(J_2 + a_2, b_2 - \sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda}) - S^*_{J_2}\log(1-T_2^{\lambda})\Big), \\ \pi(\lambda|\alpha,\beta,\text{data}) &\propto \left(\prod_{i=1}^{J_1} x_i^{\lambda-1}(1-x_i^{\lambda})^{\alpha(R_i+1)-1}\right)\left(\prod_{j=1}^{J_2} y_j^{\lambda-1}(1-y_j^{\lambda})^{\beta(S_j+1)-1}\right) \\ &\quad\times\ \lambda^{J_1+J_2+a_3-1}e^{-\lambda b_3}(1-T_1^{\lambda})^{\alpha R^*_{J_1}}(1-T_2^{\lambda})^{\beta S^*_{J_2}}.\end{aligned}$$

It is identified that the posterior pdf of λ is not a well known distribution. Therefore, we utilize the Metropolis–Hastings method with normal proposal distribution in order to generate random samples from it. Consequently, the Gibbs sampling algorithm can be proposed as follows:

1. Start with the begin value $(\alpha_{(0)}, \beta_{(0)}, \lambda_{(0)})$.

2. Set $t = 1$.

3. Generate $\lambda_{(t)}$ from $\pi(\lambda|\alpha_{(t-1)}, \beta_{(t-1)}, \text{data})$, using Metropolis–Hastings method.

4. Generate $\alpha_{(t)}$ from $\Gamma\big(J_1 + a_1, b_1 - \sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda_{(t-1)}} - R^*_{J_1}\log(1-T_1^{\lambda_{(t-1)}})\big)$.

5. Generate $\beta_{(t)}$ from $\Gamma\big(J_2 + a_2, b_2 - \sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda_{(t-1)}}) - S^*_{J_2}\log(1-T_2^{\lambda_{(t-1)}})\big)$.

6. Calculate $R_t = \frac{\alpha_t}{\alpha_t + \beta_t}$.

7. Set $t = t + 1$.

8. Repeat steps 3–7, for T times.

By applying this algorithm, the Bayes estimate of R, under the squared error loss function is resulted from

$$\widehat{R}^{\mathrm{MC}} = \frac{1}{T-M}\sum_{t=M+1}^{T} R_t, \tag{2.16}$$

where M is the burn-in period. Moreover, a $100(1-\gamma)\%$ HPD credible interval of R can be constructed by applying the method conducted by Chen and Shao [4].

3. INFERENCE ON R WITH KNOWN COMMON λ

3.1. MLE of R

Consider that $\{X_1, ..., X_n\}$ and $\{Y_1, ..., Y_m\}$ be two HP censoring samples with censoring schemes $\{N, n, T_1, J_1, R_1, ..., R_{J_1}\}$ and $\{M, m, T_2, J_2, S_1, ..., S_{J_2}\}$, respectively. Based on Section 2.1, when the common shape parameter λ is known, the MLE of R can be attained easily by the following equation:

$$\widehat{R}^{\mathrm{MLE}} = \left(1 + \frac{J_2\big(\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda}) + R^*_{J_1}\log(1-T_1^{\lambda})\big)}{J_1\big(\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda}) + S^*_{J_2}\log(1-T_2^{\lambda})\big)}\right)^{-1}. \tag{3.1}$$

In a similar manner to Section 2.4, $(\widehat{R}^{\mathrm{MLE}} - R) \xrightarrow{D} N(0, C)$, where $C = (\frac{\partial R}{\partial \alpha})^2 \frac{1}{I_{11}} + (\frac{\partial R}{\partial \beta})^2 \frac{1}{I_{22}}$, and $\frac{\partial R}{\partial \alpha}$ and $\frac{\partial R}{\partial \beta}$ are indicated in (2.12). Consequently, a $100(1-\gamma)\%$ asymptotic confidence interval for R can be constructed as

$$\Big(\widehat{R}^{\mathrm{MLE}} - z_{1-\frac{\gamma}{2}}\sqrt{\widehat{C}},\ \widehat{R}^{\mathrm{MLE}} + z_{1-\frac{\gamma}{2}}\sqrt{\widehat{C}}\Big),$$

where z_γ is 100γ-th percentile of $N(0,1)$.

3.2. Bayes estimation

In this section, we infer the Bayesian estimation and corresponding credible interval of the stress-strength parameter, when $\alpha \sim \Gamma(a_1, b_1)$ and $\beta \sim \Gamma(a_2, b_2)$ are independent random variables. With respect to the observed censoring samples, the joint posterior density function of α and β are given by:

$$\pi(\alpha, \beta|\lambda, \text{data}) = \frac{(V+b_1)^{J_1+a_1}(U+b_2)^{J_2+a_2}}{\Gamma(J_1+a_1)\Gamma(J_2+a_2)}\alpha^{J_1+a_1-1}\beta^{J_2+a_2-1}e^{-\alpha(V+b_1)-\beta(U+b_2)}, \tag{3.2}$$

where

$$V = -\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda}) - R^*_{J_1}\log(1-T_1^{\lambda}),$$
$$U = -\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda}) - S^*_{J_2}\log(1-T_2^{\lambda}).$$

Under the squared error loss function, for obtaining R Bayes estimate, we solve the following integral:

$$\widehat{R}^{\mathrm{B}} = \int_0^{\infty}\int_0^{\infty}\frac{\alpha}{\alpha+\beta}\times\pi(\alpha,\beta|\lambda,\text{data})d\alpha d\beta.$$

Now in this study, we use the idea of Kizilaslan and Nadar [8], and accordingly, obtain the R Bayes estimate as

$$\widehat{R}^{\mathrm{B}} = \begin{cases} \dfrac{(1-z)^{J_1+a_1}(J_1+a_1)}{w}\ {}_2F_1(w, J_1+a_1+1; w+1, z) & \text{if } \ |z|<1, \\[2ex] \dfrac{(J_1+a_1)}{w(1-z)^{J_2+a_2}}\ {}_2F_1(w, J_2+a_2; w+1, \dfrac{z}{1-z}) & \text{if } \ z<-1, \end{cases} \tag{3.3}$$

where $w = J_1+J_2+a_1+a_2$, $z = 1-\dfrac{V+b_1}{U+b_2}$ and

$${}_2F_1(\alpha,\beta;\gamma,z) = \frac{1}{B(\beta,\gamma-\beta)}\int_0^1 t^{\beta-1}(1-t)^{\gamma-\beta-1}(1-tz)^{-\alpha}dt, \ |z|<1,$$

is the hypergeometric series, which is quickly evaluated and readily available in standard software like MATLAB. Moreover, we construct a $100(1-\gamma)\%$ Bayesian interval for the stress-strength parameter by (L,U), where L and U are the lower and upper bounds, respectively, which indicate

$$\int_0^L f_R(R)dR = \frac{\gamma}{2}, \quad \int_0^U f_R(R)dR = 1-\frac{\gamma}{2}, \tag{3.4}$$

where $f_R(R)$ is the probability density function of R, which obtained from (3.2) as

$$f_R(R) = \frac{(1-z)^{J_1+a_1}R^{J_1+a_1-1}(1-R)^{J_2+a_2-1}(1-Rz)^{-w}}{B(J_1+a_1,J_2+a_2)}, \quad 0<R<1.$$

4. ESTIMATION OF R IN GENERAL CASE

4.1. MLE of R

The stress-strength parameter, when X and Y are two independent random variables from $\mathrm{Ku}(\alpha,\lambda_1)$ and $\mathrm{Ku}(\beta,\lambda_2)$, respectively, can be obtained as

$$\begin{aligned}
R &= P(X<Y)\\
&= \int_0^1 f_Y(y)F_X(y)dy\\
&= \int_0^1 \beta\lambda_2 y^{\lambda_2-1}(1-y^{\lambda_2})^{\beta-1}\big(1-(1-y^{\lambda_1})^{\alpha}\big)dy\\
&= 1-\int_0^1 \beta\lambda_2 y^{\lambda_2-1}(1-y^{\lambda_2})^{\beta-1}(1-y^{\lambda_1})^{\alpha}dy.
\end{aligned}$$

Assume that $\{X_1,...,X_n\}$ and $\{Y_1,...,Y_m\}$ are two HP censoring samples with censoring schemes $\{N,n,T_1,J_1,R_1,...,R_{J_1}\}$ and $\{M,m,T_2,J_2,S_1,...,S_{J_2}\}$, respectively. As a result, the likelihood function of the unknown parameters α, β, λ_1 and λ_2 can be written as

$$\begin{aligned}
L(\text{data}|\alpha,\beta,\lambda_1,\lambda_2) \;&\propto\; \alpha^{J_1}\lambda_1^{J_1}\left(\prod_{i=1}^{J_1} x_i^{\lambda_1-1}(1-x_i^{\lambda_1})^{\alpha(R_i+1)-1}\right)(1-T_1^{\lambda_1})^{\alpha R^*_{J_1}}\\
&\times\; \beta^{J_2}\lambda_2^{J_2}\left(\prod_{j=1}^{J_2} y_j^{\lambda_2-1}(1-y_j^{\lambda_2})^{\beta(S_j+1)-1}\right)(1-T_2^{\lambda_2})^{\beta S^*_{J_2}}.
\end{aligned}$$

Therefore, the log-likelihood function, along with ignoring the constant value, is as:

$$\begin{aligned}
\ell(\alpha,\beta,\lambda_1,\lambda_2) \;&=\; J_1\log(\alpha\lambda_1)+J_2\log(\beta\lambda_2)+\sum_{i=1}^{J_1}(\alpha(R_i+1)-1)\log(1-x_i^{\lambda_1})\\
&+\sum_{j=1}^{J_2}(\beta(S_j+1)-1)\log(1-y_j^{\lambda_2})+\alpha R^*_{J_1}\log(1-T_1^{\lambda_1})\\
&+\beta S^*_{J_2}\log(1-T_2^{\lambda_2})+(\lambda_1-1)\sum_{i=1}^{J_1}\log(x_i)+(\lambda_2-1)\sum_{j=1}^{J_2}\log(y_j).
\end{aligned}$$

In a similar manner as Section 2.1, $\widehat{\alpha}$ and $\widehat{\beta}$, respectively, can be obtained from

$$\begin{aligned}
\widehat{\alpha}(\lambda_1) \;&=\; -J_1\Big\{\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda_1})+R^*_{J_1}\log(1-T_1^{\lambda_1})\Big\}^{-1},\\
\widehat{\beta}(\lambda_2) \;&=\; -J_2\Big\{\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda_2})+S^*_{J_2}\log(1-T_2^{\lambda_2})\Big\}^{-1}.
\end{aligned}$$

Also, to derive $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$, respectively, we apply one numerical method like Newton–Raphson on the following equations:

$$\begin{aligned}
\frac{\partial \ell}{\partial \lambda_1} &= \frac{J_1}{\lambda_1} + \sum_{i=1}^{J_1} \log(x_i) - \sum_{i=1}^{J_1} (\alpha(R_i+1)-1) x_i^{\lambda_1} \frac{\log(x_i)}{1-x_i^{\lambda_1}} - \alpha R_{J_1}^* T_1^{\lambda_1} \frac{\log(T_1)}{1-T_1^{\lambda_1}} &= 0, \\
\frac{\partial \ell}{\partial \lambda_2} &= \frac{J_2}{\lambda_2} + \sum_{j=1}^{J_2} \log(y_j) - \sum_{j=1}^{J_2} (\beta(S_j+1)-1) y_j^{\lambda_2} \frac{\log(y_j)}{1-y_j^{\lambda_2}} - \beta S_{J_2}^* T_2^{\lambda_2} \frac{\log(T_2)}{1-T_2^{\lambda_2}} &= 0.
\end{aligned}$$

After obtaining the MLEs of α, β, λ_1, and λ_2, by using the invariance property, the MLE of R can be derived as

$$\widehat{R}^{\mathrm{MLE}} = 1 - \int_0^1 \widehat{\beta}\widehat{\lambda_2} y^{\widehat{\lambda}_2 - 1} (1 - y^{\widehat{\lambda}_2})^{\widehat{\beta}-1} (1 - y^{\widehat{\lambda}_1})^{\widehat{\alpha}} dy. \tag{4.1}$$

4.2. AMLE of R

In this section, we obtain AMLE of R. Consider $\{X_1, ..., X_n\}$ and $\{Y_1, ..., Y_m\}$ are two HP censoring samples with censoring schemes $\{N, n, T_1, J_1, R_1, ..., R_{J_1}\}$ and also by considering $\{M, m, T_2, J_2, S_1, ..., S_{J_2}\}$ from the distributions $\mathrm{Ku}(\alpha, \lambda_1)$ and $\mathrm{Ku}(\beta, \lambda_2)$, respectively, and

$$X_i' = (-\log(1 - X_i^{\lambda_1}))^{\frac{1}{\lambda_1}}, U_i = \log(X_i') \text{ and } Y_j' = (-\log(1 - Y_j^{\lambda_2}))^{\frac{1}{\lambda_2}}, V_j = \log(Y_j').$$

Based on the observed data $\{U_1, ..., U_n\}$ and $\{V_1, ..., V_m\}$, along with ignoring the constant value, the log-likelihood function is obtained as follows:

$$\begin{aligned}
\ell^*(\mu_1, \mu_2, \sigma_1, \sigma_2) &= -J_1 \log(\sigma_1) + \sum_{i=1}^{J_1} t_i - \sum_{i=1}^{J_1} (R_i+1) e^{t_i} - R_{J_1}^* e^{\delta_1} \\
&\quad - J_2 \log(\sigma_2) + \sum_{j=1}^{J_2} z_j - \sum_{j=1}^{J_2} (S_j+1) e^{z_j} - S_{J_2}^* e^{\delta_2},
\end{aligned} \tag{4.2}$$

where

$$\begin{aligned}
&t_i = \frac{u_i - \mu_1}{\sigma_1},\ z_j = \frac{v_j - \mu_2}{\sigma_2},\ \mu_1 = \frac{-\log(\alpha)}{\lambda_1},\ \mu_2 = \frac{-\log(\beta)}{\lambda_2}, \\
&\delta_p = \frac{a_p - \mu_p}{\sigma_p},\ \sigma_p = \frac{1}{\lambda_p},\ a_p = \log((-\log(1 - T_p^{\lambda_p}))^{\frac{1}{\lambda_p}}),\ p = 1, 2.
\end{aligned}$$

Now by taking derivatives due to μ_1, μ_2, σ_1 and σ_2 from (4.2), we obtain the following equations:

$$\frac{\partial \ell^*}{\partial \mu_1} = -\frac{1}{\sigma_1}\left[J_1 - \sum_{i=1}^{J_1}(R_i+1)e^{t_i} - R^*_{J_1}e^{\delta_1}\right] = 0,$$
$$\frac{\partial \ell^*}{\partial \mu_2} = -\frac{1}{\sigma_2}\left[J_2 - \sum_{j=1}^{J_2}(S_j+1)e^{z_j} - S^*_{J_2}e^{\delta_2}\right] = 0,$$
$$\frac{\partial \ell^*}{\partial \sigma_1} = -\frac{1}{\sigma_1}\left[J_1 + \sum_{i=1}^{J_1}t_i - \sum_{i=1}^{J_1}(R_i+1)t_ie^{t_i} - R^*_{J_1}\delta_1e^{\delta_1}\right] = 0,$$
$$\frac{\partial \ell^*}{\partial \sigma_2} = -\frac{1}{\sigma_2}\left[J_2 + \sum_{j=1}^{J_2}z_j - \sum_{j=1}^{J_2}(S_j+1)z_je^{z_j} - S^*_{J_2}\delta_2e^{\delta_2}\right] = 0.$$

In a similar manner as Section 2.3, we derive the AMLEs of μ_1, μ_2, σ_1 and σ_2, say $\tilde{\mu}_1$, $\tilde{\mu}_2$, $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$, respectively, by

$$\tilde{\mu}_1 = A_1 - \tilde{\sigma}_1 B_1, \qquad \tilde{\mu}_2 = A_2 - \tilde{\sigma}_2 B_2,$$
$$\tilde{\sigma}_1 = \frac{-D_1 + \sqrt{D_1^2 + 4C_1E_1}}{2C_1}, \quad \tilde{\sigma}_2 = \frac{-D_2 + \sqrt{D_2^2 + 4C_2E_2}}{2C_2},$$

where A_1, A_2, B_1, B_2, C_1, C_2, D_1, D_2, E_1, E_2 are given in Section 2.3. After achieving $\tilde{\mu}_1$, $\tilde{\mu}_2$, $\tilde{\sigma}_1$, and $\tilde{\sigma}_2$, the values of $\tilde{\alpha}$, $\tilde{\beta}$, $\tilde{\lambda}_1$, $\tilde{\lambda}_2$ and $\tilde{R}$ can be evaluated by $\tilde{\alpha} = e^{-\frac{\tilde{\mu}_1}{\tilde{\sigma}_1}}$, $\tilde{\beta} = e^{-\frac{\tilde{\mu}_2}{\tilde{\sigma}_2}}$, $\tilde{\lambda}_1 = \frac{1}{\tilde{\sigma}_1}$, $\tilde{\lambda}_2 = \frac{1}{\tilde{\sigma}_2}$ and consequently

$$\tilde{R} = 1 - \int_0^1 \tilde{\beta}\tilde{\lambda}_2 y^{\tilde{\lambda}_2-1}(1-y^{\tilde{\lambda}_2})^{\tilde{\beta}-1}(1-y^{\tilde{\lambda}_1})^{\tilde{\alpha}}dy. \tag{4.3}$$

4.3. Bayes estimation

In this section, under the squared error loss function, we infer the Bayesian estimation and corresponding credible interval of the stress-strength parameter, when the unknown parameters $\alpha \sim \Gamma(a_1, b_1)$, $\beta \sim \Gamma(a_2, b_2)$, $\lambda_1 \sim \Gamma(a_3, b_3)$ and $\lambda_2 \sim \Gamma(a_4, b_4)$ are independent random variables. In a same manner as Section 2.5, as the Bayesian estimation of R has not a closed-form, we approximate it by applying MCMC method. After simplifying the joint posterior density function of the unknown parameters, we get the posterior pdfs of α, β, λ_1 and λ_2 as:

$$\alpha|\lambda_1, \text{data} \sim \Gamma\big(J_1 + a_1, b_1 - \sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda_1}) - R^*_{J_1}\log(1-T_1^{\lambda_1})\big),$$
$$\beta|\lambda_2, \text{data} \sim \Gamma\big(J_2 + a_2, b_2 - \sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda_2}) - S^*_{J_2}\log(1-T_2^{\lambda_2})\big),$$
$$\pi(\lambda_1|\alpha, \text{data}) \propto \lambda_1^{J_1+a_3-1}\left(\prod_{i=1}^{J_1}x_i^{\lambda_1-1}(1-x_i^{\lambda_1})^{\alpha(R_i+1)-1}\right)(1-T_1^{\lambda_1})^{\alpha R^*_{J_1}}e^{-\lambda_1 b_3}$$
$$\pi(\lambda_2|\beta, \text{data}) \propto \lambda_2^{J_2+a_4-1}\left(\prod_{j=1}^{J_2}y_j^{\lambda_2-1}(1-y_j^{\lambda_2})^{\beta(S_j+1)-1}\right)(1-T_2^{\lambda_2})^{\beta S^*_{J_2}}e^{-\lambda_2 b_4}.$$

It is recognized that the posterior pdfs of λ_1 and λ_2 are not well known distributions. So, we utilize the Metropolis–Hastings method with normal proposal distribution for generating random samples from them. Therefore, the Gibbs sampling algorithm can be proposed as follows:

1. Start with the begin value $(\alpha_{(0)}, \beta_{(0)}, \lambda_{1(0)}, \lambda_{2(0)})$.

2. Set $t = 1$.

3. Generate $\lambda_{1(t)}$ from $\pi(\lambda_1|\alpha_{(t-1)}, \text{data})$, using Metropolis–Hastings method.

4. Generate $\lambda_{2(t)}$ from $\pi(\lambda_2|\beta_{(t-1)}, \text{data})$, using Metropolis–Hastings method.

5. Generate $\alpha_{(t)}$ from $\Gamma\big(J_1+a_1, b_1-\sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^{\lambda_{1(t-1)}}-R^*_{J_1}\log(1-T_1^{\lambda_{1(t-1)}})\big)$.

6. Generate $\beta_{(t)}$ from $\Gamma\big(J_2+a_2, b_2-\sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^{\lambda_{2(t-1)}})-S^*_{J_2}\log(1-T_2^{\lambda_{2(t-1)}})\big)$.

7. Calculate $R_t = 1-\int_0^1 \beta_{(t)}\lambda_{2(t)}y^{\lambda_{2(t)}-1}(1-y^{\lambda_{2(t)}})^{\beta_{(t)}-1}(1-y^{\lambda_{1(t)}})^{\alpha_t}dy$.

8. Set $t = t+1$.

9. Repeat steps 3–8, for T times.

Using this algorithm, under the squared error loss function, the R Bayes estimate will be resulted from

$$\widehat{R}^{\text{MC}} = \frac{1}{T-M}\sum_{t=M+1}^{T} R_t, \tag{4.4}$$

where M is the burn-in period. Moreover, a $100(1-\gamma)\%$ HPD credible interval of R can be constructed by applying the method accomplished by Chen and Shao [4].

5. SIMULATION STUDY AND DATA ANALYSIS

In this section, we compare the performance of different methods by Monte Carlo simulations and analyze two real data sets to illustrative aims.

5.1. Numerical experiments and discussions

In this section, we compare the behavior of various estimates by Monte Carlo simulations, under different censoring schemes. The comparison among estimates is accomplished in terms of mean squared errors (MSEs). Also, the comparison of confidence intervals is performed in terms of average lengths and coverage percentages. We apply different schemes, parameters, and hyper parameters to implement the simulation study. All results are reported based on 3000 replications. Also, the nominal level is 0.95 in comparison with the confidence intervals. We utilize the different censoring schemes as:

$$\begin{aligned}
&\text{Scheme 1: } R_1 = ... = R_{n-1} = 0,\ R_n = N - n,\\
&\text{Scheme 2: } R_1 = ... = R_n = \frac{N-n}{n},\\
&\text{Scheme 3: } R_1 = ... = R_{\frac{n}{2}} = 0,\ R_{\frac{n}{2}+1} = ... = R_n = \frac{2(N-n)}{n}.
\end{aligned}$$

We can interpret theses schemes as follows. In Scheme 1, the number of removal units at the first, second and so on until reaching the $(n-1)$-th failure times is zero and we remove all $N-n$ units at the n-th failure time. We use Scheme 2 and 3 when $N-n$ to be divisible by n, and n must be an even number. In Scheme 2, the number of removal units at the first, second and so on until reaching the (n)-th failure times is $\frac{N-n}{n}$. In Scheme 3, the number of removal units at the first, second and so on until reaching the $(\frac{n}{2})$-th failure times is zero and the number of removal units at the $\frac{n}{2}+1$, and so on up to the (n)-th failure times is $\frac{2(N-n)}{n}$. All of these schemes are considered for two values of T as 0.7 and 0.9, respectively.

In the First case, by assuming the unknown common shape parameter λ, we choose $\alpha = \beta = \lambda = 2$, without any loss of generality. Also, Bayesian inference are given in terms of three priors as: Prior 1: $a_j = 0,\ b_j = 0,\ j = 1,2,3$, Prior 2: $a_j = 1,\ b_j = 0.1,\ j = 1,2,3$, and Prior 3: $a_j = 2,\ b_j = 0.2,\ j = 1,2,3$. Moreover, we noted that the number of iterations in the MCMC method is $T = 5000$, and the threshold of burn-in is 2000. In this case, we obtained the Biases and MSEs of MLE using (2.5), AMLE using (2.10), Bayes estimates of R through Lindley's approximation and MCMC method using (2.15) and (2.16), respectively. The results are shown in Table 1. Additionally, we derived the asymptotic confidence and HPD credible intervals of R. Theses results are displayed in Table 2. By the above chosen, R was obtained equal to 0.5. Also, using the numerical method, we obtain the mean and variance of R as a random variable. Based on Priors 2 and 3, the variance of R is 0.0833 and 0.05, respectively, and the mean of R is 0.5 for both priors. So we expect that the performance of MSE is the best using Prior 3.

In the second case, by assuming the known common shape parameter λ, we choose $\alpha = \beta = \lambda = 3$, without loss of generality. Also, Bayesian inference are given in terms of three priors as: Prior 4: $a_j = 0,\ b_j = 0,\ j = 1,2$, Prior 5: $a_j = 1,\ b_j = 0.1,\ j = 1,2$, and Prior 6: $a_j = 2,\ b_j = 0.2,\ j = 1,2$. In this case, we obtained the Biases and MSEs of MLE, Bayes estimates and 95% Bayesian intervals of R using (3.1), (3.3) and (3.4), respectively. The results are indicated in Table 3. Similar to the previous case, we expect that the performance of MSE be the best using Prior 6.

In the third case, assuming the different second shape parameters λ_1 and λ_2, we choose $\alpha = \beta = \lambda_1 = \lambda_2 = 2$, without any loss of generality. Also, Bayesian inference are presented based on three priors as: Prior 7: $a_j = 0,\ b_j = 0,\ j = 1,2,3,4$, Prior 8: $a_j = 1,\ b_j = 0.1,\ j = 1,2,3,4$, and Prior 9: $a_j = 2,\ b_j = 0.2,\ j = 1,2,3,4$. Also, we noted that the number of iterations in the MCMC method is $T = 5000$, and the threshold of burn-in is 2000. In this case, we obtained the Biases and MSEs of MLE, AMLE and Bayes estimate by applying MCMC method using (4.1), (4.3) and (4.4), respectively. Also, the results are indicated in Table 4.

Table 1: Biases and MSEs for estimates of R when λ is unknown.

(N,n,T)	C.S	AMLE		MLE		Prior 1				Prior 2				Prior 3			
						MCMC		Lindley		MCMC		Lindley		MCMC		Lindley	
		|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE
(40,10,0.7)	(1,1)	0.0119	0.0179	0.0147	0.0175	0.0147	0.0103	0.0139	0.0172	0.0140	0.0099	0.0157	0.0165	0.0136	0.0095	0.0176	0.0137
	(2,2)	0.0176	0.0220	0.0187	0.0268	0.0153	0.0115	0.0175	0.0215	0.0154	0.0110	0.0178	0.0210	0.0150	0.0107	0.0181	0.0149
	(3,3)	0.0234	0.0175	0.0258	0.0179	0.0213	0.0100	0.0209	0.0161	0.0206	0.0096	0.0242	0.0140	0.0252	0.0093	0.0262	0.0123
	(1,2)	0.1521	0.0614	0.1547	0.0699	0.1383	0.0322	0.1463	0.0524	0.1364	0.0314	0.1633	0.0435	0.1340	0.0305	0.1802	0.0358
(60,10,0.7)	(1,1)	0.0015	0.0209	0.0038	0.0204	0.0047	0.0113	0.0036	0.0199	0.0039	0.0110	0.0071	0.0177	0.0041	0.0106	0.0022	0.0174
	(2,2)	0.0010	0.0195	0.0024	0.0192	0.0026	0.0107	0.0001	0.0184	0.0023	0.0104	0.0009	0.0162	0.0025	0.0101	0.0016	0.0156
	(3,3)	0.0165	0.0178	0.0163	0.0178	0.0170	0.0095	0.0152	0.0169	0.0160	0.0092	0.0160	0.0144	0.0161	0.0092	0.0167	0.0123
	(1,2)	0.0513	0.0532	0.1497	0.0531	0.1418	0.0353	0.1415	0.0526	0.1397	0.0349	0.1617	0.0420	0.1379	0.0346	0.1819	0.0358
(40,20,0.7)	(1,1)	0.0057	0.0086	0.0047	0.0085	0.0058	0.0050	0.0045	0.0076	0.0054	0.0049	0.0048	0.0068	0.0051	0.0049	0.0051	0.0061
	(2,2)	0.0075	0.0089	0.0073	0.0087	0.0062	0.0055	0.0070	0.0077	0.0063	0.0054	0.0071	0.0070	0.0068	0.0052	0.0072	0.0064
	(3,3)	0.0109	0.0063	0.0160	0.0064	0.0148	0.0042	0.0153	0.0055	0.0148	0.0042	0.0157	0.0053	0.0137	0.0040	0.0161	0.0050
	(1,2)	0.2172	0.0485	0.1702	0.0479	0.1647	0.0311	0.1651	0.0395	0.1639	0.0308	0.1753	0.0355	0.1629	0.0304	0.1856	0.0316
(60,20,0.7)	(1,1)	0.0035	0.0082	0.0031	0.0086	0.0023	0.0049	0.0030	0.0077	0.0022	0.0048	0.0033	0.0068	0.0022	0.0048	0.0035	0.0059
	(2,2)	0.0011	0.0083	0.0011	0.0087	0.0010	0.0055	0.0011	0.0082	0.0009	0.0054	0.0010	0.0073	0.0010	0.0053	0.0011	0.0065
	(3,3)	0.0023	0.0114	0.0020	0.0163	0.0017	0.0054	0.0019	0.0073	0.0018	0.0053	0.0021	0.0067	0.0017	0.0052	0.0022	0.0061
	(1,2)	0.1870	0.0466	0.1709	0.0469	0.1709	0.0352	0.1657	0.0426	0.1697	0.0348	0.01774	0.0376	0.1687	0.0345	0.1892	0.0356
(40,10,0.9)	(1,1)	0.0134	0.0172	0.0324	0.0170	0.0292	0.0097	0.0306	0.0161	0.0284	0.0094	0.0347	0.0147	0.0284	0.0092	0.0388	0.0122
	(2,2)	0.0084	0.0188	0.0031	0.0189	0.0054	0.0087	0.0030	0.0168	0.0052	0.0084	0.0058	0.0138	0.0055	0.0083	0.0087	0.0114
	(3,3)	0.0122	0.0146	0.0173	0.0143	0.0146	0.0093	0.0163	0.0133	0.0146	0.0092	0.0176	0.0129	0.0137	0.0088	0.0189	0.0109
	(1,2)	0.1526	0.0509	0.1542	0.0541	0.1384	0.0276	0.1455	0.0454	0.1366	0.0270	0.1629	0.0374	0.1347	0.0262	0.1804	0.0305
(60,10,0.9)	(1,1)	0.0040	0.0201	0.0066	0.0203	0.0022	0.0098	0.0063	0.0195	0.0028	0.0095	0.0068	0.0163	0.0032	0.0093	0.0072	0.0130
	(2,2)	0.0085	0.0189	0.0086	0.0180	0.0053	0.0098	0.0083	0.0179	0.0043	0.0095	0.0064	0.0152	0.0050	0.0092	0.0045	0.0121
	(3,3)	0.0055	0.0145	0.0138	0.0143	0.0113	0.0083	0.0130	0.0139	0.0116	0.0082	0.0138	0.0117	0.0114	0.0078	0.0147	0.0100
	(1,2)	0.1439	0.0499	0.1451	0.0475	0.1293	0.0331	0.1369	0.0453	0.1273	0.0323	0.1540	0.0373	0.1264	0.0314	0.1310	0.0330
(40,20,0.9)	(1,1)	0.0040	0.0084	0.0045	0.0084	0.0036	0.0045	0.0044	0.0074	0.0038	0.0045	0.0046	0.0066	0.0040	0.0044	0.0049	0.0060
	(2,2)	0.0059	0.0062	0.0059	0.0068	0.0044	0.0036	0.0057	0.0052	0.0037	0.0035	0.0059	0.0048	0.0038	0.0034	0.0060	0.0044
	(3,3)	0.0041	0.0048	0.0044	0.0049	0.0052	0.0036	0.0043	0.0046	0.0047	0.0035	0.0040	0.0045	0.0046	0.0034	0.0037	0.0043
	(1,2)	0.1961	0.0415	0.1581	0.0436	0.1476	0.0259	0.1460	0.0321	0.1468	0.0256	0.1551	0.0288	0.1454	0.0252	0.1461	0.0257
(60,20,0.9)	(1,1)	0.0012	0.0081	0.0010	0.0082	0.0009	0.0047	0.0010	0.0076	0.0010	0.0046	0.0011	0.0067	0.0009	0.0045	0.0009	0.0058
	(2,2)	0.0006	0.0081	0.0004	0.0080	0.0007	0.0054	0.0004	0.0079	0.0001	0.0053	0.0006	0.0071	0.0002	0.0052	0.0007	0.0063
	(3,3)	0.0010	0.0079	0.0001	0.0075	0.0004	0.0051	0.0001	0.0069	0.0006	0.0050	0.0003	0.0064	0.0003	0.0048	0.0004	0.0059
	(1,2)	0.1895	0.0462	0.1704	0.0450	0.1705	0.0251	0.1656	0.0422	0.1695	0.0243	0.1774	0.0353	0.1687	0.0240	0.1692	0.0279

Table 2: Average confidence/credible lengths and coverage percentages for estimates of R when λ is unknown.

(N, n, T)	C.S	AMLE		MLE		Prior 1		Prior 2		Prior 3	
		length	C.P	length	C.P	length	C.P	length	C.P	length	C.P
(40,10,0.7)	(1,1)	0.4374	0.8710	0.4197	0.8710	0.4061	0.9000	0.3922	0.9020	0.3791	0.9100
	(2,2)	0.4352	0.8760	0.4216	0.8750	0.4043	0.9020	0.3931	0.9070	0.3787	0.9080
	(3,3)	0.4543	0.8870	0.4270	0.8830	0.4117	0.9030	0.3956	0.9080	0.3813	0.9080
	(1,2)	0.4137	0.8940	0.4009	0.8900	0.3887	0.9010	0.3788	0.9040	0.3669	0.9060
(60,10,0.7)	(1,1)	0.4369	0.8810	0.4242	0.8860	0.4074	0.9120	0.3912	0.9130	0.3831	0.9140
	(2,2)	0.4366	0.8960	0.4221	0.8900	0.4064	0.9080	0.3915	0.9110	0.3806	0.9170
	(3,3)	0.4280	0.9090	0.4209	0.9050	0.4055	0.9200	0.3932	0.9210	0.3799	0.9260
	(1,2)	0.4329	0.9010	0.3987	0.9020	0.3764	0.9200	0.3661	0.9230	0.3605	0.9280
(40,20,0.7)	(1,1)	0.3090	0.9180	0.3055	0.9140	0.3001	0.9350	0.2926	0.9380	0.2888	0.9390
	(2,2)	0.3045	0.9290	0.3082	0.9230	0.3030	0.9340	0.2968	0.9360	0.2903	0.9400
	(3,3)	0.2989	0.9100	0.3148	0.9120	0.3099	0.9350	0.3028	0.9360	0.2947	0.9370
	(1,2)	0.2897	0.9340	0.2877	0.9340	0.2730	0.9360	0.2728	0.9380	0.2690	0.9390
(60,20,0.7)	(1,1)	0.3097	0.9240	0.3051	0.9270	0.2980	0.9310	0.2930	0.9310	0.2874	0.9330
	(2,2)	0.3065	0.9110	0.3049	0.9120	0.2983	0.9310	0.2912	0.9320	0.2888	0.9330
	(3,3)	0.3043	0.9280	0.3066	0.9230	0.3029	0.9360	0.2942	0.9370	0.2903	0.9390
	(1,2)	0.2893	0.9340	0.2807	0.9320	0.2697	0.9380	0.2614	0.9390	0.2599	0.9400
(40,10,0.9)	(1,1)	0.4370	0.8810	0.4135	0.8880	0.4019	0.9150	0.3864	0.9150	0.3780	0.9180
	(2,2)	0.4350	0.8850	0.4152	0.8880	0.4020	0.9150	0.3902	0.9160	0.3783	0.9170
	(3,3)	0.4313	0.8870	0.4250	0.8840	0.4078	0.9180	0.3948	0.9200	0.3810	0.9270
	(1,2)	0.4115	0.9000	0.3988	0.9050	0.3769	0.9150	0.3708	0.9200	0.3612	0.9210
(60,10,0.9)	(1,1)	0.4368	0.9020	0.4200	0.9080	0.4042	0.9290	0.3895	0.9300	0.3796	0.9330
	(2,2)	0.4350	0.8940	0.4219	0.8960	0.4039	0.9220	0.3906	0.9240	0.3804	0.9290
	(3,3)	0.4211	0.8900	0.4187	0.8940	0.4055	0.9210	0.3920	0.9220	0.3778	0.9270
	(1,2)	0.4319	0.9030	0.3957	0.9050	0.3717	0.9300	0.3659	0.9310	0.3567	0.9330
(40,20,0.9)	(1,1)	0.3077	0.9240	0.3040	0.9230	0.2973	0.9390	0.2920	0.9400	0.2848	0.9430
	(2,2)	0.3023	0.9320	0.3038	0.9320	0.2970	0.9390	0.2925	0.9450	0.2845	0.9480
	(3,3)	0.2909	0.9270	0.3028	0.9260	0.2963	0.9390	0.2916	0.9390	0.2844	0.9400
	(1,2)	0.2820	0.9290	0.2863	0.9210	0.2728	0.9420	0.2719	0.9460	0.2682	0.9470
(60,20,0.9)	(1,1)	0.3095	0.9360	0.3046	0.9350	0.2979	0.9400	0.2919	0.9420	0.2866	0.9490
	(2,2)	0.3065	0.9290	0.3029	0.9290	0.2977	0.9390	0.2873	0.9400	0.2859	0.9410
	(3,3)	0.2974	0.9340	0.3040	0.9360	0.2974	0.9420	0.2904	0.9440	0.2857	0.9500
	(1,2)	0.2823	0.9230	0.2767	0.9280	0.2598	0.9390	0.2562	0.9400	0.2558	0.9410

To monitor the convergence of the MCMC method, in the first and third cases, we studied the trace plots for various censoring schemes and parameters. In all cases, the trace plots indicated that the MCMC method is converged. Some of these plots are displayed in Figures 2–5. It is notable that Figures 2 and 3 have considered the problem in the first case (when the common second shape parameter is unknown), and Figures 4 and 5 have considered the problem in the third case (when all parameters are different and unknown), respectively.

Due to the information of Table 1, we observed that the Bayes estimates have the minimum value of MSEs. Also, in Bayesian inference, the informative priors performance was better than non-informative ones and the best performance, in terms of MSE, was belonged to Prior 3. Furthermore, the MCMC method performs better, in comparison with Lindley's approximation. From Table 2, we observed that the HPD credible intervals indicated a better performance compared to the asymptotic confidence intervals. Also, in Bayesian inference, the best performance belonged to Prior 3, namely, the HPD credible intervals based on Prior 3, have the smallest average lengths and largest coverage percentages.

Table 3: Biases, MSEs, Average confidence/credible lengths and coverage percentages for estimates of R when λ is known.

(N,n,T)	C.S	MLE		Asymp. C.I		Prior 4				Prior 5				Prior 6			
						Est.		C.I		Est.		C.I		Est.		C.I	
		|Bias|	MSE	length	C.P	|Bias|	MSE	length	C.P	|Bias|	MSE	length	C.P	|Bias|	MSE	length	C.P
(40,10,0.7)	(1,1)	0.0010	0.0126	0.4189	0.8730	0.0010	0.0109	0.4067	0.9230	0.0008	0.0067	0.3946	0.9240	0.0007	0.0044	0.3822	0.9290
	(2,2)	0.0054	0.0115	0.4171	0.8900	0.0052	0.0105	0.4051	0.9190	0.0040	0.0061	0.3944	0.9250	0.0032	0.0040	0.3819	0.9290
	(3,3)	0.0013	0.0127	0.4197	0.8810	0.0012	0.0117	0.4072	0.9020	0.0008	0.0068	0.3964	0.9050	0.0006	0.0045	0.3825	0.9050
	(1,2)	0.1449	0.0328	0.3819	0.8970	0.1393	0.0315	0.3759	0.9010	0.1117	0.0252	0.3758	0.9030	0.0936	0.0206	0.3687	0.9040
(60,10,0.7)	(1,1)	0.0058	0.0128	0.4190	0.9010	0.0056	0.0117	0.4067	0.9250	0.0045	0.0069	0.3958	0.9260	0.0038	0.0046	0.3816	0.9290
	(2,2)	0.0008	0.0121	0.4175	0.9040	0.0007	0.0111	0.4054	0.9140	0.0005	0.0064	0.3948	0.9250	0.0004	0.0042	0.3818	0.9290
	(3,3)	0.0018	0.0125	0.4177	0.9030	0.0018	0.0115	0.4058	0.9150	0.0014	0.0067	0.3946	0.9220	0.0012	0.0044	0.3818	0.9250
	(1,2)	0.1535	0.0394	0.3805	0.9090	0.1477	0.0379	0.3750	0.9150	0.1188	0.0306	0.3749	0.9180	0.0997	0.0253	0.3676	0.9220
(40,20,0.7)	(1,1)	0.0003	0.0059	0.3024	0.9170	0.0003	0.0056	0.2973	0.9320	0.0002	0.0042	0.2927	0.9340	0.0002	0.0033	0.2864	0.9360
	(2,2)	0.0017	0.0058	0.3079	0.9190	0.0017	0.0056	0.3028	0.9350	0.0014	0.0042	0.2973	0.9360	0.0013	0.0033	0.2916	0.9370
	(3,3)	0.0016	0.0065	0.3174	0.9180	0.0016	0.0062	0.3112	0.9320	0.0013	0.0046	0.3053	0.9350	0.0012	0.0036	0.2988	0.9370
	(1,2)	0.1477	0.0327	0.2802	0.9170	0.1447	0.0304	0.2778	0.9310	0.1280	0.0252	0.2771	0.9330	0.1149	0.0140	0.2748	0.9340
(60,20,0.7)	(1,1)	0.0001	0.0064	0.3020	0.9320	0.0001	0.0061	0.2975	0.9340	0.0006	0.0046	0.2920	0.9350	0.0003	0.0036	0.2867	0.9380
	(2,2)	0.0042	0.0064	0.3030	0.9310	0.0041	0.0061	0.2980	0.9330	0.0036	0.0046	0.2926	0.9340	0.0032	0.0036	0.2873	0.9340
	(3,3)	0.0007	0.0057	0.3070	0.9180	0.0007	0.0055	0.3016	0.9300	0.0006	0.0041	0.2966	0.9320	0.0005	0.0032	0.2900	0.9320
	(1,2)	0.1814	0.0356	0.2626	0.9300	0.1179	0.0331	0.2623	0.9350	0.1592	0.0216	0.2611	0.9360	0.1442	0.0154	0.2600	0.9370
(40,10,0.9)	(1,1)	0.0051	0.0119	0.4154	0.8900	0.0049	0.0116	0.4037	0.9150	0.0036	0.0064	0.3935	0.9190	0.0028	0.0043	0.3816	0.9210
	(2,2)	0.0019	0.0104	0.4160	0.9030	0.0019	0.0095	0.4044	0.9240	0.0013	0.0055	0.3935	0.9250	0.0010	0.0037	0.3814	0.9270
	(3,3)	0.0118	0.0120	0.4173	0.9000	0.0113	0.0110	0.4057	0.9340	0.0086	0.0064	0.3945	0.9350	0.0069	0.0042	0.3822	0.9390
	(1,2)	0.1462	0.0322	0.3790	0.9050	0.1406	0.0299	0.3736	0.9250	0.1127	0.0198	0.3733	0.9290	0.0944	0.0159	0.3675	0.9300
(60,10,0.9)	(1,1)	0.0017	0.0114	0.4182	0.8920	0.0016	0.0105	0.4060	0.9240	0.0011	0.0061	0.3951	0.9260	0.0008	0.0040	0.3815	0.9270
	(2,2)	0.0040	0.0118	0.4170	0.9050	0.0038	0.0108	0.4042	0.9240	0.0028	0.0063	0.3945	0.9250	0.0022	0.0042	0.3817	0.9290
	(3,3)	0.0031	0.0116	0.4173	0.8850	0.0030	0.0106	0.4053	0.9230	0.0025	0.0062	0.3942	0.9240	0.0021	0.0041	0.3817	0.9250
	(1,2)	0.1497	0.0393	0.3766	0.9100	0.1440	0.0378	0.3707	0.9330	0.1156	0.0304	0.3700	0.9340	0.0969	0.0251	0.3664	0.9370
(40,20,0.9)	(1,1)	0.0012	0.0058	0.3019	0.9310	0.0012	0.0056	0.2971	0.9410	0.0010	0.0042	0.2918	0.9440	0.0008	0.0033	0.2862	0.9470
	(2,2)	0.0010	0.0057	0.3028	0.9210	0.0010	0.0055	0.2981	0.9430	0.0008	0.0041	0.2925	0.9440	0.0007	0.0032	0.2869	0.9490
	(3,3)	0.0008	0.0047	0.3031	0.9270	0.0008	0.0045	0.2981	0.9410	0.0002	0.0034	0.2925	0.9430	0.0003	0.0026	0.2870	0.9450
	(1,2)	0.1620	0.0263	0.2685	0.9260	0.1588	0.0253	0.2683	0.9410	0.1415	0.0193	0.2674	0.9430	0.1278	0.0137	0.2669	0.9460
(60,20,0.9)	(1,1)	0.006	0.0063	0.3020	0.9280	0.0005	0.0061	0.2970	0.9440	0.0005	0.0045	0.2919	0.9470	0.0005	0.0035	0.2867	0.9480
	(2,2)	0.0025	0.0057	0.3020	0.9300	0.0024	0.0055	0.2970	0.9440	0.0021	0.0041	0.2921	0.9480	0.0019	0.0032	0.2866	0.9490
	(3,3)	0.0052	0.0055	0.3021	0.9280	0.0051	0.0053	0.2969	0.9430	0.0044	0.0040	0.2922	0.9460	0.0038	0.0031	0.2866	0.9500
	(1,2)	0.1818	0.0346	0.2619	0.9300	0.1782	0.0321	0.2612	0.9460	0.1596	0.0209	0.2606	0.9460	0.1447	0.0148	0.2593	0.9530

Table 4: Biases, MSEs, Average credible lengths and coverage percentages for estimates of R in general case.

(N,n,T)	C.S	AMLE		MLE		Prior 7				Prior 8				Prior 9			
						Est.		C.I		Est.		C.I		Est.		C.I	
		|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	length	C.P	|Bias|	MSE	length	C.P	|Bias|	MSE	length	C.P
(40,10,0.7)	(1,1)	0.0129	0.0234	0.0169	0.0211	0.0140	0.0111	0.4083	0.9000	0.0111	0.0080	0.3933	0.9010	0.0100	0.0060	0.3772	0.9060
	(2,2)	0.0009	0.0183	0.0033	0.0169	0.0007	0.0078	0.4114	0.9120	0.0001	0.0056	0.3948	0.9170	0.0007	0.0081	0.3825	0.9190
	(3,3)	0.0169	0.0182	0.0168	0.0194	0.0152	0.0115	0.4053	0.9000	0.0139	0.0080	0.3944	0.9060	0.0120	0.0060	0.3805	0.9100
	(1,2)	0.1954	0.0464	0.1940	0.0494	0.1395	0.0320	0.3801	0.9110	0.1210	0.0279	0.3729	0.9130	0.1073	0.0245	0.3629	0.9180
(60,10,0.7)	(1,1)	0.0316	0.0171	0.0323	0.0177	0.0095	0.0090	0.4095	0.9190	0.0077	0.0065	0.3948	0.9220	0.0064	0.0048	0.3824	0.9240
	(2,2)	0.0101	0.0223	0.0108	0.0231	0.0046	0.0110	0.4044	0.9050	0.0037	0.0079	0.3936	0.9060	0.0035	0.0060	0.3798	0.9120
	(3,3)	0.0081	0.0152	0.0084	0.0159	0.0073	0.0089	0.4080	0.9220	0.0067	0.0064	0.3905	0.9270	0.0064	0.0048	0.3799	0.9280
	(1,2)	0.2451	0.0639	0.2334	0.0662	0.1413	0.0361	0.3793	0.9200	0.1228	0.0337	0.3746	0.9240	0.1103	0.0297	0.3639	0.9260
(40,20,0.7)	(1,1)	0.0116	0.0077	0.0181	0.0079	0.0093	0.0065	0.2980	0.9370	0.0087	0.0056	0.2912	0.9390	0.0079	0.0047	0.2856	0.9400
	(2,2)	0.0165	0.0135	0.0146	0.0134	0.0041	0.0053	0.3020	0.9330	0.0037	0.0045	0.2975	0.9350	0.0032	0.0038	0.2889	0.9390
	(3,3)	0.0099	0.0171	0.0157	0.0158	0.0139	0.0049	0.3127	0.9320	0.0125	0.0040	0.3039	0.9330	0.0113	0.0034	0.2963	0.9380
	(1,2)	0.1755	0.0374	0.1715	0.0371	0.1583	0.0313	0.2686	0.9340	0.1474	0.0273	0.2666	0.9360	0.1377	0.0239	0.2632	0.9400
(60,20,0.7)	(1,1)	0.0116	0.0098	0.0103	0.0090	0.0074	0.0056	0.2984	0.9350	0.0071	0.0047	0.2930	0.9370	0.0060	0.0040	0.2874	0.9390
	(2,2)	0.0133	0.0105	0.0134	0.0102	0.0085	0.0073	0.2996	0.9320	0.0078	0.0060	0.2920	0.9350	0.0068	0.0051	0.2867	0.9380
	(3,3)	0.0180	0.0130	0.0138	0.0132	0.0127	0.0052	0.3025	0.9340	0.0114	0.0044	0.2950	0.9380	0.0107	0.0037	0.2886	0.9400
	(1,2)	0.2716	0.0517	0.2179	0.0522	0.1801	0.0358	0.2606	0.9310	0.1682	0.0317	0.2600	0.9360	0.1575	0.0282	0.2589	0.9380
(40,10,0.9)	(1,1)	0.0072	0.0121	0.0066	0.0160	0.0076	0.0090	0.4048	0.9230	0.0067	0.0066	0.3922	0.9240	0.0055	0.0050	0.3705	0.9290
	(2,2)	0.0007	0.0137	0.0023	0.0163	0.0006	0.0076	0.4076	0.9150	0.0001	0.0048	0.3940	0.9190	0.0007	0.0051	0.3821	0.9290
	(3,3)	0.0169	0.0160	0.0098	0.0167	0.0123	0.0105	0.4050	0.9210	0.0104	0.0074	0.3935	0.9220	0.0088	0.0054	0.3800	0.9290
	(1,2)	0.1497	0.0362	0.1893	0.0345	0.1285	0.0281	0.3719	0.9330	0.1108	0.0212	0.3532	0.9350	0.0980	0.0167	0.3464	0.9390
(60,10,0.9)	(1,1)	0.0160	0.0138	0.0155	0.0130	0.0039	0.0089	0.4062	0.9310	0.0022	0.0064	0.3946	0.9370	0.0023	0.0040	0.3811	0.9390
	(2,2)	0.0041	0.0220	0.0051	0.0221	0.0014	0.0104	0.4035	0.9300	0.0028	0.0073	0.3930	0.9360	0.0021	0.0051	0.3781	0.9360
	(3,3)	0.0077	0.0142	0.0081	0.0148	0.0073	0.0071	0.4030	0.9210	0.0063	0.0058	0.3811	0.9260	0.0042	0.0039	0.3774	0.9280
	(1,2)	0.2362	0.0636	0.2134	0.0630	0.1299	0.0316	0.3740	0.9150	0.1120	0.0245	0.3668	0.9170	0.0996	0.0198	0.3602	0.9190
(40,20,0.9)	(1,1)	0.0011	0.0073	0.0018	0.0064	0.0004	0.0052	0.2957	0.9400	0.0009	0.0043	0.2901	0.9430	0.0004	0.0037	0.2843	0.9460
	(2,2)	0.0154	0.0107	0.0014	0.0106	0.0029	0.0038	0.2992	0.9390	0.0025	0.0031	0.2919	0.9400	0.0023	0.0026	0.2862	0.9440
	(3,3)	0.0044	0.0146	0.0050	0.0140	0.0060	0.0031	0.2987	0.9410	0.0057	0.0025	0.2931	0.9440	0.0055	0.0021	0.2874	0.9470
	(1,2)	0.1634	0.0339	0.1486	0.0301	0.1559	0.0271	0.2677	0.9400	0.1452	0.0202	0.2654	0.9450	0.1358	0.0158	0.2630	0.9480
(60,20,0.9)	(1,1)	0.0042	0.0089	0.0049	0.0087	0.0043	0.0051	0.2973	0.9420	0.0040	0.0043	0.2923	0.9430	0.0041	0.0036	0.2848	0.9450
	(2,2)	0.0074	0.0082	0.0125	0.0073	0.0071	0.0050	0.2956	0.9400	0.0070	0.0043	0.2892	0.9450	0.0064	0.0036	0.2844	0.9470
	(3,3)	0.0024	0.0079	0.0001	0.0073	0.0001	0.0050	0.2985	0.9390	0.0007	0.0042	0.2923	0.9430	0.0007	0.0035	0.2866	0.9460
	(1,2)	0.2180	0.0514	0.2179	0.0509	0.1729	0.0271	0.2598	0.9400	0.1615	0.0203	0.2577	0.9440	0.1517	0.0161	0.2574	0.9520

As shown in Table 3, we observed that the Bayes estimates have the minimum value of MSEs. Also, in Bayesian inference, the informative priors performed better than non-informative ones and the best performance, in terms of MSE, was belonged to Prior 6. Moreover, we observed that the Bayesian credible intervals have the better performance, in comparison with the asymptotic confidence intervals. Also, in Bayesian inference, the best performance belonged to Prior 6, namely, the Bayesian credible intervals based on Prior 6 have the smallest average lengths and largest coverage percentages.

As we observe from Table 4, the Bayes estimates have the minimum value of MSEs. Also, in Bayesian inference, the informative priors perform better than non-informative ones and the best performance, in terms of MSE, was belonged to Prior 9. Moreover, we observed that HPD credible intervals based on informative priors, indicated better performance compared to non-informative ones.

To tell the truth, from Tables 1, 3 and 4, along by increasing n for fixed N and T, and also with increasing T for fixed N and n, the MSEs of all estimates decrease in all cases. This can be due to the fact in both of the above mentioned cases, some additional information is gathered. Moreover, from Tables 2, 3 and 4, with increasing n for fixed N and T, and also with increasing T for fixed N and n, the average confidence lengths decrease and the associated coverage percentages increase, in all cases.

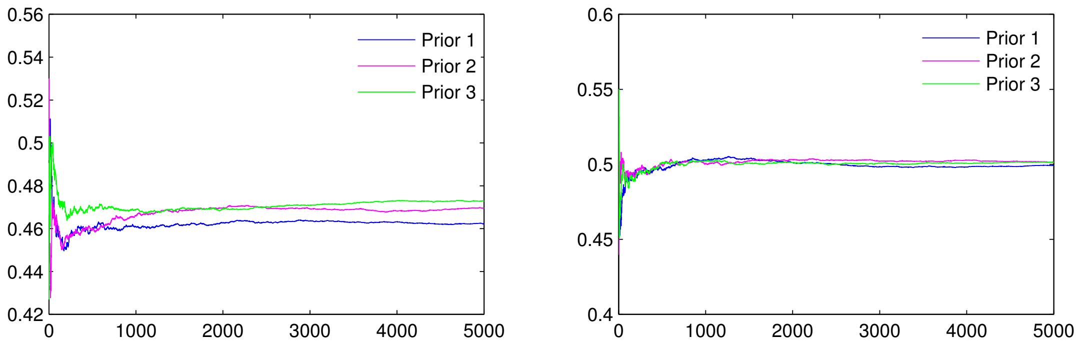

Figure 2: Trace plots with C.S $(1, 1)$ (left) and $(3, 3)$ (right), for $(N, n, T) = (40, 10, 0.7)$, in common shape parameter λ.

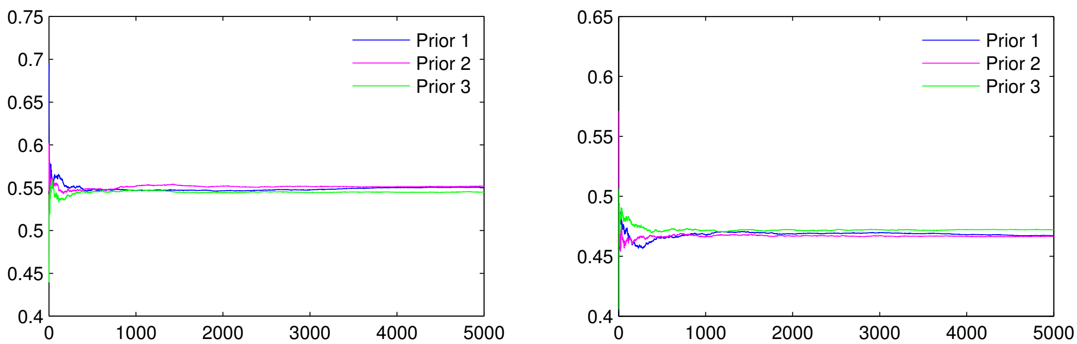

Figure 3: Trace plots with C.S $(2, 2)$ (left) and $(3, 3)$ (right), for $(N, n, T) = (60, 20, 0.9)$, in common shape parameter λ.

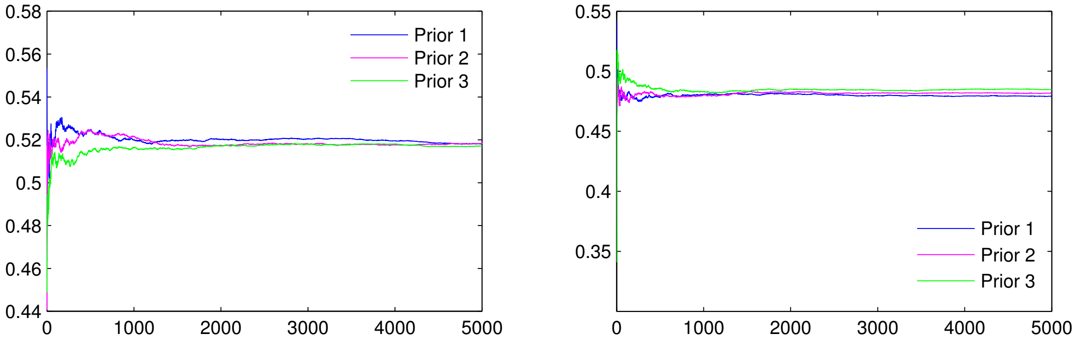

Figure 4: Trace plots with C.S $(1, 3)$ (left) and $(1, 1)$ (right), for $(N, n, T) = (40, 20, 0.7)$, in general case.

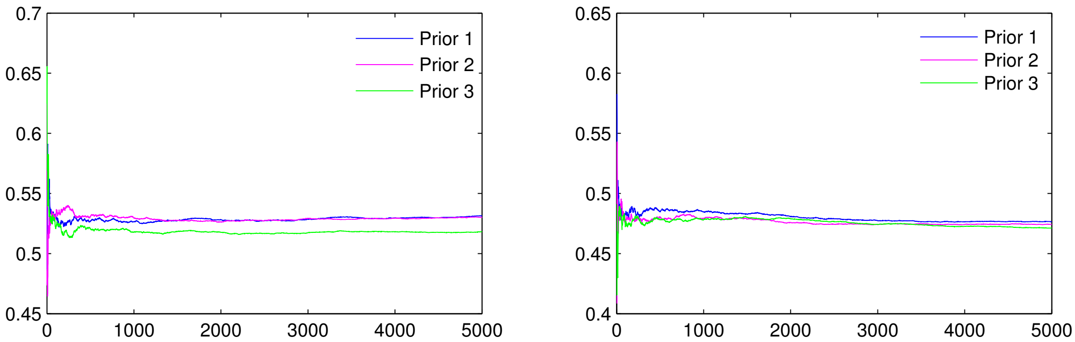

Figure 5: Trace plots with C.S $(2, 3)$ (left) and $(1, 1)$ (right), for $(N, n, T) = (60, 10, 0.9)$, in general case.

5.2. Data analysis

In this section, we analyze two pair of real data set for illustrative proposes.

Example 5.1. In the first example, we use the monthly water capacity of the Shasta reservoir in California, USA, see data in http://cdec.water.ca.gov/cgi-progs/queryMonthly?SHA. Some authors such as Sultana *et al.* [25], Kohansal [9], Kizilaslan and Nadar [8], [6] and Nadar *et al.* [19] have been studied this data, previously. From this data, we construct one scenario relating to the excessive drought. In fact, we contract that if the average water capacity in July and August of a same year is more than the water capacity in December, the excessive drought will not occur. With respect to this scenario, we consider the months July, August, and December from 1987 to 2016. So, $X_1, ..., X_{30}$ are the capacity of December and $Y_1, ..., Y_{30}$ are the average capacity of July and August from 1987 to 2016, respectively, and $R = P(X < Y)$ is the probability of non-occurrence of drought. As the range of KuD is $(0, 1)$, all data have been divided by the total capacity of Shasta reservoir, 4552000 acre-feet. This work does not make any change in statistical inference.

At first, we check that the KuD can separately analyze these data sets or not. To fit the KuD, we obtain the initial guess, in the Newton–Raphson method, by using the profile log-likelihood functions, which were indicated in Figure 6. So, we start this method by the starting values 3.45 and 3.65, for X and Y, respectively. By fitting the KuD, for X, $\widehat{\alpha}$, $\widehat{\lambda}$, the Kolmogorov–Smirnov distance and the corresponding p-value are 4.1903, 3.5000, 0.1592 and 0.3916, respectively. Also, for Y, $\widehat{\beta}$, $\widehat{\lambda}$, the Kolmogorov–Smirnov distance and the associated p-value are 3.7828, 3.7700, 0.1218 and 0.7195, respectively. In terms of the p-values, we identify that the KuD provides suitable fits for the data sets. Figures 7 and 8 indicated the empirical distribution functions, PP-plots, and PP-plots with simulated envelope, for X and Y, respectively.

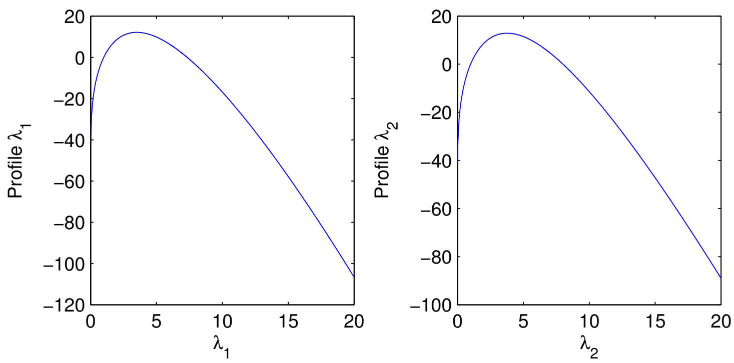

Figure 6: Profile log-likelihood function of λ for X (left) Y (right).

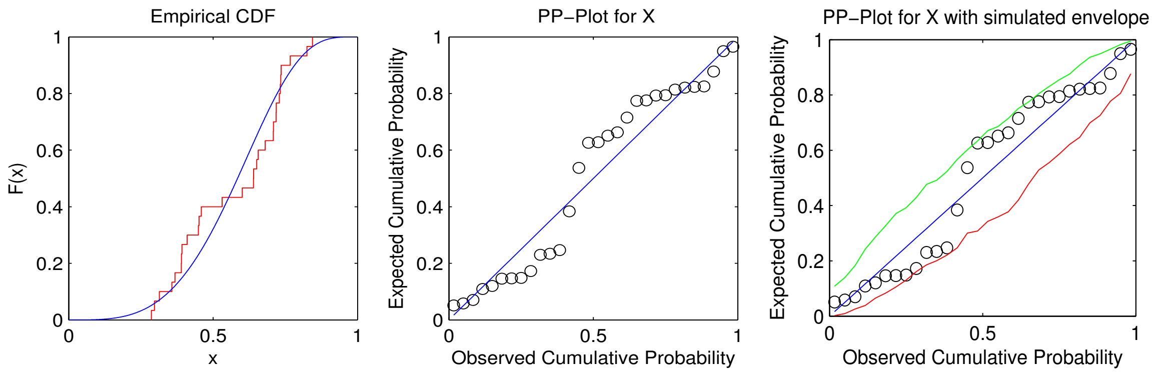

Figure 7: Empirical distribution function (left), PP-plot (center) and PP-plots with simulated envelope (right) for X.

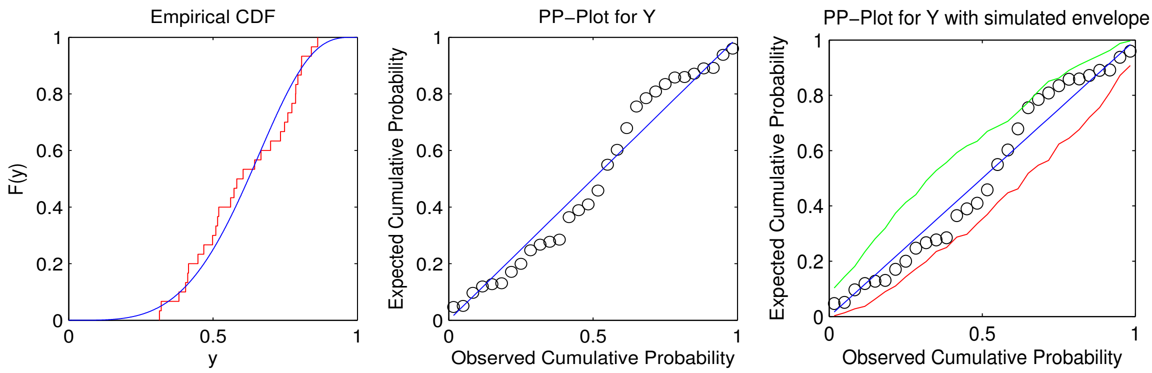

Figure 8: Empirical distribution functions (left)PP-plot (center) and PP plots with simulated envelope (right) for Y.

For the illustrative proposes, we consider two different HP censoring schemes for X and Y as follows:

$$\text{Scheme 1:} \ [1^{*10}, 0^{*10}], \ T_1 = T_2 = 0.9,$$
$$\text{Scheme 2:} \ [2^{*10}], \qquad T_1 = T_2 = 0.5.$$

In the first case, when the common shape parameter λ is unknown, for complete data sets, and Schemes 1 and 2, we obtained the ML, AML and Bayes estimates of R with non-informative priors assumption, i.e., $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$ by applying Lindley's approximation and MCMC method. Also, we derived the 95% asymptotic and HPD intervals. The results are listed in Table 5.

Table 5: The ML, AML, Bayes estimates and different confidence/credible intervals of R, in Example 5.1.

		MLE	Asymp. (MLE)	AMLE	Asymp. (AMLE)	Bayes		HPD
						MCMC	Lindley	
λ	Complete	0.5522	(0.4268,0.6776)	0.5641	(0.4403,0.6879)	0.5520	0.5511	(0.4258,0.6707)
	Scheme 1	0.5520	(0.3983,0.7057)	0.5369	(0.3865,0.6927)	0.5523	0.5503	(0.3985,0.7036)
	Scheme 2	0.5723	(0.3563,0.7882)	0.5200	(0.3013,0.7388)	0.5727	0.5673	(0.3530,0.7687)
λ_1, λ_2	Complete	0.5617	—	0.5971	—	0.5647	—	(0.4372,0.6848)
	Scheme 1	0.5533	—	0.5593	—	0.5534	—	(0.3974,0.7027)
	Scheme 2	0.5777	—	0.4899	—	0.5779	—	(0.3501,0.7657)

As we observe, the second shape parameters of two data sets are not exactly same. As a result, in the second case, when the shape parameters λ_1 and λ_2 are different and unknown, for complete data sets, Schemes 1 and 2, we obtained the ML, AML and Bayes estimates of R with non-informative priors assumption, i.e., $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = a_4 = b_4 = 0$, respectively. Also, we derived 95% HPD credible intervals. Theses results are presented in Table 5. By comparing the two schemes, we observed that estimators have smaller standard errors in Scheme 1, compared to Scheme 2, as it was expected. It is notable that the estimation methods which presented a better performance in the simulations are more reliable than the others. So, the results based on the Bayesian estimations and in Bayesian estimation the results obtained by the MCMC method are more preferred, in comparison with the others. Also, we would like to use the HPD credible intervals as the best intervals.

Example 5.2. In the second example, we use the lifetime data for insulation specimens. The length of time was observed until each specimen failed or "broke down". Also, the results for seven groups of specimens, tested at voltages ranging from 26 to 38 kilovolts (kV) were presented. We consider the data sets for 34 kV and 36 kV, reported in Lawless [15], as the strength and stress variables, respectively. Therefore, the parameter $R = P(X < Y)$ can be investigated as the probability of insulation resistance. For the same reason as it was earlier explained in Example 5.1, we have converted all data between 0 and 1. Recently, Kizilaslan and Nadar [7] considered this data set.

At first, we must check that the KuD can analyze these data sets, separately. By fitting the KuD, for X, $\widehat{\alpha}$, $\widehat{\lambda}$, the Kolmogorov–Smirnov distance and the corresponding p-value are

9.7733, 0.84, 0.2103 and 0.4592, respectively. Also, for Y, $\widehat{\beta}$, $\widehat{\lambda}$, the Kolmogorov–Smirnov distance and the associated p-value are 0.8963, 0.3736, 0.2756 and 0.0911, respectively. In terms of the p-values, we observe that the KuD provides suitable fits for the data sets.

For the illustrative proposes, we consider the HP censoring scheme as Scheme 3: $[1^{*5}, 0^{*5}]$, $T_1 = 0.1$ and $[1^{*9}, 0^{*1}]$, $T_2 = 0.2$ for X and Y, respectively.

In the first case, when the common shape parameter λ is unknown, for complete data sets and Scheme 3, we obtained the ML, AML, and Bayes estimates of R with non-informative priors assumption, i.e., $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = 0$ by applying Lindley's approximation and MCMC method. Also, we derived the 95% asymptotic and HPD intervals. These obtained results are listed in Table 6.

As indicated, the second shape parameters of two data sets are not similar. So, when the shape parameters λ_1 and λ_2 are different and unknown, for complete data sets, Schemes 1 and 2, we obtained the ML, AML and Bayes estimates of R with non-informative priors assumption, i.e., $a_1 = b_1 = a_2 = b_2 = a_3 = b_3 = a_4 = b_4 = 0$. Also, we derived 95% HPD credible intervals. These results are given in Table 6.

Table 6: The ML, AML, Bayes estimates and different confidence/credible intervals of R, in Example 5.2.

		MLE	Asymp. (MLE)	AMLE	Asymp. (AMLE)	Bayes		HPD
						MCMC	Lindley	
λ	Complete	0.8007	(0.6763,0.9252)	0.7034	(0.5944,0.8619)	0.8016	0.7892	(0.6798,0.8938)
	Scheme 3	0.6368	(0.4151,0.8614)	0.6739	(0.4131,0.9048)	0.6326	0.6273	(0.3851,0.8183)
λ_1, λ_2	Complete	0.7127	—	0.6058	—	0.7252	—	(0.5979,0.8360)
	Scheme 3	0.6371	—	0.6760	—	0.6351	—	(0.3989,0.8234)

To see a motivation based on real data set that presents the need for the new methodology, we consider the progressive scheme, one of the most applicable censoring scheme, for this data set. Comparison between two methodologies (HP and progressive schemes) is performed by obtaining the values of Akaike information criterion (AIC), Bayesian information criterion (BIC) and Hannan–Quinn information criterion (HQC). We have shown the results in Table 7. From Table 7, by ignoring minor differences, we see that the new methodology (results based on HP scheme) is better than the previous one (results based on the progressive scheme.)

Table 7: AIC, BIC and HQC in comparison of two methodology, in Example 5.2.

		HP				Progressive			
		MLE	AMLE	Lindley	MCMC	MLE	AMLE	Lindley	MCMC
λ	AIC	−42.9761	−37.8133	−42.9575	−42.9762	−42.0119	−37.5979	−42.0107	−42.0119
	BIC	−40.4765	−35.3137	−40.4579	−40.4765	−39.0247	−34.6107	−39.0235	−39.0249
	HQC	−42.7276	−37.5648	−42.7090	−42.7277	−41.4288	−37.0148	−41.4276	−41.4289
λ_1, λ_2	AIC	−41.0665	−37.4275	—	−41.0906	−40.0246	−35.5724	—	−40.0373
	BIC	−37.7337	−34.0946	—	−37.7678	−36.0417	−31.5895	—	−36.0544
	HQC	−40.7352	−37.0962	—	−40.7694	−39.2471	−34.7949	—	−39.2597

6. CONCLUSION

In this paper, we obtain different estimates of the stress-strength parameter, under the hybrid progressive censored scheme, at the time that stress and strength are considered as two independent Kumaraswamy random variables. The problem is going to be solved in three cases. First, when $X \sim \mathrm{Ku}(\alpha, \lambda)$ and $Y \sim \mathrm{Ku}(\beta, \lambda)$, we derive ML, AML and two approximated Bayes estimates by applying Lindley's approximation and MCMC method, due to the lack of explicit forms. Also, we consider the existence and uniqueness of the MLE and construct the asymptotic and HPD intervals for R. Second, when the common second shape parameter, λ, is known, we obtain the MLE and exact Bayes estimate of R. Third, in general case, when $X \sim \mathrm{Ku}(\alpha, \lambda_1)$ and $Y \sim \mathrm{Ku}(\beta, \lambda_2)$, we provide ML, AML and Bayesian inferences of R, respectively.

From the simulation results, which were obtained using the Monte Carlo method, in point estimates, we observed that the Bayes estimates have the minimum value of MSEs. Also, in Bayesian inference, the informative priors perform better than non-informative ones. Furthermore, the MCMC method performs better than Lindley's approximation. In interval estimates, we observed that the HPD credible intervals have a better performance in comparison with the asymptotic confidence intervals. Also, in Bayesian inference, the HPD credible intervals based on informative priors have the smallest average lengths and largest coverage percentages.

A. APPENDIX

Proof of Theorem 2.1: By a simple method, we can rewrite $G(\lambda)$ as:

$$G(\lambda) = \frac{J_1}{\lambda} + G_1(\lambda) + J_1\frac{G_2(\lambda)}{G_3(\lambda)} + \frac{J_2}{\lambda} + H_1(\lambda) + J_2\frac{H_2(\lambda)}{H_3(\lambda)},$$

where

$$\begin{aligned}
G_1(\lambda) &= \sum_{i=1}^{J_1} \frac{\log(x_i)}{1-x_i^\lambda}, \quad G_2(\lambda) = \sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\frac{\log(x_i)}{1-x_i^\lambda} + R_{J_1}^* T_1^\lambda \frac{\log(T_1)}{1-T_1^\lambda},\\
G_3(\lambda) &= \sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^\lambda) + R_{J_1}^*\log(1-T_1^\lambda),\\
H_1(\lambda) &= \sum_{j=1}^{J_2} \frac{\log(y_j)}{1-y_j^\lambda}, \quad H_2(\lambda) = \sum_{j=1}^{J_2}(S_j+1)y_j^\lambda\frac{\log(y_j)}{1-y_j^\lambda} + S_{J_2}^* T_2^\lambda \frac{\log(T_2)}{1-T_2^\lambda},\\
H_3(\lambda) &= \sum_{j=1}^{J_2}(S_j+1)\log(1-y_j^\lambda) + S_{J_2}^*\log(1-T_2^\lambda).
\end{aligned}$$

We observe that $\lim_{\lambda\to 0^+} G(\lambda) = \infty$ and $\lim_{\lambda\to\infty} G(\lambda) < 0$. Consequently, $G(\lambda)$ has at least one root in $(0, \infty)$ by the intermediate value theorem. So, it is enough to show that $G'(\lambda) < 0$.

We can obtain $G'(\lambda)$, after accomplishing some steps, as:

$$\begin{aligned}
G'(\lambda) &= -\frac{1}{\lambda^2}\left\{G_4(\lambda) - J_1\frac{G_3(\lambda)G_5(\lambda) + \big(G_2(\lambda)\big)^2}{\big(G_3(\lambda)\big)^2}\right\} \\
&\quad -\frac{1}{\lambda^2}\left\{H_4(\lambda) - J_2\frac{H_3(\lambda)H_5(\lambda) + \big(H_2(\lambda)\big)^2}{\big(H_3(\lambda)\big)^2}\right\},
\end{aligned}$$

where

$$\begin{aligned}
G_4(\lambda) &= J_1 - \sum_{i=1}^{J_1} x_i^\lambda\left(\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)^2, \quad H_4(\lambda) = J_2 - \sum_{j=1}^{J_2} y_j^\lambda\left(\frac{\log(y_j^\lambda)}{1-y_j^\lambda}\right)^2, \\
G_5(\lambda) &= \sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\left(\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)^2 + R_{J_1}^* T_1^\lambda\left(\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2, \\
H_5(\lambda) &= \sum_{j=1}^{J_2}(S_j+1)y_j^\lambda\left(\frac{\log(y_j^\lambda)}{1-y_j^\lambda}\right)^2 + S_{J_2}^* T_2^\lambda\left(\frac{\log(T_2^\lambda)}{1-T_2^\lambda}\right)^2.
\end{aligned}$$

It can be observed that $G_4(\lambda) > 0$, as $f(x) = x\big(\frac{\log(x)}{1-x}\big)^2$, so $f(x) < 1$ for $x \in (0,1)$. Moreover,

$$\begin{aligned}
(G_2(\lambda))^2 &= \left(\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)^2 + \left(R_{J_1}^* T_1^\lambda\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2 \\
&\quad + 2\left(\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)\left(R_{J_1}^* T_1^\lambda\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right) \\
&\le \left(\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\right)\left(\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\left(\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)^2\right) + \left(R_{J_1}^* T_1^\lambda\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2 \\
&\quad + \sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\left(R_{J_1}^* T_1^\lambda\left(\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2\right) + \sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\big(R_{J_1}^* T_1^\lambda\big) \\
&\le \left(-\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\log(1-x_i^\lambda)\right)\left(\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\left(\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)^2\right) \\
&\quad + R_{J_1}^* T_1^\lambda\left(\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2\big(-R_{J_1}^*\log(1-T_1^\lambda)\big) \\
&\quad - \sum_{i=1}^{J_1}(R_i+1)\log(1-x_i^\lambda)\left(R_{J_1}^* T_1^\lambda\left(\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2\right) \\
&\quad + \sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\big(-R_{J_1}^*\log(1-T_1^\lambda)\big) \\
&= \left[-\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\log(1-x_i^\lambda) - R_{J_1}^*\log(1-T_1^\lambda)\right] \\
&\quad \times\left[\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\left(\frac{\log(x_i^\lambda)}{1-x_i^\lambda}\right)^2 + R_{J_1}^* T_1^\lambda\left(\frac{\log(T_1^\lambda)}{1-T_1^\lambda}\right)^2\right] = -G_3(\lambda)G_5(\lambda).
\end{aligned}$$

The above equations have been obtained by applying the Cauchy–Schwarz inequality and $x < -\log(1-x)$, $x \in (0,1)$. Consequently, $G'(\lambda) < 0$ and the proof is completed. □

B. APPENDIX

We compute $\tilde{\mu}_1$, $\tilde{\mu}_2$ and $\tilde{\sigma}$ at

$$
\begin{aligned}
A_1 &= \frac{\sum_{i=1}^{J_1}(R_i+1)\beta_i u_i + R^*_{J_1}\beta^*_{J_1}a_1}{\sum_{i=1}^{J_1}(R_i+1)\beta_i + R^*_{J_1}\beta^*_{J_1}}, \quad B_1 = \frac{\sum_{i=1}^{J_1}\alpha_i - \sum_{i=1}^{J_1}R_i(1-\alpha_i) - R^*_{J_1}(1-\alpha^*_{J_1})}{\sum_{i=1}^{J_1}(R_i+1)\beta_i + R^*_{J_1}\beta^*_{J_1}},\\
A_2 &= \frac{\sum_{j=1}^{J_2}(S_j+1)\bar{\beta}_j v_j + S^*_{J_2}\bar{\beta}^*_{J_2}a_2}{\sum_{j=1}^{J_2}(S_j+1)\bar{\beta}_j + S^*_{J_2}\bar{\beta}^*_{J_2}}, \quad B_2 = \frac{\sum_{j=1}^{J_2}\bar{\alpha}_j - \sum_{j=1}^{J_2}S_j(1-\bar{\alpha}_j) - S^*_{J_2}(1-\bar{\alpha}^*_{J_2})}{\sum_{j=1}^{J_2}(S_j+1)\bar{\beta}_j + S^*_{J_2}\bar{\beta}^*_{J_2}},\\
D_1 &= \sum_{i=1}^{J_1}\alpha_i u_i - A_1B_1\left(\sum_{i=1}^{J_1}(R_i+1)\beta_i + R^*_{J_1}\beta^*_{J_1}\right) - \sum_{i=1}^{J_1}R_i u_i(1-\alpha_i)\\
&\quad - R^*_{J_1}(1-\alpha^*_{J_1})a_1, \qquad C_1 = J_1,\\
D_2 &= \sum_{j=1}^{J_2}\bar{\alpha}_j v_j - A_2B_2\left(\sum_{j=1}^{J_2}(S_j+1)\bar{\beta}_j + S^*_{J_2}\bar{\beta}^*_{J_2}\right) - \sum_{j=1}^{J_2}S_j v_j(1-\bar{\alpha}_j)\\
&\quad - S^*_{J_2}(1-\bar{\alpha}^*_{J_2})a_2, \qquad C_2 = J_2,\\
E_1 &= \sum_{i=1}^{J_1}(R_i+1)\beta_i(u_i-A_1)^2 + R^*_{J_1}\beta^*_{J_1}(a_1-A_1)^2,\\
E_2 &= \sum_{j=1}^{J_2}(S_j+1)\bar{\beta}_j(v_j-A_2)^2 + S^*_{J_2}\bar{\beta}^*_{J_2}(a_2-A_2)^2.
\end{aligned}
$$

C. APPENDIX.

For three parameters case, we compute (2.14) at $\widehat{\theta} = (\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3)$, where

$$
\begin{aligned}
d_i &= \rho_1\sigma_{i1} + \rho_2\sigma_{i2} + \rho_3\sigma_{i3}, \quad i = 1,2,3,\\
d_4 &= u_{12}\sigma_{12} + u_{13}\sigma_{13} + u_{23}\sigma_{23},\\
d_5 &= \frac{1}{2}(u_{11}\sigma_{11} + u_{22}\sigma_{22} + u_{33}\sigma_{33}),\\
A &= \ell_{111}\sigma_{11} + 2\ell_{121}\sigma_{12} + 2\ell_{131}\sigma_{13} + 2\ell_{231}\sigma_{23} + \ell_{221}\sigma_{22} + \ell_{331}\sigma_{33},\\
B &= \ell_{112}\sigma_{11} + 2\ell_{122}\sigma_{12} + 2\ell_{132}\sigma_{13} + 2\ell_{232}\sigma_{23} + \ell_{222}\sigma_{22} + \ell_{332}\sigma_{33},\\
C &= \ell_{113}\sigma_{11} + 2\ell_{123}\sigma_{12} + 2\ell_{133}\sigma_{13} + 2\ell_{233}\sigma_{23} + \ell_{223}\sigma_{22} + \ell_{333}\sigma_{33}.
\end{aligned}
$$

In our case, for $(\theta_1,\theta_2,\theta_3)\equiv(\alpha,\beta,\lambda)$ and $u=R=\frac{\alpha}{\alpha+\beta}$, we have

$$\begin{aligned}
\rho_1 &= \frac{a_1-1}{\alpha}-b_1, \quad \rho_2=\frac{a_2-1}{\beta}-b_2, \quad \rho_3=\frac{a_3-1}{\lambda}-b_3,\\
\ell_{11} &= -\frac{J_1}{\alpha^2}, \qquad \ell_{22}=-\frac{J_2}{\beta^2}, \qquad \ell_{12}=\ell_{21}=0,\\
\ell_{13} &= \ell_{31}=-\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\frac{\log(x_i)}{1-x_i^\lambda}-R^*_{J_1}T_1^\lambda\frac{\log(T_1)}{1-T_1^\lambda},\\
\ell_{23} &= \ell_{32}=-\sum_{j=1}^{J_2}(S_j+1)y_j^\lambda\frac{\log(y_j)}{1-y_j^\lambda}-S^*_{J_2}T_2^\lambda\frac{\log(T_2)}{1-T_2^\lambda},\\
\ell_{33} &= -\frac{J_1+J_2}{\lambda^2}-\sum_{i=1}^{J_1}\big(\alpha(R_i+1)-1\big)x_i^\lambda\left(\frac{\log(x_i)}{1-x_i^\lambda}\right)^2-\alpha R^*_{J_1}T_1^\lambda\left(\frac{\log(T_1)}{1-T_1^\lambda}\right)^2\\
&\quad -\sum_{j=1}^{J_2}\big(\beta(S_j+1)-1\big)y_j^\lambda\left(\frac{\log(y_j)}{1-y_j^\lambda}\right)^2-\beta S^*_{J_2}T_2^\lambda\left(\frac{\log(T_2)}{1-T_2^\lambda}\right)^2,
\end{aligned}$$

σ_{ij}, $i,j=1,2,3$ are obtained using ℓ_{ij}, $i,j=1,2,3$ and

$$\begin{aligned}
\ell_{111} &= \frac{2J_1}{\alpha^3}, \qquad \ell_{222}=\frac{2J_2}{\beta^3}\\
\ell_{133} &= \ell_{331}=\ell_{313}=-\sum_{i=1}^{J_1}(R_i+1)x_i^\lambda\left(\frac{\log(x_i)}{1-x_i^\lambda}\right)^2-R^*_{J_1}T_1^\lambda\left(\frac{\log(T_1)}{1-T_1^\lambda}\right)^2,\\
\ell_{233} &= \ell_{332}=\ell_{323}=-\sum_{j=1}^{J_2}(S_j+1)y_j^\lambda\left(\frac{\log(y_j)}{1-y_j^\lambda}\right)^2-S^*_{J_2}T_2^\lambda\left(\frac{\log(T_2)}{1-T_2^\lambda}\right)^2,\\
\ell_{333} &= \frac{2(J_1+J_2)}{\lambda^3}-\sum_{i=1}^{J_1}\big(\alpha(R_i+1)-1\big)x_i^\lambda(1+x_i^\lambda)\left(\frac{\log(x_i)}{1-x_i^\lambda}\right)^3\\
&\quad -\sum_{j=1}^{J_2}\big(\beta(S_j+1)-1\big)y_j^\lambda(1+y_j^\lambda)\left(\frac{\log(y_j)}{1-y_j^\lambda}\right)^3-\alpha R^*_{J_1}T_1^\lambda(1+T_1^\lambda)\left(\frac{\log(T_1)}{1-T_1^\lambda}\right)^3\\
&\quad -\beta S^*_{J_2}T_2^\lambda(1+T_2^\lambda)\left(\frac{\log(T_2)}{1-T_2^\lambda}\right)^3,
\end{aligned}$$

and other $\ell_{ijk}=0$. Moreover, $u_3=u_{i3}=0$, $i=1,2,3$, and u_1, u_2 are given in (2.12). Also, $u_{11}=\frac{-2\beta}{(\alpha+\beta)^3}$, $u_{12}=u_{21}=\frac{\alpha-\beta}{(\alpha+\beta)^3}$, $u_{22}=\frac{2\alpha}{(\alpha+\beta)^3}$. So,

$$\begin{aligned}
&d_4=u_{12}\sigma_{12}, \qquad d_5=\frac{1}{2}(u_{11}\sigma_{11}+u_{22}\sigma_{22}),\\
&A=\ell_{111}\sigma_{11}+\ell_{331}\sigma_{33}, \quad B=\ell_{222}\sigma_{22}+\ell_{332}\sigma_{33}, \quad C=2\ell_{133}\sigma_{13}+2\ell_{233}\sigma_{23}+\ell_{333}\sigma_{33}.
\end{aligned}$$

ACKNOWLEDGMENTS

The author of this study would like to thank the Editor, Associate Editor, and referees for their careful reading and comments, which greatly improved the quality of this paper.

REFERENCES

[1] Ahmadi, K. and Ghafouri, S. (2019). Reliability estimation in a multicomponent stress-strength model under generalized half-normal distribution based on progressive type-II censoring, *Journal of Statistical Computation and Simulation*, **89**, 2505–2548.

[2] Babayi, S.; Khorram, E. and Tondro, F. (2014). Inference of $R = P[X < Y]$ for generalized logistic distribution, *Statistics*, **48**, 862–871.

[3] Birnbaum, Z.W. (1956). *On a use of the Mann–Whitney statistic*. In "Proceedings of the third Berkley Symposium in Mathematics, Statistics and Probability, Vol. 1, 13–17.

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

[5] Epstein, B. (1954). Truncated life tests in the exponential case, *The Annals of Mathematical Statistics*, **25**, 555–564.

[6] Kizilaslan, F. and Nadar, M. (2016). Estimation and prediction of the Kumaraswamy distribution based on record values and inter-record times, *Journal of Statistical Computation and Simulation*, **86**, 2471–2493.

[7] Kizilaslan, F. and Nadar, M. (2017). Statistical inference of $P(X < Y)$ for the Burr Type XII distribution based on records, *Hacettepe Journal of Mathematics and Statistics*, **46**, 713–742.

[8] Kizilaslan, F. and Nadar, M. (2018). Estimation of reliability in a multicomponent stress-strength model based on a bivariate Kumaraswamy distribution, *Statistical Papers*, **59**, 307–340.

[9] Kohansal, A. (2019). On estimation of reliability in a multicomponent stress-strength model for a Kumaraswamy distribution based on progressively censored sample, *Statistical Papers*, **60**, 2185–2224.

[10] Kohansal, A. (2020). Bayesian and classical estimation of $R = P(X < Y)$ based on Burr type XII distribution under hybrid progressive censored samples, *Communications in Statistics – Theory and Methods*, **49**, 1043–1081.

[11] Kohansal, A. and Nadarajah, S. (2019). Stress-strength parameter estimation based on Type-II hybrid progressive censored samples for a Kumaraswamy distribution, *IEEE Transactions on Reliability*, **68**, 1296–1310.

[12] Kohansal, A. and Rezakhah, S. (2019). Inference of $R = P(Y < X)$ for two-parameter Rayleigh distribution based on progressively, *Statistics*, **53**, 81–100.

[13] Kohansal, A. and Shoaee, Sh. (2021). Bayesian and classical estimation of reliability in a multicomponent stress-strength model under adaptive hybrid progressive censored data, *Statistical Papers*, **62**, 309–359.

[14] Kundu, D. and Joarder, A. (2006). Analysis of type-II progressively hybrid censored data, *Computational Statistics and Data Analysis*, **50**, 2509–2528.

[15] Lawless, J.F. (2003). *Statistical Models and Methods for Lifetime Data*, 2nd ed., Hoboken, NJ, Wiley.

[16] Lindley, D.V. (1980). Approximate Bayesian methods, *Trabajos de Estadistica*, **3**, 281–288.

[17] Maurya, R.K. and Tripathi, Y.M. (2020). Reliability estimation in a multicomponent stress-strength model for Burr XII distribution under progressive censoring, *Brazilian Journal of Probability and Statistics*, **34**, 345–369.

[18] Nadar, M.; Kizilaslan, F. and Papadopoulos, A. (2014). Classical and Bayesian estimation of $P(Y < X)$ for Kumaraswamy's distribution, *Journal of Statistical Computation and Simulation*, **84**, 1505–1529.

[19] Nadar, M.; Papadopoulos, A. and Kizilaslan, F. (2013). Statistical analysis for Kumaraswamy's distribution based on record data, *Statistical Papers*, **54**, 355–369.

[20] Rasethuntsa, T.R. and Nadar, M. (2018). Stress-strength reliability of a non-identical-component-strengths system based on upper record values from the family of Kumaraswamy generalized distributions, *Statistics*, **52**, 684–716.

[21] Rezaei, S.; Tahmasbi, R. and Mahmoodi, M. (2010). Estimation of $P[Y < X]$ for generalized Pareto distribution, *Journal of Statistical Planning and Inference*, **140**, 480–494.

[22] Shoaee, Sh. and Khorram, E. (2015). Stress-strength reliability of a two-parameter bathtub-shaped lifetime distribution based on progressively censored samples, *Communications in Statistics – Theory and Methods*, **44**, 5306–5328.

[23] Shoaee, Sh. and Khorram, E. (2016). Statistical inference of $\mathcal{R} = P(Y < X)$ for Weibull distribution under Type-II progressively hybrid censored data, *Journal of Statistical Computation and Simulation*, **86**, 3815–3834.

[24] Singh, B. and Goel, R. (2018). Reliability estimation of modified Weibull distribution with Type-II hybrid censored data, *Iranian Journal of Science and Technology, Transactions A: Science*, **42**, 1395–1407.

[25] Sultana, F.; Tripathi, Y.M.; Rastogi, M.K. and Wu, S. (2018). Parameter estimation for the Kumaraswamy distribution based on hybrid censoring, *American Journal of Mathematical and Management Sciences*, **37**, 243–261.

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 85–100

THE SOLUTION TO A DIFFERENTIAL–DIFFERENCE EQUATION ARISING IN OPTIMAL STOPPING OF A JUMP–DIFFUSION PROCESS

Authors: Cláudia Nunes
– Department of Mathematics and CEMAT, Instituto Superior Técnico,
Av. Rovisco Pais, 1049–001 Lisboa, Portugal
cnunes@math.tecnico.ulisboa.pt

Rita Pimentel
– RISE SICS Västerås AB,
Stora Gatan 36, SE-722 12 Västerås, Sweden
rita.duarte.pimentel@ri.se

Ana Prior
– Department of Mathematics, ISEL,
R. Conselheiro Emídio Navarro 1, 1959-007 Lisboa, Portugal
afpontes@adm.isel.pt

Received: September 2018 Revised: July 2019 Accepted: November 2019

Abstract:

- In this paper we present a solution to a second order differential–difference equation that occurs in different contexts, specially in control engineering and finance. This equation leads to an ordinary differential equation, whose homogeneous part is a Cauchy–Euler equation. We derive a particular solution to this equation, presenting explicitly all the coefficients.
  The differential–difference equation is motivated by investment decisions addressed in the context of real options. It appears when the underlying stochastic process follows a jump–diffusion process, where the diffusion is a geometric Brownian motion and the jumps are driven by a Poisson process. The solution that we present — which takes into account the geometry of the problem — can be written backwards, and therefore its analysis is easier to follow.

Keywords:

- *differential–difference equation; differential equation, jump–diffusion process.*

AMS Subject Classification:

- 37H10, 60G40.

1. MOTIVATION AND RELATED WORK

In this paper we present a solution to a second order differential–difference equation. This equation may appear when one solves an optimal stopping problem in which the state process follows a jump–diffusion process, where the diffusion is a geometric Brownian motion and the jumps are driven by a Poisson process. The main difficulty of working with this type of equation is due to the jump process, which makes the equation not local in one point — see, for instance, Murto [25]. This characteristic is not universal, i.e., there are optimal stopping problems involving jump–diffusions processes for which the differential–difference equation does not exhibit this behavior. Thus, on these cases to find a closed form solution to the differential–difference equation can be easier. However, as we will see later, this is not the case when the jumps may lead directly to the stopping region, across the boundary.

The seminal works in financial options — such as the classical work of Black and Scholes [5], where for the first time a pricing formula was derived — and in real options — as the seminal book of Dixit and Pindyck [12] — assume that the sample path of the involved state process is continuous, with probability one.

In recent times investors often need to take decisions facing uncertainty and there is higher likelihood of financial crashes, which are the climax of the so-called log-periodic power law signatures associated with speculative bubbles (see Johansen and Sornette [18]). One example of this occurred in February 2015, when due to a cyber-attack, a high-frequency trading company started uncontrollably buying oil futures, causing a downward jump in the oil prices[1]. Here, a crash is a significant drop in the total value of the market, creating a situation wherein the majority of investors are trying to flee the market at the same time and consequently incurring massive losses. Indeed, in the presence of a crash investors likely take the decision to sell their assets. As the crash means that there is a significant drop, we borrow the probabilistic terminology and we call it a *jump* (in the above example, a downward jump).

The sudden changes in the state variable can also be found when one decides about investments in projects, often addressed in the context of real options. In this context, usually the temporal term is long, and therefore unexpected events may occur, leading to a disruption of the market. One example of a disrupt event is the introduction or the abolition of public subsidies. There are many economical sectors where subsidies play an important role, such as agriculture.

Due to the interest of the equation that we solve in this paper in the framework of real options, we mainly focus in problems and questions arising in such context. The following are examples of decisions regarding investments where investors face the likelihood of sudden events.

It is well established that agricultural pricing policies (taxes, subsidies) have a substantial influence on farmer production decisions[2]. For example, USA has been supporting farming since early times. But after several decades, these incentive policies have proved to

[1] https://www.businessinsider.com/investigation-into-hft-firm-for-using-an-algo-gone-wild-that-caused-oil-trading-mayhem-in-just-5-seconds-2010-8

[2] http://www.pbl.nl/en/publications/the-impact-of-taxes-and-subsidies-on-crop-yields

be unsuccessful [3]. In 2005 Bush administration decided to change the farm incentive policy, cutting in agricultural subsidies[4]. Evidently, this decision led to changes in private investment farming projects.

Another area where subsidies play an important role is the renewable energy (RE) sector. In an effort to reach the ambitious targets of the EU Strategic Energy Technology Plan (SET-Plan), EU member states have implemented support mechanisms of various forms (e.g., price mechanisms, like carbon tax or permit trading schemes) intended to incentive and accelerate adoption of RE technologies. These climate change policies have introduced a new factor that has to be included into the investment decision and have become a major source of uncertainty in energy strategy. The problem is that policies designed to stimulate the investment in green energies have frequently and unexpectedly been changed for a number of reasons. For instance, change of governments, collapse of the international cooperation for reducing GHG emissions, arrival of new information about climate sensitivity, and fiscal pressure. In the last decade we have seen many studies on the impact of wrong investment decisions. We refer, for instance, to Boomsma and Linnerud [6], Boomsma *et al.* [7], and Hagspiel *et al.* [14].

These examples show that when taking decisions regarding investments in new projects, the investor needs to take into account these sudden changes. The area of real options soon realized the importance of such events, and therefore the interest of real options literature in problems involving jump–diffusion processes is not new. We refer to Kou [21] for a survey on jump–diffusion models for finance engineering. In the area of real options, there has been an increasing interest about jump–diffusion processes in the context of technology adoption (see, for instance, Hagspiel *et al.* [16]).

Furthermore, Kwon [22] and Hagspiel *et al.* [15] consider a combination of a continuous process with a jump-process, but they do not consider a sequence of innovations arriving over time. Instead, they assume a one-single innovation opportunity, with other involved options (like the option to exit the market). Kwon [22] work is generalized in Hagspiel *et al.* [15], by considering capacity optimization, and by Matomäki [23], considering different stochastic processes representing the profit uncertainty.

In another context, Couto *et al.* [11] and Nunes and Pimentel [26] consider the investment problem in a high-speed railway service, assuming that both the demand and the investment cost are modeled by jump–diffusion processes. Although these papers start by assuming two sources of uncertainty, they end up with the study of a one-dimensional problem. This happens because they assume that the value of the firm is homogeneous, and therefore it is possible to consider a change of variables that will turn the two-dimensional problem in a one-dimensional one. Murto [25] also consider two stochastic processes, in order to model technological and revenue uncertainties, motivated by wind power investment. He assumes that the investment cost depends on the technological progress, driven by a pure Poisson process, whereas the price of the output is a geometric Brownian motion. As the value of the project is homogeneous, the same type of approach as in Nunes and Pimentel [26] is proposed.

In all the above examples, it is of the most crucial importance to assess the impact of the jumps in the decision, and, in particular, in case the jumps anticipate the optimal decision.

[3] https://grist.org/article/farm_bill2/
[4] https://www.agpolicy.org/weekpdf/258.pdf

Moreover, the impact of such jumps has to be reflected in the *value* of the project, which is a quantitative measure of the value that the firm has as a result of its option to invest. Under an optimal strategy in terms of the investment timing, such value is, before the investment, solution of a differential equation (that we will present in section 2). Mathematically, the possibility of occurrence of jumps leads to this value being solution of particular types of differential equations.

Our contribution to the state of the art is two-fold: on one side, we provide an analytical solution to a non-homogeneous differential equation. As it turns out, some optimal stopping problems found in real options lead to a differential–difference equations that are exactly as the form of such differential equation, for a subset of the state space. Therefore, one may use this analytical solution to provide a characterization of the value of a firm, which is given by a piecewise function.

The paper is organized as follows: in Section 2 we motivate the differential–difference equation that we address in this paper, presenting also the basic assumptions. In Section 3 we show how we can find a general solution for such equation, using a backwards procedure. This procedure presents the solution as a piecewise function. For each branch, the function is the solution of a non-homogeneous differential equation. Therefore, in Section 4 we provide the particular solution to it. Finally, in Section 5 we conclude.

2. DIFFERENTIAL–DIFFERENCE EQUATION

In order to motivate the meaningfulness of the differential–difference equation solved in this paper, we consider that we want to derive the value of a firm that has the option to undertake an investment. As we briefly explain in this section, these type of problems leads to a variational inequality known as the Hamilton–Jacobi–Bellman (HJB, for short) equation, where one of the members is a differential–difference equation. To solve such equation, we also need to be able to find the solution of a differential equation of the following type:

$$x^2y''(x) + axy'(x) + by(x) = Ax^\alpha(\ln x)^n, \tag{2.1}$$

with $x > 0$, $a, b \in \mathbb{R}$, $\alpha, A \in \mathbb{R} \setminus \{0\}$ and $n \in \mathbb{N}_0$.

We note that the corresponding homogeneous equation to (2.1) is an Euler–Cauchy equation and its solution is known. The difficulty lays in the particular solution, consequence of the non-homogeneous term, $Ax^\alpha(\ln x)^n$.

The result that we provide in this paper is *per se* interesting, as it provides a contribution to the area of ordinary differential equations (ODE). Besides this contribution, being able to compute the solution of such equation is also relevant for the applications. Next we motivate the mathematical problem by an investment problem, using the terminology and notation of real options.

Real options is a theory on how to make decisions under uncertainty about future returns. These decisions share the following two characteristics: they are irreversible and can be postponed.

One of the most relevant problems in real options regards the characterization of the optimal time to undertake some investment decision. This leads to an *optimal stopping problem*, which is formally defined as follows: given a stochastic process $\mathbf{X} = \{X(t), t > 0\}$, find $V(x)$ and $\tau^\star \in \mathcal{T}$ such that

$$V(x) = \sup_{\tau \in \mathcal{T}} \mathbb{E}^x \left[e^{-r\tau} g(X(\tau)) \chi_{\{\tau < +\infty\}} \right], \quad x \in \mathbb{R}^+, \tag{2.2}$$

with $\mathcal{T}$ being the set of all stopping times adapted to the filtration generated by the process $\mathbf{X}$, $r > 0$ states for the discount factor and $\chi_{\{A\}}$ represents the indicator function on set A. The function g is usually called *running function*, which accounts for the return of the investment.

The class of stochastic processes that lead to the type of equations that we study in this paper — equation (2.1) — is an one-dimensional jump–diffusion, which is the strong solution of the following stochastic differential equation:

$$\frac{dX(t)}{X(t^-)} = \mu dt + \sigma dW(t) + \kappa dN(t),$$

with initial value $X(0) = x > 0$, where $\{W(t), t > 0\}$ is a standard one-dimensional Brownian motion, and $\{N(t), t > 0\}$ is a centered time-homogeneous Poisson process, with intensity $\lambda > 0$. Moreover, μ is the drift of the process $\mathbf{X}$, $\sigma > 0$ is its volatility and κ is the multiplicative factor, in case a jump occurs. The notation $X(t^-)$ means that whenever there is a jump, the value of the process before the jump is considered. Motivated by the references mentioned in Section 1, we assume that the jumps are multiplicative and with constant magnitude.

One way to solve the optimal stopping problem defined in (2.2) is to solve the variational inequality HJB (we do not provide further details, referring instead to Peskir and Shiryaev [28]). In this case, the corresponding HJB equation is the following:

$$\min \{ rV(x) - \mathcal{L}V(x), V(x) - g(x) \} = 0, \tag{2.3}$$

where $\mathcal{L}$ is the infinitesimal generator of the process $\mathbf{X}$. As $\mathbf{X}$ is a jump–diffusion process, it follows that

$$\mathcal{L}v(x) = \frac{\sigma^2}{2} x^2 v''(x) + (\mu - \lambda\kappa) x v'(x) + \lambda \left(v(x(1+\kappa)) - v(x) \right), \tag{2.4}$$

for $v \in C^1$ and $x \in \mathbb{R}^+$ (see Øksendal and Sulem [27] for more details).

In general, the use of these inequalities leads to a differential equation, which in some cases may be solved analytically. Besides the possible difficulty to find the analytical solution to the differential equation, one faces also the problem to find the boundary conditions, as the set of values where the differential equation holds is also unknown. For this reason the problem presented in (2.2) when solved by the use of variational inequalities is known in the literature as a *free boundary problem*.

Considering an investment problem, the differential equation holds in the region where it is not optimal to stop (in our case to invest). For that reason, this region is usually called *continuation region*, and in opposition its complementary is called *stopping region*. In some cases, one can provide a guess for the shape of the continuation set. For example, if g is a non-decreasing function, the firm takes the decision to invest for large values of x, whereas for small values of x the firm postpones its investment decision. Thus, the stopping region is of the form $\mathcal{S} = [x^*, +\infty)$ and the continuation region is $\mathcal{C} = (0, x^*)$, where x^* is the exercise threshold.

When $\mathbf{X}$ is a jump–diffusion process with positive jumps, the stopping region can be reached in two different ways:

(i) Either due to a continuous change, caused by the diffusion. In this case the state process *hits* the boundary threshold x^*.

(ii) Or due to the occurrence of a jump. In this case the state process *crosses* the boundary threshold.

In the literature, the majority of the authors address the case that either there is just the jump process (for which it is possible to solve the corresponding difference equation, as there is no differential part) — this is the case of Huisman [17] — or the process is a jump–diffusion but the jumps always lead to the continuation region — which is the case of Nunes and Pimentel [26].

Our work is related with Merton [24], who considers a model to price American call options. He assumes multiplicative independent and identically distributed jumps and $g(x) = \max(x - K, 0)$ (the payoff of an American call option). For this case, he provides in Equation (16) a semi-analytical result, as it involves a series with infinite number of terms that depend, each one, on the cumulative distribution of a normal random variable. More recently, Murto [25] considers a problem with a similar setting as ours. However, in view of the impossibility to derive an analytical solution, he provides solutions only for some particular cases (namely, if the volatility parameter of the diffusion is zero, or when the jump process is in fact deterministic, with an exponential decay).

In the current paper, we assume a non-decreasing g function. Then it follows that on the one hand, in the stopping region V is equal to g, i.e. $V(x) = g(x)$ for $x \geq x^\star$. On the other hand, in the continuation region the value function V must be the solution of the left-hand side of the HJB Equation (2.3), which combined with Equation (2.4), leads to the following equation:

$$x^2 V''(x) + a\, xV'(x) + b\, V(x) - c\, V(x(1+k)) = 0, \tag{2.5}$$

where $a = \frac{2(\mu-\lambda k)}{\sigma^2}$, $b = -\frac{2(r+\lambda)}{\sigma^2}$ and $c = -\frac{2\lambda}{\sigma^2}$. This is called in the literature *mixed partial differential–difference equation*, and it is known to be difficult to solve (see Merton [24]).

3. BACKWARDS ANALYSIS

In this section we provide a backwards procedure that can be used to solve the Equation (2.5). This procedure is motivated by the geometry of the stopping/continuation regions previously presented, when g is non-decreasing.

Firstly, we note that the homogeneous part of Equation (2.5) has an analytical solution, hereby denoted by V_h, which is given by

$$V_h(x) = \delta_1 x^{\beta_1} + \delta_2 x^{\beta_2}, \tag{3.1}$$

where β_1 and β_2 are the roots of the characteristic polynomial

$$Q(\beta) = \beta\,(\beta - 1) + a\beta + b. \tag{3.2}$$

In our case, given that $b < 0$, there are two distinct real roots:

$$\beta_1 = \frac{1}{2}\left[1 - a + \sqrt{(1-a)^2 - 4b}\right] > 0; \tag{3.3}$$

$$\beta_2 = \frac{1}{2}\left[1 - a - \sqrt{(1-a)^2 - 4b}\right] < 0. \tag{3.4}$$

As presented before, $V(x) = g(x)$ for $x \in [x^\star, +\infty)$. Therefore, one needs to solve the problem for $0 < x < x^\star$. For that, we start by considering $x \in \left[\frac{x^\star}{1+\kappa}, x^\star\right)$, meaning that $x(1+\kappa) \geq x^\star$. So, the interval $\left[\frac{x^\star}{1+\kappa}, x^\star\right)$ is the set of values of x where stopping will surely happen if a jump occurs. Thus $V(x(1+\kappa)) = g(x(1+\kappa))$. In this case Equation (2.5) can be re-written as

$$x^2 V''(x) + a\, xV'(x) + b\, V(x) = c\, g(x(1+\kappa)).$$

and therefore its solution, hereby denoted by V_1, is given by

$$V(x) := V_1(x) = V_h(x) + V_p^1(x) = \delta_1 x^{\beta_1} + \delta_2 x^{\beta_2} + f_g^1(x), \tag{3.5}$$

Note that the superscript in V_p^1 and f_g^1 represents how many jumps we are away from the stopping region[5] (see Figure 1 for an illustration). Moreover, the bottom index in f_g^1 emphasizes that this function depends explicitly on g.

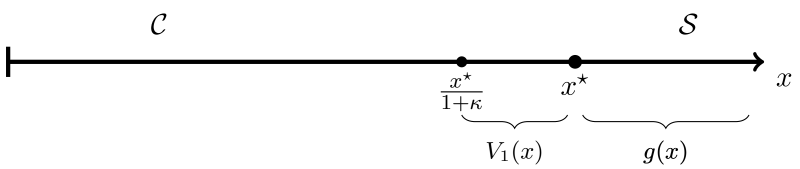

Figure 1: Representation of V in the last interval before stopping.

Next we derive the value of V when we are two jumps away from the stopping region. Following the same notation, we denote this function by V_2, defined for $x \in \left[\frac{x^\star}{(1+\kappa)^2}, \frac{x^\star}{1+\kappa}\right)$. In this case $x(1+\kappa) \in \left[\frac{x^\star}{1+\kappa}, x^\star\right)$, so $V(x(1+\kappa)) = V_1(x(1+\kappa))$. This means that (2.5) can be re-written as follows:

$$x^2 V''(x) + a\, xV'(x) + b\, V(x) = c\, V_1(x(1+\kappa)).$$

The homogeneous part of the previous equation is the same as before, and thus the solution is provided in (3.1). We just need to take into account the particular solution, which we denote by V_p^2. This particular solution depends on V_p^1 (and thus depends on g) but also depends on V_h (then also depends on the roots of Q, β_1 and β_2), as V_1 is given by (3.5). Therefore, both the homogeneous and the particular solution for this case *share* the powers β_1 and β_2. Using Theorem 3.5 of Sabuwala and De Leon [29], we end up with the following particular solution:

$$V_p^2(x) = \eta_1^2 \ln x\; x^{\beta_1} + \eta_2^2 \ln x\; x^{\beta_2} + f_g^2(x).$$

[5] We use this type of notation for all particular solutions.

We write f_g^2 to denote the part of the solution that depends strictly on g (following the same reasoning as for f_g^1), whereas η_1^2 and η_2^2 depend on the parameters from the homogeneous solution. So, for $x \in \left[\frac{x^\star}{(1+\kappa)^2}, \frac{x^\star}{1+\kappa}\right)$ (see Figure 2 for an illustration), we have

$$V(x) := V_2(x) = \delta_1 x^{\beta_1} + \delta_2 x^{\beta_2} + \eta_1^2 \ln x \, x^{\beta_1} + \eta_2^2 \ln x \, x^{\beta_2} + f_g^2(x). \tag{3.6}$$

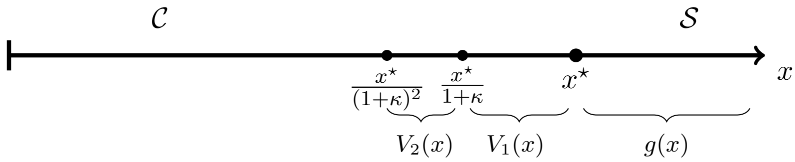

Figure 2: Representation of V in the last two intervals before stopping.

Proceeding one step back, we determine the value of V when we are three jumps away from the stopping region, which we call V_3. When $x \in \left[\frac{x^\star}{(1+\kappa)^3}, \frac{x^\star}{(1+\kappa)^2}\right)$, then $x(1+\kappa) \in \left[\frac{x^\star}{(1+\kappa)^2}, \frac{x^\star}{1+\kappa}\right)$ and $V(x(1+\kappa)) = V_2(x(1+\kappa))$. Then, Equation (2.5) is re-written as

$$x^2 V''(x) + a\, xV'(x) + b\, V(x) = c\, V_2(x(1+\kappa)). \tag{3.7}$$

As before, the homogeneous equation is the same and therefore V_h is part of the solution of this equation. Once more, the problem is reduced to the derivation of a particular solution, which is not trivial, as the function V_2 involves polynomials of power β_1 and β_2 multiplied by a logarithm (see Equation (3.6)). After some calculations, one may find that the particular solution of (3.7) is of the following form:

$$V_p^3(x) = \eta_1^3 \ln x \, x^{\beta_1} + \eta_2^3 \ln x \, x^{\beta_2} + \eta_3^3 (\ln x)^2 x^{\beta_1} + \eta_4^3 (\ln x)^2 x^{\beta_2} + f_g^3(x).$$

Also here f_g^3 stands for the part of the solution that depends strictly on g whereas $\eta_1^3, \eta_2^3, \eta_3^3$ and η_4^3 depend on the parameters from the homogeneous solution. As previously, for $x \in \left[\frac{x^\star}{(1+\kappa)^3}, \frac{x^\star}{(1+\kappa)^2}\right)$, we have

$$\begin{aligned} V(x) := V_3(x) \;=\; & \delta_1 x^{\beta_1} + \delta_2 x^{\beta_2} + \eta_1^3 \ln x \, x^{\beta_1} + \eta_2^3 \ln x \, x^{\beta_2} \\ & + \eta_3^3 (\ln x)^2 x^{\beta_1} + \eta_4^3 (\ln x)^2 x^{\beta_2} + f_g^3(x). \end{aligned}$$

A similar reasoning applies for other intervals of x. When we are i (with $i \in \mathbb{N}$) jumps away from the stopping region, we have $\frac{x^\star}{(1+\kappa)^i} \le x < \frac{x^\star}{(1+\kappa)^{i-1}}$ and V is represented by V_i, which may be obtained using a similar procedure as the one used for V_1, V_2 and V_3. Indeed, V is a piecewise function, given by

$$V(x) = \begin{cases} V_i(x) & \text{if} \quad \frac{x^\star}{(1+\kappa)^i} \le x < \frac{x^\star}{(1+\kappa)^{i-1}} \\ g(x) & \text{if} \quad x \ge x^\star \end{cases},$$

where

$$V_i(x) = \delta_1 x^{\beta_1} + \delta_2 x^{\beta_2} + V_p^i(x)$$

with

$$V_p^1(x) = f_g^1(x) \quad \text{and}$$

$$V_p^i(x) = \sum_{j=1}^{i-1} \left[\eta_{2j-1}^i \, x^{\beta_1} + \eta_{2j}^i \, x^{\beta_2}\right] (\ln x)^j + f_g^i(x), \text{ for } i \in \mathbb{N} \setminus \{1\} . \tag{3.8}$$

Clearly, one needs to find functions that are solutions of certain differential equations, that depend intrinsically on the function g, considered in the definition of the problem.

For example, for

$$g(x) = \rho x^\theta - I \tag{3.9}$$

we obtain the following particular solutions V_p^i, for $i = 1, 2, 3$:

$$V_p^1(x) = \xi_1^1 x^\theta + \xi_2^1, \quad \text{with} \quad \xi_1^1 = \frac{c\rho(1+\kappa)^\theta}{Q(\theta)}, \ \xi_2^1 = -\frac{cI}{b}.$$

$$\begin{aligned}
V_p^2(x) = &\eta_1^2 \ln x \, x^{\beta_1} + \eta_2^2 \ln x \, x^{\beta_2} + \xi_1^2 x^\theta + \xi_2^2, \quad \text{with} \\
&\eta_1^2 = \delta_1 \frac{c(1+\kappa)^{\beta_1}}{Q'(\beta_1)}, \ \eta_2^2 = \delta_2 \frac{c(1+\kappa)^{\beta_2}}{Q'(\beta_2)}, \\
&\xi_1^2 = \rho \left[\frac{c(1+\kappa)^\theta}{Q(\theta)}\right]^2, \ \xi_2^2 = -\left(\frac{c}{b}\right)^2 I.
\end{aligned}$$

$$\begin{aligned}
V_p^3(x) = &\eta_1^3 \ln x \, x^{\beta_1} + \eta_2^3 \ln x \, x^{\beta_2} + \eta_3^3 (\ln x)^2 \, x^{\beta_1} + \eta_4^3 (\ln x)^2 \, x^{\beta_2} + \xi_1^3 x^\theta + \xi_2^3, \quad \text{with} \\
&\eta_1^3 = \delta_1 \frac{c(1+\kappa)^{\beta_1}}{Q'(\beta_1)} \left[1 + \frac{c(1+\kappa)^{\beta_1}}{Q'(\beta_1)} \left(\ln(1+\kappa) - \frac{1}{Q'(\beta_1)}\right)\right], \\
&\eta_2^3 = \delta_2 \frac{c(1+\kappa)^{\beta_2}}{Q'(\beta_2)} \left[1 + \frac{c(1+\kappa)^{\beta_2}}{Q'(\beta_2)} \left(\ln(1+\kappa) - \frac{1}{Q'(\beta_2)}\right)\right], \\
&\eta_3^3 = \frac{\delta_1}{2} \left[\frac{c(1+\kappa)^{\beta_1}}{Q'(\beta_1)}\right]^2, \ \eta_4^3 = \frac{\delta_2}{2} \left[\frac{c(1+\kappa)^{\beta_2}}{Q'(\beta_2)}\right]^2, \\
&\xi_1^3 = \rho \left[\frac{c(1+\kappa)^\theta}{Q(\theta)}\right]^3, \ \xi_2^3 = -\left(\frac{c}{b}\right)^3 I.
\end{aligned}$$

For simplicity, in the above calculations we assume that θ is not a root of the characteristic polynomial Q. This example is motivated by the relevance of this analysis in real options context. In fact, functions such that the one presented in (3.9) are frequently used in this context and describe the profit of a firm. This function is called in the literature an iso-elastic demand function (see, for instance, Nunes and Pimentel [26]).

This example also shows that a more systematic way to find the solution to the nonhomogeneous differential Equation (2.1) is quite valuable. We address this issue in the next section.

4. MAIN RESULTS

We want to find a particular solution to the Equation (2.1). The type of solution is understandable from the special case solved at the end of the previous section. However, a systematic way to obtain all the coefficients is not so easy to develop.

We start deriving a recursive expression for the particular solution of (2.1). Later, using this result, we will be able to present explicit expressions for the involved coefficients.

Theorem 4.1 (recursive)**.** *Consider the second order ODE presented in* (2.1), *with the corresponding characteristic polynomial Q given by* (3.2). *Then the following cases occur:*

- *If α is not a root of Q, the particular solution of* (2.1) *is*

$$y_p(x) = x^\alpha \sum_{i=0}^{n} c_i \,(\ln x)^i\,,$$

where $c_n = \frac{A}{Q(\alpha)}$, $c_{n-1} = -nA\frac{Q'(\alpha)}{Q(\alpha)^2}$ *and* $c_i = -\frac{i+1}{Q(\alpha)}\left[Q'(\alpha)c_{i+1} + (i+2)c_{i+2}\right]$ *for* $i = 0, 1, 2, ..., n-2$.

- *If α is a simple root of Q, the particular solution of* (2.1) *is*

$$y_p(x) = x^\alpha \sum_{i=0}^{n} c_i \,(\ln x)^{i+1}\,,$$

where $c_n = \frac{A}{(n+1)Q'(\alpha)}$ *and* $c_i = -\frac{i+2}{Q'(\alpha)}c_{i+1}$, *for* $i = 0, 1, 2, ..., n-1$.

- *If α is a root of Q with multiplicity two, the particular solution of* (2.1) *is*

$$y_p(x) = x^\alpha c_n \,(\ln x)^{n+2}\,,$$

where $c_n = \frac{A}{(n+1)(n+2)}$.

Proof: We start by proposing that the particular solution of Equation (2.1) is of the form $y_p(x) = x^\alpha P(x)$. Calculating first and second derivatives, we obtain

$$\begin{aligned} y_p'(x) &= x^{\alpha-1}\left[xP'(x) + \alpha P(x)\right] \\ y_p''(x) &= x^{\alpha-2}\left[x^2P''(x) + 2\alpha x P'(x) + \alpha(\alpha-1)P(x)\right], \end{aligned}$$

from where

$$x^2 y_p''(x) + axy_p'(x) + by_p(x) \;=\; x^\alpha\left[x^2P''(x) + \left(Q'(\alpha)+1\right)xP'(x) + Q(\alpha)P(x)\right].$$

Thus $P(x)$ is such that

$$x^2P''(x) + \left(Q'(\alpha)+1\right)xP'(x) + Q(\alpha)P(x) = A(\ln x)^n. \tag{4.1}$$

Taking into account whether $Q(\alpha)$ is null or not, we end up with different cases, described hereafter:

1. If α is not a root of Q, then $P(x) = \sum_{i=0}^{n} c_i (\ln x)^i$, as we prove next. For that, we compute the first and second derivatives:

$$\begin{aligned} P'(x) &= \frac{1}{x}\sum_{i=1}^{n} ic_i (\ln x)^{i-1} \\ P''(x) &= \frac{1}{x^2}\left[\sum_{i=2}^{n} i(i-1)c_i(\ln x)^{i-2} - \sum_{i=1}^{n} ic_i(\ln x)^{i-1}\right]. \end{aligned}$$

   Thus, $x^2P''(x) + (Q'(\alpha)+1)\,xP'(x) + Q(\alpha)P(x)$ is given by

$$\begin{aligned} &\sum_{i=0}^{n-2}\left[(i+2)(i+1)c_{i+2} + Q'(\alpha)(i+1)c_{i+1} + Q(\alpha)c_i\right](\ln x)^i \\ &+ \left[Q'(\alpha)nc_n + Q(\alpha)c_{n-1}\right](\ln x)^{n-1} + Q(\alpha)c_n(\ln x)^n. \end{aligned}$$

   Therefore, (4.1) holds if $Q(\alpha)c_n = A$, $Q'(\alpha)nc_n + Q(\alpha)c_{n-1} = 0$ and $(i+2)(i+1)c_{i+2} + Q'(\alpha)(i+1)c_{i+1} + Q(\alpha)c_i = 0$, for $i = 0, 1, ..., n-2$, which leads to the result.

2. If α is a root of Q with multiplicity one, then $P(x) = \sum_{i=0}^{n} c_i (\ln x)^{i+1}$. In fact, calculating first and second derivatives, we obtain

$$\begin{aligned} P'(x) &= \frac{1}{x}\sum_{i=0}^{n}(i+1)c_i (\ln x)^i, \\ P''(x) &= \frac{1}{x^2}\left[\sum_{i=1}^{n}(i+1)ic_i(\ln x)^{i-1} - \sum_{i=0}^{n}(i+1)c_i(\ln x)^i\right]. \end{aligned}$$

   Given that $Q(\alpha) = 0$, then $x^2P''(x) + (Q'(\alpha)+1)\,tP'(x) + Q(\alpha)P(x)$ is given by $\sum_{i=0}^{n-1}\left[(i+2)c_{i+1} + Q'(\alpha)c_i\right](\ln x)^i + Q'(\alpha)(n+1)c_n(\ln x)^n$.
   Assuming that α has multiplicity one we have $Q'(\alpha) \neq 0$. Thus, in order to have (4.1), we need to set that $Q'(\alpha)(n+1)c_n = A$ and $(i+2)c_{i+1} + Q'(\alpha)c_i = 0$, for $i = 0, 1, ..., n-1$, and the result follows.

3. If α is a root of Q with multiplicity two, then $P(x) = c_n (\ln x)^{n+2}$ as

$$\begin{aligned} P'(x) &= \frac{1}{x}c_n(n+2)(\ln x)^{n+1}, \\ P''(x) &= \frac{1}{x^2}c_n(n+2)\left[(n+1)(\ln x)^n - (\ln x)^{n+1}\right]. \end{aligned}$$

   Since $Q(\alpha) = 0$ and $Q'(\alpha) = 0$, then $x^2P''(x) + (Q'(\alpha)+1)\,tP'(x) + Q(\alpha)P(x)$ is given by $c_n(n+2)(n+1)(\ln x)^n$. Finally, in order to have (4.1) we conclude that $c_n = \frac{A}{(n+1)(n+2)}$. □

This theorem is useful in two ways: first it provides a way to compute (recursively) the particular solution of the differential equation (2.1). Second, it provides the tool to derive explicit expressions for the involved coefficients. In the following theorem we present such result.

Theorem 4.2 (non-recursive)**.** *Consider the second order ODE presented in (2.1), with the corresponding characteristic polynomial Q given by (3.2).*

- *If α is not a root of Q, the particular solution of (2.1) is given by $y_p(x) = x^\alpha \sum_{i=0}^{n} c_i (\ln x)^i$, with*

$$c_i = (-1)^{n-i}\,\frac{n!}{i!}\,\frac{A}{Q(\alpha)^{n-i+1}}\,\sum_{\substack{j=0\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}} (-1)^j \binom{n-i-j}{j}\, Q'(\alpha)^{n-i-2j}\, Q(\alpha)^j, \tag{4.2}$$

  for $i = 0, 1, 2, ..., n$, where $\binom{k}{r} = \frac{k!}{r!(k-r)!}$, with $k \geq r \geq 0$.

- *If α is a simple root of Q, the particular solution of (2.1) is $y_p(x) = x^\alpha \sum_{i=0}^{n} c_i (\ln x)^{i+1}$, with*

$$c_i = (-1)^{n-i}\,\frac{n!}{(i+1)!}\,\frac{A}{Q'(\alpha)^{n-i+1}},\ \text{for } i = 0, 1, 2, ..., n. \tag{4.3}$$

- *If α is a root of Q with multiplicity two, the particular solution of (2.1) is $y_p(x) = x^\alpha c_n (\ln x)^{n+2}$, with $c_n = \frac{A}{(n+1)(n+2)}$.*

Proof: The last case coincides with the one presented in Theorem 4.1. For the other two cases, we use backwards mathematical induction to prove it, taking advantage of the recursive solutions presented in Theorem 4.1.

1. If α is not a root of Q, we already know that, the particular solution is of the form $y_p(x) = x^\alpha \sum_{i=0}^{n} c_i (\ln x)^i$, where $c_n = \frac{A}{Q(\alpha)}$, $c_{n-1} = -nA\frac{Q'(\alpha)}{Q(\alpha)^2}$ and $c_i = -\frac{i+1}{Q(\alpha)}\left[Q'(\alpha)c_{i+1} + (i+2)c_{i+2}\right]$ for $i = 0, 1, 2, ..., n-2$. We want to prove that, for $i = 0, 1, 2, ..., n$, the coefficients c_i can be written in the general form presented in (4.2).
   Using backwards mathematical induction we have two base cases to be verified, c_n and c_{n-1}, which we know from Theorem 4.1 that are $\frac{A}{Q(\alpha)}$ and $-nA\frac{Q'(\alpha)}{Q(\alpha)^2}$, respectively. Taking into account (4.2), we have

$$\begin{aligned} c_n &= (-1)^0\,\frac{n!}{n!}\,\frac{A}{Q(\alpha)}\,(-1)^0\binom{0}{0}\,Q'(\alpha)^0\,Q(\alpha)^0 = \frac{A}{Q(\alpha)},\\ c_{n-1} &= (-1)\,\frac{n!}{(n-1)!}\,\frac{A}{Q(\alpha)^2}\,(-1)^0\binom{1}{0}\,Q'(\alpha)^1\,Q(\alpha)^0 = -nA\frac{Q'(\alpha)}{Q(\alpha)^2}, \end{aligned}$$

   which means that the base cases are verified. For the inductive step, we assume that, for $i = 0, 1, 2, ..., n-2$, c_{i+1} and c_{i+2} are given by (4.2), and we want to prove that c_i is also given by (4.2).
   From Theorem 4.1, we know that $c_i = -\frac{i+1}{Q(\alpha)}\left[Q'(\alpha)c_{i+1} + (i+2)c_{i+2}\right]$ for $i = 0, 1, 2, ..., n-2$. Plugging the expressions of c_{i+1} and c_{i+2}, which are defined by (4.2), in the expression of c_i we obtain

$$\begin{aligned} -\frac{i+1}{Q(\alpha)}\Bigg[& Q'(\alpha)(-1)^{n-i-1}\frac{n!}{(i+1)!}\frac{A}{Q(\alpha)^{n-i}}\sum_{\substack{j=0\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}-\frac{1}{2}}(-1)^j\binom{n-i-j-1}{j}Q'(\alpha)^{n-i-2j-1}Q(\alpha)^j\\ & +(i+2)(-1)^{n-i-2}\frac{n!}{(i+2)!}\frac{A}{Q(\alpha)^{n-i-1}}\sum_{\substack{j=0\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}-1}(-1)^j\binom{n-i-j-2}{j}Q'(\alpha)^{n-i-2j-2}Q(\alpha)^j\Bigg]. \end{aligned}$$

Rearranging the terms and changing the variable in the second sum, we get

$$(-1)^{n-i}\frac{n!}{i!}\frac{A}{Q(\alpha)^{n-i+1}}\left[\sum_{\substack{j=0\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}-\frac{1}{2}}(-1)^j\binom{n-i-j-1}{j}Q'(\alpha)^{n-i-2j}\,Q(\alpha)^j\right.$$
$$\left.+\sum_{\substack{j=1\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}}(-1)^j\binom{n-i-j-1}{j-1}Q'(\alpha)^{n-i-2j}\,Q(\alpha)^j\right].$$

Joining the two sums and taking into account some permutation's properties, we end up with the following expression:

$$(-1)^{n-i}\frac{n!}{i!}\frac{A}{Q(\alpha)^{n-i+1}}\left[\sum_{\substack{j=1\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}-\frac{1}{2}}(-1)^j\binom{n-i-j}{j}Q'(\alpha)^{n-i-2j}\,Q(\alpha)^j\right.$$
$$\left.+\,Q'(\alpha)^{n-i}+(-1)^{\frac{n-i}{2}}\,Q(\alpha)^{\frac{n-i}{2}}\chi_{\{n-i\text{ is even}\}}\right].$$

Finally, we conclude that

$$c_i\;=\;(-1)^{n-i}\frac{n!}{i!}\frac{A}{Q(\alpha)^{n-i+1}}\sum_{\substack{j=0\\ j\in\mathbb{N}_0}}^{\frac{n-i}{2}}(-1)^j\binom{n-i-j}{j}Q'(\alpha)^{n-i-2j}\,Q(\alpha)^j,$$

which coincides with the expression given by (4.2). Thus the proof for the first case is finished.

2. If α is a root of Q with multiplicity one, as we proved before, the particular solution is of the form $y_p(x)=x^\alpha\sum_{i=0}^n c_i\,(\ln x)^{i+1}$, where $c_n=\frac{A}{(n+1)Q'(\alpha)}$ and $c_i=-\frac{i+2}{Q'(\alpha)}c_{i+1}$, for $i=0,1,2,...,n-1$. We want to prove that we can write the coefficients c_i in the general way presented in (4.3).
As before, we use backwards mathematical induction. Starting with the base case and taking into account (4.3), we have

$$c_n=(-1)^0\,\frac{n!}{(n+1)!}\,\frac{A}{Q'(\alpha)}=\frac{A}{(n+1)Q'(\alpha)},$$

which coincides with the expression given by Theorem 4.1. Thus, the base case is verified. To prove the induction step, for $i=0,1,2,...,n-1$, we assume that c_{i+1} is given by (4.3) and we want to prove that c_i is also given by (4.3).
From Theorem 4.1, we know that $c_i=-\frac{i+2}{Q'(\alpha)}c_{i+1}$, for $i=0,1,2,...,n-1$. Plugging in c_i the expression of c_{i+1}, which is given by (4.3), we obtain

$$c_i=-\frac{i+2}{Q'(\alpha)}\,(-1)^{n-i-1}\,\frac{n!}{(i+2)!}\,\frac{A}{Q'(\alpha)^{n-i}}=(-1)^{n-i}\,\frac{n!}{(i+1)!}\,\frac{A}{Q'(\alpha)^{n-i+1}},$$

and therefore the induction step is proved. With this we conclude the proof. □

A special case of the previous theorem is when $n = 0$. In this case the Equation (2.1) is simply

$$x^2 y''(x) + axy'(x) + by(x) = Ax^\alpha.$$

Using the results proved before, the corresponding particular solution is given by

$$y_p(x) = \varphi\, x^\alpha \,(\ln x)^r \,,$$

where $\varphi = \frac{A}{Q^{(r)}(\alpha)}$[6], with r[7] being the multiplicity of α as a root of Q.

In the following corollary, we use the results presented in Theorem 4.2 for the case that the non-homogeneous part of the differential equation is a sum of power and log functions (as it is the case, for example, of (3.8)).

Corollary 4.1. *Consider the following second order differential equation:*

$$x^2 y''(x) + axy'(x) + by(x) = \sum_{k=1}^{m} A_k x^{\alpha_k} (\ln x)^{n_k} \,, \tag{4.4}$$

with $x > 0$, $a, b \in \mathbb{R}$, $\alpha_k, A_k \in \mathbb{R} \setminus \{0\}$ and $n_k \in \mathbb{N}_0$, for $k = 1, 2, ..., m$, with $m \in \mathbb{N}$. Then the particular solution of (4.4) is of the form $y_p(x) = \sum_{k=1}^m y_{p_k}(x)$[8], where $y_{p_k}(x)$ is the solution of the equation

$$x^2 y_k''(x) + axy_k'(x) + by_k(x) = A_k x^{\alpha_k} (\ln x)^{n_k} \,,$$

which is presented in Theorem 4.2.

Proof: The result follows from the superposition principle. □

5. CONCLUSIONS

In this paper we provide a solution to a differential–difference equation that can be found, for instance, when one studies an investment problem with the underlying following a jump–diffusion process. This problem is particularly important from the point of view of the application, as nowadays the prices and demand are often subject to external shocks that cause a disruptive behavior on the state variables. Analytical solutions or quasi-analytical solutions are scarce or even non-existent. Our results contribute to the state of the art in this area.

As our results show, the solution to the differential–difference equation is a piecewise function, where each branch depends on the next one. Therefore, to find the expression for each branch a non-homogeneous ODE needs to be solved. In this paper we also provide the expression for each coefficient involved in the particular solution of this family of ODEs.

[6] $Q^{(r)}(\alpha)$ is the derivative of order r of Q w.r.t. α. In particular, if $r = 0$ we consider that $Q^{(r)}(\alpha)$ is exactly $Q(\alpha)$.

[7] r can take the values $0, 1$ or 2. We consider $r = 0$ when α is not a root of Q.

[8] Note that y_p has at least m parcels and at most $m + \sum_{k=1}^m n_k$ parcels. When $\alpha_1 = \alpha_2 = ... = \alpha_m$ are roots of Q all with multiplicity two, y_p has m parcels. Oppositely, when none of the α_k (with $k = 1, 2, ..., m$) has multiplicity two, y_p has $m + \sum_{k=1}^m n_k$ parcels.

As future work, we want to apply these results to solve the original optimal stopping problem. We highlight that this is a challenging question, as in order to find the optimal value function, we need to use enough conditions to define all the unknown parameters of the solution. Indeed, the expressions that we provide in this paper define classes of solutions, and only considering the boundary and initial conditions we are able to derive the solution.

ACKNOWLEDGMENTS

Most of Rita Pimentel's research was supported by her Ph.D. research grant from FCT with reference number SFRH/BD/97259/2013. The final part of the work was carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship Programme.

Most of Cláudia Nunes's research was supported by the FCT research grants FARO_PTDC/EGE-ECO/30535/2017 and UID/MULTI/04621/2019. She also acknowledge Invest ExL Project No.: 268093/E20, funded by the Norwegian research agency.

The authors also thank to Carlos Oliveira for the discussions and all invaluable suggestions.

REFERENCES

[1] Applebaum, D. (2004). *Lévy processes and stochastic calculus*, Cambridge University Press.

[2] Arino, O.; Hbid, M.L. and Dads, A. (2007). *Delay differential equations and applications.* In "Proceedings of the NATO Advanced Study Institute", Marrakech, Morocco, 9–21 September 2002, Springer Science & Business Media.

[3] Balcer, Y. and Lippman, S. (1984). Technological expectations and adoption of improved technology, *Journal of Economic Theory*, **34**(2), 292–318.

[4] Bass, R. (2004). Stochastic differential equations with jumps, *Probab. Surveys*, **1**, 1–19.

[5] Black, F. and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal Of Political Economy*, **81**(3), 637–654.

[6] Boomsma, T. and Linnerud, K. (2015). Market and policy risk under different renewable electricity support schemes, *Energy*, **89**, 435–448.

[7] Boomsma, T.; Meade, N. and Fleten, S.-E. (2012). Renewable energy investments under different support schemes: a real options approach, *European Journal Of Operational Research*, **220**(1), 225–237.

[8] Boyce, W. and DiPrima, R. (2008). *Elementary Differential Equations and Boundary Value Problems*, 9th edition, Wiley.

[9] Chen, G. and Wang, D. (2002). The Cauchy problem for the Euler equations for compressible fluids, *Handbook Of Mathematical Fluid Dynamics*, **1**, 421–543.

[10] Chern, H.; Hwang, H. and Tsai, T. (2002). An asymptotic theory for Cauchy–Euler differential equations with applications to the analysis of algorithms, *J. Algorithms*, **44**(1), 177–225.

[11] Couto, G.; Nunes, C. and Pimentel, P. (2015). High-speed rail transport valuation and conjecture shocks, *The European Journal of Finance*, **21**, 791–805.

[12] Dixit, A. and Pindyck, R. (1994). *Investment Under Uncertainty*, Princeton University Press.

[13] Farzin, Y.; Huisman, K. and Kort, P. (1998). Optimal timing of technology adoption, *Journal Of Economic Dynamics & Control*, **22**(5), 779–799.

[14] Hagspiel, V.; Dalby, P.; Leth-Olsen, T. and Thijssen, J. (2018). Green investment under policy uncertainty and Bayesian learning, *Energy*, **161**, 1262–1281.

[15] Hagspiel, V.; Huisman, K.; Kort, P. and Nunes, C. (2016). How to escape a declining market: capacity investment or exit?, *European Journal of Operational Research*, **254**(1), 40–50.

[16] Hagspiel, V.; Huisman, K. and Nunes, C. (2015). Optimal technology adoption when the arrival rate of new technologies changes, *European Journal Of Operational Research*, **243**(3), 897–911.

[17] Huisman, K. (2013). *Technology Investment: A Game Theoretic Real Options Approach*, "Theory and Decision Library", Series C, Vol. 28, Springer Science & Business Media.

[18] Johansen, A. and Sornette, D. (2010). Shocks, crashes and bubbles in financial markets, *Brussels Economic Review*, **53**(2), 201–253.

[19] Johnson, T. and Zervos, M. (2007). The solution to a second order linear ordinary differential equation with a non-homogeneous term that is a measure, *Stochastics – An International Journal of Probability and Stochastic Processes*, **79**(3–4), 363–382.

[20] Kohler, W. and Johnson, L. (2006). *Elementary Differential Equations*, 2nd edition, Addison-Wesley.

[21] Kou, S. (2007). Jump–diffusion models for asset pricing in financial engineering, *Handbooks in Operations Research and Management Science*, **15**, 73–116.

[22] Kwon, D. (2010). Invest or exit? Optimal decisions in the face of a declining profit stream, *Operations Research*, **58**(3), 638–649.

[23] Matomäki, P. (2013). *On Two-sided Controls of a Linear Diffusion*, Turku School of Economics.

[24] Merton, R. (1976). Option pricing when underlying stock returns are discontinuous, *Journal Of Financial Economics*, **3**(1–2), 125–144.

[25] Murto, P. (2007). Timing of investment under technological and revenue-related uncertainties, *Journal of Economic Dynamics And Control*, **31**(5), 1473–1497.

[26] Nunes, C. and Pimentel, R. (2017). Analytical solution for an investment problem under uncertainties with shocks, *European Journal of Operational Research*, **259**(3), 1054–1063.

[27] Øksendal, B. and Sulem-Bialobroda, A. (2009). *Applied Stochastic Control of Jump Diffusions*, 3rd edition, Springer, Berlin, Germany.

[28] Peskir, G. and Shiryaev, A. (2006). Optimal Stopping and Free-Boundary Problems, Springer.

[29] Sabuwala, A. and De Leon, D. (2011). Particular solution to the Euler–Cauchy equation with polynomial non-homogeneities, *Discrete And Continuous Dynamical Systems*, Supplement 2011, 1271–1278.

[30] Sennewald, K. and Wälde, K. (2006). "Ito's lemma" and the Bellman equation for Poisson processes: an applied view, *Journal Of Economics*, **89**(1), 1–36.

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 101–115

FLEXIBLE ROBUST MIXTURE REGRESSION MODELING

Authors: Marcus G. Lavagnole Nascimento
– Departamento de Métodos Estatísticos, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
marcus@dme.ufrj.br

Carlos A. Abanto-Valle
– Departamento de Métodos Estatísticos, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil
cabantovalle@im.ufrj.br

Received: April 2019 Revised: December 2019 Accepted: December 2019

Abstract:

- This paper provides a flexible methodology for the class of finite mixture of regressions with scale mixture of skew-normal errors (SMSN-FMRM) introduced by [42], relaxing the constraints imposed by the authors during the estimation process. Based on the data augmentation principle and Markov chain Monte Carlo (MCMC) algorithms, a Bayesian inference procedure is developed. A simulation study is implemented in order to understand the possible effects caused by the restrictions and an example with a well known dataset illustrates the performance of the proposed methods.

Keywords:

- *finite mixture of regressions; scale mixture of skew-normal distributions; Markov chain Monte Carlo.*

AMS Subject Classification:

- 62F15, 62J05, 62H30.

1. INTRODUCTION

Finite mixture regression models (FMRM) provide a flexible tool for modeling data that arise from a heterogeneous population, where a single regression model is not enough for capturing the complexities of the conditional distribution of the observed sample given the features. FMRM of Gaussian distributions, using maximum likelihood methods for parameter estimation, have been extensively used in the literature in different fields like marketing [11, 12], economics [10, 21], agriculture [36], psychometrics [28], among others.

From a Bayesian perspective, there is a wide range of nonparametric methods, in particular, methods in which the error follows a mixture of Dirichlet process [27] or a mixture of Polya trees [22]. However, in comparison with these methodologies, the finite mixture of regressions presents the advantage of classifying the observations over the components of the mixture in a natural way. This classification, in a range of applications, is the main topic of interest and provides for practitioners a clear interpretation of the results, besides facilitating the implementation.

Extensions of FMRM of Gaussian distributions have been proposed to broaden the applicability of the model to more general structures like skewed or heavy tailed errors. In this regard, [4] modified the EM algorithm for normal mixtures, replacing the least squares criterion in the M step with a robust one. [33] and [41], in turn, implemented an estimation procedure for finite mixture of linear regression models assuming that the error terms follow a Laplace and a Student-t distribution, respectively. As an attempt to accommodate asymmetric observations, [29] introduced a FMRM based on skew-normal distributions [1].

More recently, as an attractive way to deal with skewness and heavy tails simultaneously, [42] introduced a finite mixture regression model based on scale mixtures of skew-normal distributions [6, SMSN] as follow:

$$f(y_i|\mathbf{x}_i,\boldsymbol{\vartheta},\boldsymbol{\eta}) = \sum_{j=1}^{G} \eta_j g(y_i|\mathbf{x}_i,\boldsymbol{\theta}_j), \tag{1.1}$$

where the probability density function $g(\cdot|\mathbf{x}_i,\boldsymbol{\theta}_j)$ comes from the same member of the $\mathrm{SMSN}(\mathbf{x}_i\boldsymbol{\beta}_j+\mu_j,\sigma_j^2,\lambda_j,\nu_j)$ family, $\boldsymbol{\theta}_j=(\boldsymbol{\beta}_j,\sigma_j^2,\lambda_j,\nu_j)$ is the specific parametric vector for the component j, $\eta_j>0$, $j=1,...,G$, $\sum_{j=1}^{G}\eta_j=1$, $\boldsymbol{\vartheta}$ and $\boldsymbol{\eta}$ denote the unknown parameters with $\boldsymbol{\vartheta}=(\boldsymbol{\theta}_1,...,\boldsymbol{\theta}_G)$ and $\boldsymbol{\eta}=(\eta_1,...,\eta_G)$. However, [42] impose the constraints $\tau_1^2=\cdots=\tau_G^2$ and $\nu_1=\cdots=\nu_G$ about the parameters during the estimation procedure in which $\tau_j^2=\sigma_j^2(1-\delta_j^2)$ and $\delta_j=\lambda_j/(\sqrt{1+\lambda_j^2})$.

The aim of this paper, therefore, is to provide a flexible version for the mixture of regressions based on scale mixtures of skew-normal distributions introduced by [42], relaxing the restrictions described above and verifying empirically how our ideas improve the estimation process. Bayesian inference is developed applying ideas like the data augmentation principle, stochastic representation in terms of a random-effects model [2, 23], standard hierarchical representation of a finite mixture model [14] and MCMC methods.

The remainder of the paper is organized as follows. Section 2 is related to the development of a flexible methodology for the mixture regression model based on scale mixture of skew-normal (SMSN-FMRM) distributions from a Bayesian perspective. In order to make comparisons between the methodology proposed in the present work and the one proposed by [42] feasible, Sections 3 and 4 present the analysis of a simulation study and a real dataset respectively. Finally, some concluding remarks and suggestions for future developments are given in Section 5.

2. MIXTURE REGRESSION MODEL BASED ON SCALE MIXTURE OF SKEW-NORMAL DISTRIBUTIONS

2.1. The model

Let $\mathbf{y} = (y_1, ..., y_n)^T$ given $\mathbf{x} = (\mathbf{x}_1^T, ..., \mathbf{x}_n^T)^T$ be a random sample from a G-component mixture model, $\mathbf{x}_i$ is a p-dimensional vector of explanatory variables, and consider a mixture regression model in which the random errors follow a scale mixtures of skew-normal distributions (SMSN-FMRM) as defined by the equation 1.1. Let $\mathbf{S} = (\mathbf{S}_1, ..., \mathbf{S}_n)$ be the allocation vector, i. e., the vector containing the information about in which group the observation y_i of the random variable Y_i is. The indicator variable $\mathbf{S}_i = (S_{i1}, ..., S_{iG})^T$, with

$$S_{ij} = \begin{cases} 1, & \text{if } Y_i \text{ belongs to component } j \\ 0, & \text{otherwise} \end{cases}$$

and $\sum_{j=1}^{G} S_{ij} = 1$. Given the weights vector $\boldsymbol{\eta}$, the latent variables $\mathbf{S}_1, ..., \mathbf{S}_n$ are independent with multinomial distribution

$$p(\mathbf{S}_i|\boldsymbol{\eta}) = \eta_1^{S_{i1}} \eta_2^{S_{i2}} \cdots (1 - \eta_1 - \cdots - \eta_{G-1})^{S_{iG}}.$$

The joint density of $\mathbf{Y} = (Y_1, ..., Y_n)$ and $\mathbf{S} = (\mathbf{S}_1, ..., \mathbf{S}_n)$ is given by

$$f(\mathbf{y}, \mathbf{s}|\mathbf{x}, \boldsymbol{\vartheta}, \boldsymbol{\eta}) = \prod_{j=1}^{G} \prod_{i=1}^{n} \left[\eta_j g(y_i|\mathbf{x}_i, \boldsymbol{\theta}_j)\right]^{S_{ij}}.$$

From the stochastic representation in terms of a random-effects model introduced by [2] and [23], a random variable drawn from the scale mixture of skew-normal distributions has a hierarchical representation. Hence, the individual Y_i belonging to the j-th component can be written as

$$\begin{aligned} Y_i|S_{ij} = 1, \mathbf{x}_i, w_i, u_i, \boldsymbol{\theta}_j &\sim N(\mathbf{x}_i\boldsymbol{\beta}_j + \mu_j + \sigma_j\delta_j w_i, k(u_i)\sigma_j\sqrt{1-\delta_j^2}), \\ W_i|S_{ij} = 1, u_i &\sim TN_{[0,+\infty)}(0, k(u_i)), \\ U_i|S_{ij} = 1, \nu_j &\sim h(\cdot; \nu_j), \end{aligned}$$

where $\mu_j = -\sqrt{\frac{2}{\pi}} m_{1,j}\sigma_j\delta_j$, $m_1 = E[U^{-1/2}]$, which corresponds to the regression model where the error distribution has zero mean and hence the regression parameters are all comparable.

Thus, the joint density of $\mathbf{Y}$ and the latent variables $\mathbf{S}$, $\mathbf{W}$ and $\mathbf{U}$ is

$$f(\mathbf{y},\mathbf{s},\mathbf{w},\mathbf{u}|\mathbf{x},\boldsymbol{\vartheta},\boldsymbol{\eta}) = \prod_{j=1}^{G}\left[\prod_{i=1}^{n}[\eta_j f(y_i|\boldsymbol{\theta}_j,\mathbf{x}_i,w_i,u_i)f(w_i|u_i)f(u_i|\nu_j)]^{S_{ij}}\right].$$

In this article, $k(U) = U^{-1}$ is used since it leads to good mathematical properties. Without loss of generality, the distributions skew normal [1, SN], skew-t [3, ST] and skew-slash [39, SSL] are considered here, it means that mixing variables are chosen as: $U = 1$, $U \sim G(\frac{\nu}{2},\frac{\nu}{2})$ and $U \sim Be(\nu,1)$, where $G(\cdot,\cdot)$ and $Be(\cdot,\cdot)$ indicate the gamma and beta distributions respectively.

As in [17], we introduce a new parameterization in terms of the component-specific parameters $\boldsymbol{\theta}_j^* = (\boldsymbol{\beta}_j,\psi_j,\tau_j^2,\nu_j)$, where $\psi_j = \sigma_j\delta_j$ and $\tau_j^2 = \sigma_j^2(1-\delta_j^2)$. The original parametric vector $\boldsymbol{\theta}_j = (\boldsymbol{\beta}_j,\sigma_j^2,\lambda_j,\nu_j)$, on its turn, is recovered through

$$\lambda_j = \frac{\psi_j}{\tau_j},\ \ \sigma_j^2 = \tau_j^2 + \psi_j^2,$$

since $\psi_j/\tau_j = \sigma_j\delta_j/(\sigma_j\sqrt{1-\delta_j^2}) = \lambda_j$ and $\tau_j^2+\psi_j^2 = \sigma_j^2(1-\delta_j^2)+\sigma_j^2\delta_j^2 = \sigma_j^2$.

2.2. Bayesian inference

Performing a Bayesian analysis, an important step is the priors distributions selection. In the context of finite mixture models, in particular, mixture regression models, a special attention on these choices is quite relevant since it is not possible to choose an improper prior because it implies in an improper posterior density [16]. In addition, as pointed by [25], it is recommended to avoid be as "noninformative as possible" by choosing large prior variances because the number of components is highly influenced by the prior choices. Consequently, in order to avoid identifiability problems, it was adopted the hierarchical priors introduced by [31] for mixtures of normal distributions to reduce sensitivity with respect to choosing the prior variances.

Hence, considering the parametric vector $\boldsymbol{\theta}_j^* = (\boldsymbol{\beta}_j,\psi_j,\tau_j^2,\nu_j)$ for an arbitrary mixture component j, the prior set was specified as: $\boldsymbol{\eta} \sim D(e_0,...,e_0)$, $(\boldsymbol{\beta}_j,\psi_j)|\tau_j^2 \sim N_{p+1}(\mathbf{b}_0,\tau_j^2\mathbf{B}_0)$, $\tau_j^2|C_0 \sim IG(c_0,C_0)$ and $C_0 \sim G(h_0,H_0)$, where e_0, $\mathbf{b}_0 \in \mathbb{R}^{(p+1)}$, $\mathbf{B}_0 \in \mathbb{R}^{(p+1)\times(p+1)}$, c_0, h_0 and H_0 are known hyper parameters, $N_q(\cdot,\cdot)$, $D(\cdot,...,\cdot)$ and $IG(\cdot,\cdot)$ indicate the q-variate normal, the Dirichlet and inverse gamma distributions. Considering the parameter ν priors, $p(\nu_j) \propto \nu_j/(\nu_j+d)^3\mathbb{1}_{(2,40)}(\nu_j)$ [26] and $\nu_j \sim G_{(1,40)}(\alpha,\gamma)$, where α and γ are known hyper parameters and $G_A(\cdot,\cdot)$ denotes the truncated gamma on set A, are specified for the ST-FMRM and SSL-FMRM respectively.

The Bayesian approach for estimating the parameters uses the data augmentation principle [35], which considers $\mathbf{W},\mathbf{U}$ and $\mathbf{S}$ as latent unobserved variables. The joint posterior density of parameters and latent variables can be written as

$$p(\boldsymbol{\vartheta}^*,\boldsymbol{\eta},\mathbf{w},\mathbf{u},\mathbf{s}|\mathbf{y},\mathbf{x}) \propto \left\{\prod_{j=1}^{G}\left[\prod_{i=1}^{n}\left[\eta_j f(y_i|\boldsymbol{\theta}_j^*,\mathbf{x}_i,w_i,u_i)f(w_i|u_i)f(u_i|\nu_j)\right]^{S_{ij}}\right]p(\boldsymbol{\theta}_j^*)\right\}p(\boldsymbol{\eta}),$$

where $p(\boldsymbol{\theta}_j^*) = p(\boldsymbol{\beta}_j, \psi_j|\tau_j^2)p(\tau_j^2|C_0)p(C_0)p(\nu_j)$ and $\boldsymbol{\vartheta}^* = (\boldsymbol{\theta}_1^*, ..., \boldsymbol{\theta}_G^*)$. In light of the data augmentation technique, conditional on the allocation vector $\mathbf{S}$, the parameters estimation may be executed independently for each parametric component $\boldsymbol{\theta}_j^*$ and for the weights distribution $\boldsymbol{\eta}$. As a consequence, the full conditionals of the parameters and the latent unobserved variables for the mixture regression models based on the SMSN distributions are written as follows:

$$p(\boldsymbol{\eta}|\mathbf{s}) \propto p(\mathbf{s}|\boldsymbol{\eta})p(\boldsymbol{\eta}) \tag{2.1}$$

$$p(w_i|S_{ij}=1,\cdots) \propto \left[f(y_i|\boldsymbol{\theta}_j^*, \mathbf{x}_i, w_i, u_i)f(w_i|u_i)\right]^{S_{ij}}, \tag{2.2}$$

$$p(u_i|S_{ij}=1,\cdots) \propto \left[f(y_i|\boldsymbol{\theta}_j^*, \mathbf{x}_i, w_i, u_i)f(w_i|u_i)f(u_i|\nu_j)\right]^{S_{ij}}, \tag{2.3}$$

$$p(\boldsymbol{\beta}_j, \psi_j|\cdots) \propto \prod_{\{i:S_{ij}=1\}} f(y_i|\boldsymbol{\theta}_j^*, \mathbf{x}_i, w_i, u_i)p(\boldsymbol{\beta}_j, \psi_j|\tau_j^2), \tag{2.4}$$

$$p(\tau_j^2|\cdots) \propto \prod_{\{i:S_{ij}=1\}} f(y_i|\boldsymbol{\theta}_j^*, \mathbf{x}_i, w_i, u_i)p(\tau_j^2|C_0), \tag{2.5}$$

$$p(C_0|\cdots) \propto \prod_{j=1}^{G} p(\tau_j^2|C_0)p(C_0), \tag{2.6}$$

$$p(\nu_j|\cdots) \propto \prod_{\{i:S_{ij}=1\}} f(u_i|\nu_j)p(\nu_j). \tag{2.7}$$

Additional details about the derivations of the full conditionals are available in Appendix A.1.

In furtherance of making Bayesian analysis feasible for parameter estimation in the SMSN-FMRM class of models, random samples from the posterior distributions of $(\boldsymbol{\vartheta}, \boldsymbol{\eta}, \mathbf{w}, \mathbf{u}, \mathbf{s})$ given $(\mathbf{y}, \mathbf{x})$ are drawn through Monte Chain Monte Carlo simulation methods. Algorithm 1 describes the sampling scheme from the full conditionals distributions of the parameters and the latent unobserved variables.

Algorithm 1. MCMC for finite mixture of scale mixtures of skew-normal.

1. Set $k = 1$ and get starting values for $\mathbf{S}^{(0)}$, $(\boldsymbol{\theta}_1^{*(0)}, ..., \boldsymbol{\theta}_G^{*(0)})$, $\boldsymbol{\eta}^{(0)}$, $\mathbf{w}^{(0)}$ and $\mathbf{u}^{(0)}$;

2. Parameter simulation conditional on the classification $\mathbf{S}^{(k-1)}$:

  2.1. Sample $\boldsymbol{\eta}^{(k)}$ from $p(\boldsymbol{\eta}|\mathbf{s}^{(k-1)})$;

  2.2. Sample the component latent variables $w_i^{(k)}$ and $u_i^{(k)}$, $i = 1, ..., n$, from the full conditionals (2.2)–(2.3) and the component parameters $\boldsymbol{\beta}_j^{(k)}, \psi_j^{(k)}, \tau_j^{2(k)}, \nu_j^{(k)}$, $j = 1, ..., G$, from the full conditionals (2.4)–(2.7).

3. Sample $S_i^{(k)}$ independently for each $i = 1, ..., n$ from

$$\Pr(S_{il}=1|y_i, \mathbf{x}_i, \boldsymbol{\vartheta}^*) = \frac{g(y_i|\mathbf{x}_i, \boldsymbol{\theta}_l^*)\Pr(S_{il}=1|\boldsymbol{\vartheta}^*)}{\sum_{j=1}^{G} g(y_i|\mathbf{x}_i, \boldsymbol{\theta}_j^*)\Pr(S_{ij}=1|\boldsymbol{\vartheta}^*)}.$$

4. Set $k = k + 1$ and repeat the steps $2, 3$ and 4 until convergence is achieved.

Introduced by [30] into the mixture models background, the term *label switching* refers to the invariance of the mixture likelihood function under relabeling the components. Considering the maximum likelihood estimation, where we are looking for the corresponding modes of the likelihood function, label switching is not an object of interest. From the Bayesian point of view, however, it is a topic of concern because the labeling of the unobserved categories changes during the sample process of the mixture posterior distribution. Post-processed the MCMC, in order to deal with the label switching problem, the Kullback–Leibler algorithm [34] is applied over this paper.

3. SIMULATION STUDY

In this section, a simulated scenario is considered for three purposes:

- **(i)** verifying if the true parameter values are recovered accurately by using the methodology described on Section 2;
- **(ii)** comparing the estimation performance of the unconstrained and constrained models;
- **(iii)** formulating a sensitivity analysis study to the hyperparameters specification.

To that end, datasets are artificially generated as follow:

$$\begin{cases} Y_i = \mathbf{x}_i\boldsymbol{\beta}_1 + \varepsilon_1, & S_{i1} = 1, \\ Y_i = \mathbf{x}_i\boldsymbol{\beta}_2 + \varepsilon_2, & S_{i2} = 1, \end{cases}$$

where S_{ij} is a component indicator of Y_i with $\Pr(S_{ij} = 1) = \eta_j$, $j = 1, 2$, $\mathbf{x}_i = (1, x_{i1})$, $i = 1, ..., n$. Finally, ε_1 and ε_2 follow a distribution in the SMSN family. According to this procedure, 100 random samples of size $n = 500$ are generated from the SN-FMRM, ST-FMRM and SSL-FMRM models with the following parameter values: $\boldsymbol{\beta}_1 = (\beta_{01}, \beta_{11})^T = (20, 0)^T$, $\boldsymbol{\beta}_2 = (\beta_{02}, \beta_{12})^T = (-4, 3)^T$, $\sigma_1^2 = 1$, $\sigma_2^2 = 4$, $\lambda_1 = 0$, $\lambda_2 = 5$, $\eta_1 = 0.4$, $\eta_2 = 0.6$. In addition, for the ST-FMRM and SSL-FMRM models, $\boldsymbol{\nu} = (\nu_1, \nu_2) = (8, 3)$ and $\boldsymbol{\nu} = (6, 2)$, respectively.

During the estimation process for the SMSN-FMRM models, the unconstrained version proposed in this paper and the constrained version of [42] were considered and it was adopted the four different hyperparameters specifications described in Table 1 for both. For each sample, 20000 iterations from Algorithm 1 were conducted. The first 10000 were discarded as a burn-in period. In order to reduce the autocorrelation within the successive values of the simulated chain, it was required a thin equals to 10. Finally, based on 1000 records, the posterior mean were obtained.

Table 1: Prior sets hyperparameters specifications.

Specification	e_0	$\mathbf{b}_0$	$\mathbf{B}_0$	c_0	h_0	H_0	d	α	γ
P_1	4	(0,0,0)	Diag(100,100,100)	0.01	0.01	0.01	$4/(1+\sqrt{2})$	6	1
P_2	4	(0,0,0)	Diag(10,10,10)	0.01	0.01	0.01	$4/(1+\sqrt{2})$	6	1
P_3	4	(0,0,0)	Diag(100,100,100)	2.5	0.75	$\frac{0.75}{0.5s_y^2}$	$4/(1+\sqrt{2})$	6	1
P_4	4	(0,0,0)	Diag(100,100,100)	0.01	0.01	0.01	$9/(1+\sqrt{2})$	4	1

Table 2: MSE and coverage percentage in parenthesis for the MCMC estimates based on the 100 samples from the SMSN-FMRM.

Parameters		SN-FMRM		ST-FMRM		SSL-FMRM	
		$\tau_1^2 \neq \tau_2^2$	$\tau_1^2 = \tau_2^2$	$\tau_1^2 \neq \tau_2^2, \nu_1 \neq \nu_2$	$\tau_1^2 = \tau_2^2, \nu_1 = \nu_2$	$\tau_1^2 \neq \tau_2^2, \nu_1 \neq \nu_2$	$\tau_1^2 = \tau_2^2, \nu_1 = \nu_2$
$\beta_{0,1}$	P_1	0.0143(1.00)	0.0148(0.99)	0.0221(0.99)	0.0271(0.96)	0.0234(1.00)	0.0458(0.97)
	P_2	0.0222(0.98)	0.0225(0.97)	0.0311(0.98)	0.0293(1.00)	0.0426(0.98)	0.0497(0.97)
	P_3	0.0253(1.00)	0.0272(0.98)	0.0286(0.99)	0.0364(0.94)	0.0312(0.99)	0.0434(0.97)
	P_4	—	—	0.0228(0.99)	0.0284(0.98)	0.0378(0.99)	0.0499(0.97)
$\beta_{1,1}$	P_1	0.0000(0.97)	0.0001(0.98)	0.0001(0.96)	0.0001(0.96)	0.0001(0.93)	0.0001(0.94)
	P_2	0.0001(0.95)	0.0001(0.93)	0.0002(0.95)	0.0002(0.92)	0.0002(0.95)	0.0002(0.92)
	P_3	0.0001(0.94)	0.0001(0.89)	0.0002(0.95)	0.0002(0.95)	0.0001(0.97)	0.0001(0.94)
	P_4	—	—	0.0001(0.99)	0.0001(0.94)	0.0002(0.91)	0.0002(0.92)
$\beta_{0,2}$	P_1	0.0142(0.94)	0.0170(0.97)	0.0454(0.84)	0.0545(0.84)	0.1869(0.91)	0.1932(0.84)
	P_2	0.0156(0.94)	0.0204(0.99)	0.0351(0.94)	0.0697(0.83)	0.1461(0.91)	0.2866(0.65)
	P_3	0.0157(0.93)	0.0163(0.96)	0.0369(0.91)	0.0477(0.90)	0.1502(0.91)	0.1502(0.90)
	P_4	—	—	0.0316(0.90)	0.0429(0.88)	0.1708(0.96)	0.1170(0.95)
$\beta_{1,2}$	P_1	0.0000(0.94)	0.0001(0.96)	0.0001(0.97)	0.0001(0.99)	0.0001(0.90)	0.0001(0.92)
	P_2	0.0000(0.96)	0.0001(0.95)	0.0001(0.97)	0.0001(0.99)	0.0001(0.99)	0.0001(0.99)
	P_3	0.0000(0.95)	0.0000(0.98)	0.0001(0.97)	0.0001(0.96)	0.0001(0.98)	0.0001(0.99)
	P_4	—	—	0.0001(0.90)	0.0001(0.98)	0.0001(0.95)	0.0001(0.96)
σ_1^2	P_1	0.0956(0.95)	0.9566(0.34)	0.0523(0.99)	0.0756(0.98)	0.0943(0.99)	0.4890(0.77)
	P_2	0.1233(0.37)	0.0823(0.89)	0.2337(0.17)	0.0335(0.98)	0.1374(0.42)	0.0612(0.96)
	P_3	0.1385(0.90)	0.8905(0.29)	0.0600(0.99)	0.1026(0.95)	0.1015(0.98)	0.4174(0.84)
	P_4	—	—	0.0593(0.98)	0.1311(0.95)	0.0348(1.00)	0.2948(0.89)
σ_2^2	P_1	0.1760(0.91)	0.5010(0.62)	1.3980(0.84)	0.7495(0.86)	2.5937(0.84)	1.5439(0.85)
	P_2	0.2358(0.82)	1.9505(0.09)	0.7075(0.90)	0.5111(0.94)	2.3527(0.72)	1.8668(0.74)
	P_3	0.2076(0.89)	0.5872(0.55)	0.9655(0.87)	0.7671(0.85)	2.3347(0.80)	1.3687(0.88)
	P_4	—	—	0.9523(0.91)	0.7049(0.88)	1.4102(0.92)	0.9150(0.88)
λ_1	P_1	0.0648(1.00)	2.4698(0.54)	0.0622(1.00)	0.6500(0.85)	0.1128(1.00)	1.7950(0.62)
	P_2	0.0122(1.00)	0.0820(1.00)	0.0128(1.00)	0.0589(1.00)	0.0142(1.00)	0.0965(1.00)
	P_3	0.1465(1.00)	2.3544(0.48)	0.0781(0.99)	0.6287(0.89)	0.1241(1.00)	1.2843(0.78)
	P_4	—	—	0.0620(1.00)	0.7134(0.83)	0.0547(1.00)	1.3541(0.72)
λ_2	P_1	1.6120(0.96)	6.1614(0.00)	3.1617(0.98)	6.3709(0.00)	2.3855(0.94)	4.1220(0.04)
	P_2	1.8628(0.52)	14.1231(0.00)	0.7802(0.92)	10.7803(0.00)	0.5909(0.94)	8.9064(0.00)
	P_3	1.0375(0.86)	6.6829(0.00)	0.9961(0.96)	6.6883(0.01)	0.7847(0.97)	4.6046(0.02)
	P_4	—	—	3.2051(0.96)	6.4518(0.00)	1.8351(1.00)	4.7205(0.01)
η_1	P_1	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
	P_2	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
	P_3	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
	P_4	—	—	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
η_2	P_1	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
	P_2	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
	P_3	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
	P_4	—	—	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)	0.0000(1.00)
ν_1	P_1	—	—	3.6146(0.97)	19.8156(0.14)	1.3108(1.00)	7.2708(0.73)
	P_2			11.0161(1.00)	14.5225(0.25)	1.9544(1.00)	3.2971(0.89)
	P_3			5.1347(0.96)	20.0174(0.12)	1.7550(1.00)	8.0294(0.66)
	P_4			5.9129(0.97)	19.0113(0.16)	4.9511(0.97)	10.5920(0.46)
ν_2	P_1	—	—	1.0621(1.00)	0.9324(0.92)	3.3490(0.86)	4.7921(0.65)
	P_2			1.7930(0.95)	3.4168(0.63)	4.4666(0.69)	14.4817(0.18)
	P_3			1.0371(0.99)	1.3937(0.92)	3.0896(0.79)	3.7640(0.67)
	P_4			1.9016(0.95)	2.2195(0.91)	1.0183(0.96)	1.7100(0.84)

Table 2 shows the mean squared error (MSE) and coverage percentage for the MCMC estimates based on the 100 samples, in which the coverage percentage is the proportion of the time that the credibility interval contains the true value of interest. The first important fact that is possible to observe from the table is that with high probability the true parameter values are recovered, particularly if the unconstrained methodology is considered. Comparing the unconstrained methodology proposed in this work with the restricted version, there is a significant improvement on the MSE and coverage percentage, specially for the scale, symmetry and kurtosis parameters. Taking λ_2, for example, the coverage percentage is zero or almost zero in all cases and the MSE is more than ten times greater in specific cases.

Taking the hyperparameters specification P_1 as a baseline, a sensitivity analysis study is built. The specification P_2 consists in reducing the values of $\mathbf{B}_0$, and almost no impact on the results of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ is observed, however, looking to the unconstrained model, a significant decrease in the coverage percentage for the scale and symmetry parameters is noticed. The specification P_3 follows [31], the results are similar compared with the P_1 ones, but there is a gain on the MSE for λ_2 in the heavy tailed distributions and unconstrained model. Lastly, a degradation on the MSE for $\boldsymbol{\nu}$ is noted when the changes made in P_4 for d, α and γ are assumed.

4. EMPIRICAL ANALYSIS

In order to explore the interval memory hypothesis and the partial matching hypothesis, [9] designed an experiment in which a pure fundamental tone with electronically generated overtones added was played to a trained musician. The overtones were determined by a stretching ratio, corresponding to the harmonic pattern usually heard in traditional definite pitched instruments. The musician was asked to tune an adjustable tone to the octave above the fundamental tone and 150 trials were recorded as the ratio of the adjusted tone to the fundamental.

This dataset has been analysed in many articles which explored the mixture of linear regression framework [13, 38, 24]. More recently, [41] fitted a robust mixture regression model using the t-distribution and [42], a robust mixture regression based on the SMSN class of distributions. Conducive to make comparisons with the results in [42] possible, the methods proposed in this paper are applied to the tone perception data.

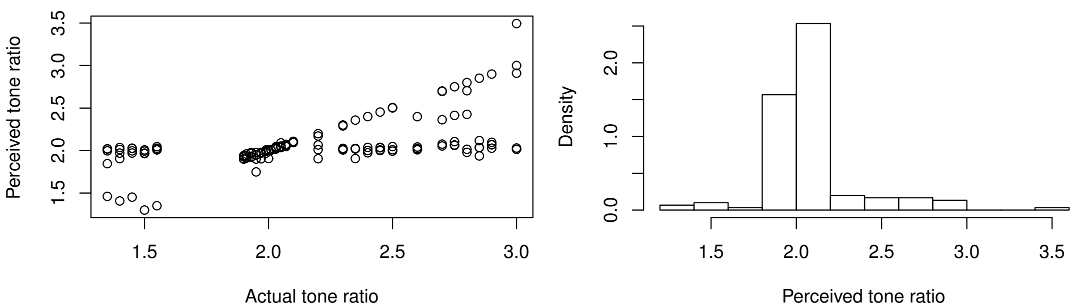

Figure 1: Tone perception data scatterplot and histogram.

Considering the estimation process for the SN-FMRM, ST-FMRM and SSL-FMRM, the hyperparameters specification P_3 presented in Table 1 was chosen. From the MCMC scheme described in Section 2.2, 20000 iterations were drawn. The first 10000 draws were discarded as a burn-in period. In order to reduce the autocorrelation between successive values of the simulated chain, only every 10th values of the chain were stored and from the resulting 1000 we calculated the posterior estimates. It is worth mentioning that, because of the two well defined components, the label switching problem was not identified.

Table 3: Estimation results for fitting the SMSN-FMRM under analysis to the tone data. First row: maximum a posteriori. Second row: 95% high posterior density credibility interval. Third row: convergence test Z-scores.

Parameters	N-FMRM	T-FMRM	SL-FMRM	SN-FMRM	ST-FMRM	SSL-FMRM
$\beta_{0,1}$	1.9107 (1.8586,1.9569) −1.2777	1.9325 (1.8832,1.9771) 0.0250	1.9167 (1.8703,1.9689) −0.0878	1.9044 (1.8532,1.9664) −0.7281	1.9291 (1.8757,1.9846) 0.2527	1.9147 (1.8679,1.9703) −0.2157
$\beta_{1,1}$	0.0457 (0.0243,0.0688) 1.0459	0.0387 (0.0175,0.0595) −0.2666	0.0425 (0.0196,0.0649) −0.2561	0.0447 (0.0205,0.0672) 0.4088	0.0365 (0.0151,0.0618) −0.2404	0.0431 (0.0202,0.0641) 1.1750
$\beta_{0,2}$	−0.0188 (−0.2054,0.2059) −0.7409	0.0153 (−0.0186,0.0704) 1.2719	0.0477 (−0.0317,0.1359) −0.6623	0.0208 (−0.2457,0.2495) −0.1860	0.0136 (−0.0358,0.0849) −0.9211	0.0194 (−0.1075,0.1276) −1.1195
$\beta_{1,2}$	0.9893 (0.9070,1.0802) 0.3946	0.9928 (0.9669,1.0079) −1.5883	0.9745 (0.9304,1.0061) 1.0861	0.9796 (0.8899,1.0971) 0.3949	0.9869 (0.9615,1.0141) 0.0831	0.9729 (0.9228,1.0212) 1.7043
σ_1^2	0.0027 (0.0019,0.0036) −0.4449	0.0020 (0.0012,0.0029) 1.7121	0.0019 (0.0014,0.0029) −1.5685	0.0028 (0.0019,0.0042) 0.6334	0.0021 (0.0013,0.0035) 0.4865	0.0022 (0.0015,0.0034) 1.8521
σ_2^2	0.0173 (0.0105,0.02676) 0.1553	0.0005 (0.0002,0.0010) 1.5999	0.0011 (0.0004,0.0026) −0.9927	0.0269 (0.0127,0.0621) 1.1119	0.0009 (0.0003,0.0024) 0.2782	0.0032 (0.0008,0.0141) −0.2783
λ_1	—	—	—	0.0800 (−0.7634,0.7341) −0.3516	−0.0972 (−0.8113,0.5411) −1.6838	0.0186 (−0.7843,0.5725) 0.1532
λ_2	—	—	—	1.0045 (−1.7427,2.7095) −0.7809	−0.3676 (−1.3333,0.0821) −1.3453	−1.2264 (−2.6623,0.3076) 0.4094
η_1	0.6908 (0.6030,0.7733) −0.2578	0.5606 (0.4700,0.6516) 1.6709	0.5805 (0.4820,0.6876) 1.7261	0.7045 (0.6103,0.7901) 0.3072	0.5691 (0.4538,0.6564) 0.4209	0.6296 (0.5223,0.7383) −0.6418
η_2	0.3091 (0.2266,0.3969) 0.2578	0.4393 (0.3483,0.5299) −1.6709	0.4194 (0.3123,0.5179) −1.7261	0.2954 (0.2098,0.3896) −0.3072	0.4308 (0.3435,0.5461) −0.4209	0.3703 (0.2616,0.4776) 0.6418
ν_1	—	3.0280 (2.0015,24.7743) 0.7383	5.8212 (2.1481,11.7897) −1.0693	—	5.5252 (2.0678,21.7135) 1.5870	6.2337 (3.1571,11.5048) 1.4383
ν_2	—	2.1162 (2.0001,2.6451) 0.8492	1.4630 (1.4000,1.7509) 0.9953	—	2.1281 (2.0000,2.6977) −1.8332	1.5494 (1.4000,3.0780) −0.3276
WAIC_1 WAIC_2	−263.9868 −288.2918	**−349.6941** **−372.0548**	−301.1313 −329.3142	−253.9442 −290.7716	−329.4679 −361.6124	−283.6500 −330.5183

Table 3 contains the maximum a posteriori estimation of the parameters of the models under analysis: SN-FMRM, ST-FMRM and SSL-FMRM in addition to their corresponding 95% high posterior density credibility interval and the Z-scores for the convergence test introduced by [20]. Additionally, in order to compare the fit of the different models, two versions proposed by [19] of the Watanabe–Akaike Information Criterion [40, WAIC] were computed, indicating that the T-FMRM has the best fitting, conclusion that goes in opposition to the ST-FMRM model observed by [42]. More details about these criteria are available in Appendix A.2. Figure 2 illustrates the scatterplots of the dataset with the six fitted models and the equivalent 95% high posterior density credibility intervals.

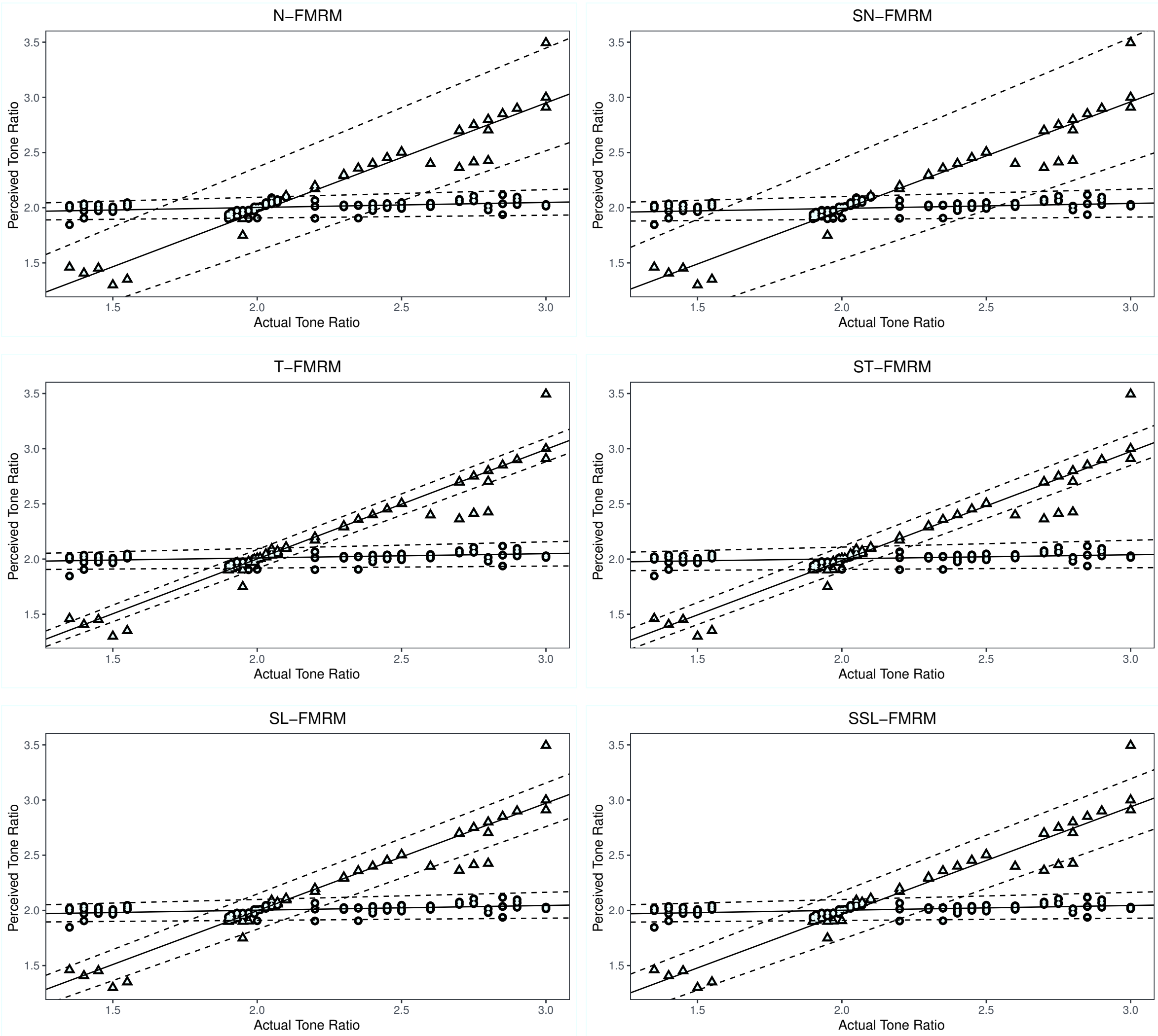

Figure 2: Tone perception data scatterplot and the fitted SMSN-FMRM models.

In comparison with [42], the coefficients $\boldsymbol{\beta}$ estimates are quite similar. However, for the parameters $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$, in line with the results observed on the previous section, the estimates diverge. [42] outcomes point to the presence of asymmetry for at least of one the components when the SN-FMRM, ST-FMRM and SSL-FMRM are considered. Nevertheless, as Figure 3 illustrates, when the flexible version proposed in this paper is applied, it is possible to verify that the introduction of a skewness parameter is not effective considering the dataset under analysis.

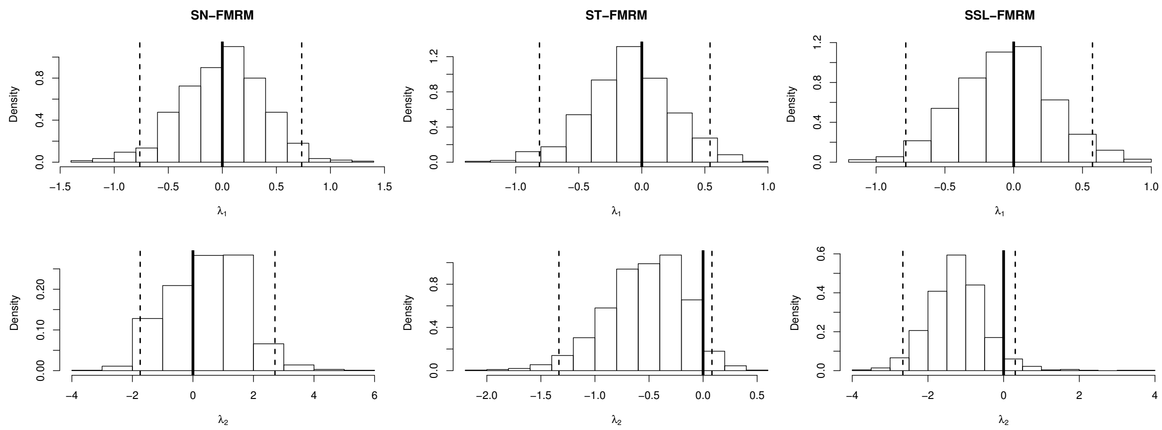

Figure 3: Skewness parameters posterior samples.

5. CONCLUSION

In this work a flexible Bayesian methodology is developed for the mixture regression models based on scale mixtures of skew-normal distributions proposed by [42] with the aim of understanding the possible effects caused by the restrictions commonly imposed in the context of robust mixture regression modeling. The tone perception data and an artificial dataset are analysed in order to verify the advantages that the additional flexibility introduced by the methodology developed in this article has. In fact, this paper presents divergent results in comparison with [42] and the empirical analysis illustrates the possible effects of imposing constraints for this class of models.

Extensions of the contributions made in this article are possible. First, the number of components might be consider as an unknown quantity of interest, estimating it in a full Bayesian framework. Also the proposed methods may be extended to multivariate settings, such as the recent proposals of [18] for mixtures of multivariate Student-t distributions and to models capable to deal with longitudinal data as discussed in [37]. Contemplating extensions able to deal with nonlinear effects of the covariates [7, 8, 5] is also a stimulating topic for further research.

A. APPENDIX

A.1. Mixture regression based on scale mixtures of skew-normal full conditional distributions

Considering the SN-FMRM model and assuming $\mathbf{F}_{n\times(p+1)} = (\mathbf{x}\ \mathbf{w})$, for each $j = 1, ..., G$, construct a matrix $\mathbf{F}_j \in \mathbb{R}^{N_j\times(p+1)}$, $N_j = \sum_{i=1}^{n} S_{ij}$. Similarly, construct an observation matrix $\mathbf{y}_j \in \mathbb{R}^{N_j\times 1}$. Hence, by the Bayes theorem, the full conditionals are:

- $\boldsymbol{\eta}|\mathbf{s} \sim D(e_0 + N_1, ..., e_0 + N_G)$;
- $(\boldsymbol{\beta}_j, \psi_j)|\mathbf{s}, \mathbf{y}, \mathbf{w}, \tau_k^2 \sim N_{p+1}(\mathbf{b}_j, \mathbf{B}_j)$,
  $\mathbf{B}_j = \left(\frac{1}{\tau_j^2}\mathbf{B}_0^{-1} + \frac{1}{\tau_j^2}(\mathbf{F}_j{}^T\mathbf{F}_j)\right)^{-1}$,
  $\mathbf{b}_j = \mathbf{B}_j\left(\frac{1}{\tau_j^2}\mathbf{B}_0^{-1}\mathbf{b}_0 + \frac{1}{\tau_j^2}(\mathbf{F}_j{}^T(\mathbf{y}_k - \mu_k))\right)$;
- $\tau_j^2|\mathbf{s}, \mathbf{y}, \mathbf{w}, C_0, \boldsymbol{\beta}_j, \psi_j \sim IG(c_j, C_j)$,
  $c_j = c_0 + \frac{N_j}{2} + \frac{1}{2}$,
  $C_j = C_0 + \frac{(\mathbf{y}_j - \mathbf{F}_j\boldsymbol{\beta}_j^* - \mu_j)^T(\mathbf{y}_j - \mathbf{F}_j\boldsymbol{\beta}_j^* - \mu_j) + (\boldsymbol{\beta}_j^* - \mathbf{b}_0)^T\mathbf{B}_0^{-1}(\boldsymbol{\beta}_j^* - \mathbf{b}_0)}{2}$;
- $C_0|\tau_1^2, ..., \tau_G^2 \sim G(h, H)$,
  $h = h_0 + Gc_0$,
  $H = H_0 + \sum_{j=1}^{G}\frac{1}{\tau_j^2}$;

where $\boldsymbol{\beta}_j^* = (\boldsymbol{\beta}_j\ \psi_j)^T$. Considering now the latent variable $\mathbf{W}$:

- $W_i|S_{ij} = 1, y_i, \boldsymbol{\beta}_j, \psi_j, \tau_j^2 \sim TN_{[0,+\infty)}(a, A)$,
  $a = \frac{(y_i - \mathbf{x}_i\boldsymbol{\beta}_j - \mu_j)\psi_j}{\tau_j^2 + \psi_j^2}$,
  $A = \frac{\tau_j^2}{\tau_j^2 + \psi_j^2}$.

For the ST-FMRM and the SSL-FMRM models the full conditionals are almost the same, the difference is that $\mathbf{F}$ is replaced by $\mathbf{F}^w_{n\times(p+1)} = (\sqrt{\mathbf{u}}\mathbf{x}\ \sqrt{\mathbf{u}}\mathbf{w})$ and $\mathbf{y}$, by $\mathbf{y}^w = \sqrt{\mathbf{u}}\mathbf{y}$, where $\sqrt{\mathbf{u}}$ is the square root element by element. Considering now the latent variable $\mathbf{W}$:

- $W_i|S_{ij} = 1, y_i, u_i, \boldsymbol{\beta}_j, \psi_j, \tau_j^2 \sim TN_{[0,+\infty)}(a, A/u_i)$.

Lastly, for the latent variable $\mathbf{U}$ and the parameters ν:

- Skew-t:
  $U_i|S_{ij} = 1, y_i, w_i, \nu_j, \boldsymbol{\beta}_j, \psi_j, \tau_j^2 \sim G\left(\frac{\nu_j}{2} + 1, \frac{\nu_j}{2} + \frac{(y_i - \mu_j - \mathbf{x}_i\boldsymbol{\beta}_j - \psi_j w_i)^2}{2\tau_j^2} + \frac{w_i^2}{2}\right)$;
- Skew-slash:
  $U_i|S_{ij} = 1, y_i, w_i, \nu_j, \boldsymbol{\beta}_j, \psi_j, \tau_j^2 \sim G_{(0,1)}\left(\nu_j + 1, \frac{(y_i - \mu_j - \mathbf{x}_i\boldsymbol{\beta}_j - \psi_j w_i)^2}{2\tau_j^2} + \frac{w_i^2}{2}\right)$,
  $\nu_j|\mathbf{s}, \mathbf{u} \sim G_{(2,40)}(\alpha + N_j, \gamma - \sum_{i:S_{ij}=1} u_i)$.

For the degrees of freedom in skew-t is not possible to find a closed form to the full conditionals, so a Metropolis–Hastings step is required. To sample ν_j, $j = 1, ..., G$ a normal log random walk proposal was used

$$\log(\nu_j^{new} - 2) \sim N(\log(\nu_j - 2), c_{\nu_j}) \tag{A.1}$$

with adaptive width parameter c_{ν_j} [32]. The proposal was shifted away from 0, as it is advisable to avoid values for ν_j that are close to 0, see [15].

A.2. Watanabe–Akaike information criterion

Define the predictive accuracy of the fitted model to data as

$$p(\mathbf{y}) = \sum_{i=1}^{n} \log \int f(y_i|\boldsymbol{\theta})p(\boldsymbol{\theta}|\mathbf{y})d\boldsymbol{\theta}.$$

To compute this predictive density, it is possible to evaluate the expectation using draws from the usual posterior simulations:

$$\overline{p(\mathbf{y})} = \sum_{i=1}^{n} \log \left(\frac{1}{T} \sum_{t=1}^{T} f(y_i|\boldsymbol{\theta}^{(t)}) \right).$$

Introduced by [40], the Watanabe–Akaike information criterion (WAIC) consists on the posterior predictive density in addition to a correction for effective number of parameters to adjust for overfitting. [19] describes two adjustments. The first one is a difference:

$$WAIC_1^* = 2 \sum_{i=1}^{n} \left(\log \left(E_{(\boldsymbol{\theta}|\mathbf{y})} f(y_i|\boldsymbol{\theta}) \right) - E_{(\boldsymbol{\theta}|\mathbf{y})} \left(\log(f(y_i|\boldsymbol{\theta})) \right) \right),$$

which can be computed from simulations by replacing the expectations by averages over the posterior draws, it means,

$$\overline{WAIC_1^*} = 2 \sum_{i=1}^{n} \left(\log \left(\frac{1}{T} \sum_{t=1}^{T} f(y_i|\boldsymbol{\theta}^{(t)}) \right) - \frac{1}{T} \sum_{t=1}^{T} \log f(y_i|\boldsymbol{\theta}^{(t)}) \right).$$

The second is based on the variance of individual terms in the log predictive density summed over the n data observations:

$$WAIC_2^* = \sum_{i=1}^{n} \mathrm{var}_{(\boldsymbol{\theta}|\mathbf{y})} \left(\log f(y_i|\boldsymbol{\theta}) \right).$$

In practice, the posterior variance of the log predictive density for each data point y_i, that is, $V_{t=1}^{T} \log f(y_i|\theta^{(t)})$, where $V_{t=1}^{T}$ is the sample variance, $V_{t=1}^{T} a_{(t)} = \frac{1}{T-1} \sum_{t=1}^{T} (a_{(t)} - \bar{a})^2$. Summing over all the data observations, the effective number of parameters is:

$$\overline{WAIC_2^*} = \sum_{i=1}^{n} V_{t=1}^{T} \left(\log f(y_i|\theta^{(t)}) \right).$$

Finally, either $WAIC_1^*$ or $WAIC_2^*$ are applied as a bias correction:

$$WAIC_q = -2(p(\mathbf{y}) - WAIC_q^*). \tag{A.2}$$

REFERENCES

[1] Azzalini, A. (1985). A class of distributions which includes the normal ones, *Scandinavian Journal of Statistics*, **12**(2), 171–178.

[2] Azzalini, A. (1986). Further results on a class of distributions which includes the normal ones, *Statistica*, **46**(2), 199–208.

[3] Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution, *Journal of the Royal Statistical Society, Series B*, **65**(2), 367–389.

[4] Bai, X.; Yao, W. and Boyer, J.E. (2012). Robust fitting of mixture regression models, *Computational Statistics and Data Analysis*, **56**(7), 2347–2359.

[5] Baldacchino, T.; Worden, K. and Rowson, J. (2017). Robust nonlinear system identification: Bayesian mixture of experts using the t-distribution, *Mechanical Systems and Signal Processing*, **85**, 977–992.

[6] Branco, M.D. and Dey, D.K. (2001). A general class of multivariate skew-elliptical distributions, *Journal of Multivariate Analysis*, **79**(1), 99–113.

[7] Carvalho, A.X. and Tanner, M.A. (2006). Modeling nonlinearities with mixtures-of-experts of time series models, *International Journal of Mathematics and Mathematical Sciences*, **2006**, 1–22.

[8] Chamroukhi, F. (2016). Robust mixture of experts modeling using the t-distribution, *Neural Networks*, **79**, 20–36.

[9] Cohen, E.A. (1984). Some effects of inharmonic partials on interval perception, *Music Perception*, **1**(3), 323–349.

[10] Cosslett, S.R. and Lee, L.F. (1985). Serial correlation in latent discrete variable models, *Journal of Econometrics*, **27**(1), 79–97.

[11] DeSarbo, W.S. and Cron, W.L. (1988). A maximum likelihood methodology for clusterwise linear regression, *Journal of Classification*, **5**, 249–282.

[12] DeSarbo, W.S.; Wedel, M.; Vriens, M. and Ramaswamy, V. (1992). Latent class metric conjoint analysis, *Marketing Letters*, **3**(3), 273–288.

[13] DeVeaux, R.D. (1989). Mixtures of linear regressions, *Computational Statistics and Data Analysis*, **8**(3), 227–245.

[14] Diebolt, J. and Robert, C.P. (1994). Estimation of finite mixture distributions through Bayesian sampling, *Journal of the Royal Statistical Society, Series B*, **56**(2), 363–375.

[15] Fernandez, C. and Steel, M.F.J. (1999). Multivariate Student-t regression models: pitfalls and inference, *Biometrika*, **86**(1), 153–167.

[16] Fruhwirth-Schnatter, S. (2006). *Finite Mixture and Markov Switching Models*, Springer, New York.

[17] Fruhwirth-Schnatter, S. and Pyne, S. (2010). Bayesian inference for finite mixtures of univariate and multivariate skew-normal and skew-t distributions, *Biostatistics*, **11**(2), 317–336.

[18] Galimberti, G. and Soffritti, G. (2014). A multivariate linear regression analysis using finite mixtures of t distributions, *Computational Statistics and Data Analysis*, **71**, 138–150.

[19] Gelman, A.; Hwang, J. and Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models, *Statistics and Computing*, **24**, 997–1016.

[20] Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments, *Bayesian Statistics*, **4**, 169–193.

[21] HAMILTON, J.D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle, *Econometrica*, **57**(2), 357–384.

[22] HANSON, T. and JOHNSON, W.O. (2002). Modeling regression error with a mixture of Polya trees, *Journal of the American Statistical Association*, **97**(460), 1020–1033.

[23] HENZE, N. (1986). A probabilistic representation of the skew-normal distribution, *Scandinavian Journal of Statistics*, **13**(4), 271–275.

[24] HUNTER, D.R. and YOUNG, D.S. (2012). Semiparametric mixtures of regressions, *Journal of Nonparametric Statistics*, **24**(1), 19–38.

[25] JENNISON, C. (1997). Discussion of the paper by Richardson and Green, *Journal of the Royal Statistical Society, Series B*, **59**(4), 778–779.

[26] JUÁREZ, M.A. and STEEL, M.F.J. (2010). Model-based clustering of non-Gaussian panel data based on skew-t distributions, *Journal of Business & Economic Statistics*, **28**(1), 52–66.

[27] KUO, L. and MALLICK, B.K. (1997). Bayesian semiparametric inference for the accelerated failure time model, *Canadian Journal of Statistics*, **25**(4), 457–472.

[28] LIU, M.; HANCOCK, G.R. and HARRING, J.R. (2011). Using finite mixture modeling to deal with systematic measurement error: a case study, *Journal of Modern Applied Statistical Methods*, **10**(1), 249–261.

[29] LIU, M. and LIN, T.I. (2014). A skew-normal mixture regression model, *Educational and Psychological Measurement*, **74**(1), 139–162.

[30] REDNER, R.A. and WALKER, H. (1984). Mixture densities, maximum likelihood and the EM algorithm, *SIAM Review*, **26**(2), 195–239.

[31] RICHARDSON, S. and GREEN, P.J. (1997). On Bayesian analysis of mixtures with an unknown number of components, *Journal of the Royal Statistical Society, Series B*, **59**(4), 731–792.

[32] SHABY, B.A. and WELLS, M.T. (2010). *Exploring an adaptive Metropolis algorithm*, Technical Report, Duke University, Department of Statistical Science.

[33] SONG, W.; YAO, W. and XING, Y. (2014). Robust mixture regression model fitting by Laplace distribution, *Computational Statistics and Data Analysis*, **71**, 128–137.

[34] STEPHENS, M. (2000). Dealing with label switching in mixture models, *Journal of the Royal Statistical Society, Series B*, **62**(4), 795–809.

[35] TANNER, M.A. and WONG, W.H. (1987). The calculation of posterior distributions by data augmentation, *Journal of the American Statistical Association*, **82**(398), 528–540.

[36] TURNER, T.R. (2000). Estimating the propagation rate of a viral infection of potato plants via mixtures of regressions, *Applied Statistics*, **49**(3), 371–384.

[37] VERBEKE, G. and LESAFFRE, E. (1996). A linear mixed-effects model with heterogeneity in the random-effects population, *Journal of the American Statistical Association*, **91**(433), 217–221.

[38] VIELE, K. and TONG, B. (2002). Modeling with mixtures of linear regressions, *Statistics and Computing*, **12**, 315–330.

[39] WANG, J. and GENTON, M.G. (2006). The multivariate skew-slash distribution, *Journal of Statistical Planning and Inference*, **136**(1), 209–220.

[40] WATANABE, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory, *Journal of Machine Learning Research*, **11**, 3571–3594.

[41] YAO, W.; WEI, Y. and YU, C. (2014). Robust mixture regression using the t-distribution, *Computational Statistics and Data Analysis*, **71**, 116–127.

[42] ZELLER, C.B.; CABRAL, C.R.B. and LACHOS, V.H. (2016). Robust mixture regression modeling based on scale mixtures of skew-normal distributions, *TEST*, **25**, 375–396.

REVSTAT – Statistical Journal
Volume 20, Number 1, January 2022, 117–134

ECONOMIC AND ECONOMIC-STATISTICAL DESIGNS OF MULTIVARIATE COEFFICIENT OF VARIATION CHART

Authors: Wei Chun Ng
– School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia

Michael B.C. Khoo *
– School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Minden, Penang, Malaysia
mkbc@usm.my

Zhi Lin Chong
– Department of Physical and Mathematical Science, Faculty of Science, Universiti Tunku Abdul Rahman, 31900 Kampar, Perak Darul Ridzuan, Malaysia

Ming Ha Lee
– Faculty of Engineering, Computing and Science, Swinburne University of Technology Sarawak Campus, 93350 Kuching, Sarawak, Malaysia

Received: June 2019 Revised: December 2019 Accepted: December 2019

Abstract:

- From the economic perspective, cost minimization is an important part of Statistical Process Control (SPC). The conventional approach in SPC focuses on monitoring the process mean and variance for possible shifts. In some processes, such as clinical and financial investments, the process mean and variance are not independent of one another. Thus, a separate monitoring of the mean and variance using two different control charts is not meaningful. Therefore, the coefficient of variation chart that measures the ratio of the process variance to the mean needs to be employed. In multivariate SPC, the quality characteristics that jointly control the process quality are correlated. Thus, the multivariate coefficient of variation (MCV) chart is used in process monitoring to monitor the process MCV. This work studies the economic and economic-statistical designs of the MCV chart. Optimal parameters that minimize the cost function of the MCV chart are computed. Furthermore, it is shown that adding statistical constraints to the economic design of the MCV chart improves the chart's statistical performance with only a minimal increase in cost.

Keywords:

- *multivariate coefficient of variation (MCV); economic design; economic-statistical design; cost model.*

AMS Subject Classification:

- 62P20, 62P30, 91B02.

*Corresponding author.

1. INTRODUCTION

The coefficient of variation (CV) chart is commonly used in SPC for processes which require the reproducibility of measuring tools or methods [3, 20]. Operators usually demand a lower CV profile for better equipment and/or method precision while maintaining the accuracy of the process with an in-control state [8, 17]. Examples of the use of CV are laboratory assay techniques in medicine and biology [19, 36], monitoring the associated stand-alone risk in actuarial finance [24], factory processes in mechanical industries [4], to name a few.

Kang *et al.* [9] proposed the first Shewhart-type univariate CV chart. Since then, the univariate CV charts continue to receive attention among researchers (see [4] and [28], to name a few) but not the multivariate CV (MCV) chart. Yeong *et al.* [32] was the first to propose a control chart for the MCV. More recent studies on MCV charts include studies by Giner-Bosch *et al.* [6] on the EWMA MCV chart and Nguyen *et al.* [16] on one-sided synthetic MCV charts. Some crucial applications of MCV in laboratories and industries are in the correlation of phenotypic variation [25], affymetrix gene expression [7], comparison of serum protein electrophoresis techniques [35], multivariate gage repeatability and reproducibility studies [18, 27], and several others.

The advancement in hardware technologies enabled more automation techniques to be easily applied in various aspects of living. Newly developed equipment and methods can produce large pool of useful data and results with high efficiency. The generalization of CV to the multivariate setting is required to accommodate the part-to-part variability measurements and the correlations of higher dimensional variables. However, the definition of MCV is not as straight forward as that of the univariate CV, i.e. lacking in the generality. Currently, the available definitions of MCV were those by Reyment [21], Van Valen [29], Voinov and Nikulin [30], and Albert and Zhang [2]. Similar to existing MCV type control charts (see for example, Yeong *et al.* [32], Abbasi and Adegoke [1], Khaw *et al.* [11] and Khatun *et al.* [10]), this work adopts the Voinov and Nikulin's [30] definition of MCV.

A pure statistical design of a control chart may not be cost effective in industrial practices. An optimal economic design of a control chart will enhance the competency of the chart from the cost perspective [26]. The idea of an economic model was first presented by Duncan [5], and later improved by Lorenzen and Vance [13]. Saniga [23] expanded the model by incorporating statistical constraints into the cost function, resulting in an economic-statistical model. The unified cost model by Lorenzen and Vance [13] is widely accepted and used in many types of control charts. Some published works which are closely related to this study include Linderman and Love [12] and Molnau *et al.* [14] on economic and economic-statistical designs of multivariate EWMA control chart.

Despite being over three decades old, the Lorenzen and Vance's [13] model is one of the most inclusive cost models in the literature, where it considers all possible sources of cost assumptions, phases of a process and evaluations of expenses. As the Lorenzen and Vance's [13] model is easy to be implemented, it continues to be adopted by researchers until now. Some of the recent works that adopted the Lorenzen and Vance's [13] model are Safe *et al.* [22] and Wan and Zhu [31] who used the model on variable sampling interval type control charts; and Ng *et al.* [15] who employed the model on auxiliary information based $\bar{X}$, synthetic and EWMA charts. Note that the numerical example presented in Lorenzen and Vance [13] and adopted by the above-mentioned researchers, to name a few, is based on a real casting operation process from the General Motors Company.

This study proposes the economic and economic-statistical designs of MCV chart as they are currently not available in the literature. In each of the designs, optimal parameters will be computed to minimize the cost. A comparison between purely economic design and economic-statistical design will also be presented.

This paper is organized in the following order: The properties of MCV and the MCV chart will be explained in Section 2. Following that is a brief review on Lorenzen and Vance [13] cost model in Section 3. Subsequently, a set of numerical examples along with comparisons of different parameter settings and designs are given in Section 4. A sum up of the paper with some general remarks and findings are given in Section 5.

2. PROPERTIES OF MCV AND MCV CHART

Section 2.1 discusses the cumulative distribution function (cdf) and inverse cdf of the sample MCV derived by Yeong *et al.* [32] while the MCV chart is discussed in Section 2.2.

2.1. Distribution of the sample MCV

Suppose that a random vector, $\mathbf{X_i}$, in a sample of size n with mean vector, $\boldsymbol{\mu}$ and covariance matrix, $\boldsymbol{\Sigma}$ follows a p-variate normal distribution, i.e. $\mathbf{X}_i \sim N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\mathbf{X}_i^\mathsf{T} = (X_{i1}, X_{i2}, ..., X_{ip})$, for $1 \leq i \leq n$. A general definition of the population MCV by Voinov and Nikulin [30] is

$$\gamma = \left(\boldsymbol{\mu}^\mathsf{T} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}\right)^{-\frac{1}{2}}. \tag{2.1}$$

Yeong *et al.* [32] derived an estimator of the process MCV, $\hat{\gamma}$ based on Equation (2.1), where $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are estimated using the sample mean vector, $\bar{\mathbf{X}}$ and the sample covariance matrix, $\mathbf{S}$, respectively. Here,

$$\bar{\mathbf{X}}^\mathsf{T} = \left(\frac{1}{n}\sum_{i=1}^{n} X_{i1}, \frac{1}{n}\sum_{i=1}^{n} X_{i2}, ..., \frac{1}{n}\sum_{i=1}^{n} X_{ip}\right), \tag{2.2}$$

and

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n} (\mathbf{X}_i - \bar{\mathbf{X}})(\mathbf{X}_i - \bar{\mathbf{X}})^\mathsf{T}. \tag{2.3}$$

Then, $\hat{\gamma}$ takes the form

$$\hat{\gamma} = (\bar{\mathbf{X}}^\mathsf{T} \mathbf{S}^{-1} \bar{\mathbf{X}})^{-\frac{1}{2}}. \tag{2.4}$$

The cdf of $\hat{\gamma}$ was derived by Yeong *et al.* [32] to be

$$F_{\hat{\gamma}}(x|n,p,\delta) = 1 - F_F\left(\frac{n(n-p)}{(n-1)px^2}|p,n-p,\delta\right), \tag{2.5}$$

where $F_F(\cdot|p,n-p,\delta)$ is the non-central F distribution with p and $n-p$ degrees of freedom and non-centrality parameter $\delta = n\boldsymbol{\mu}^\mathsf{T}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}$ (which can be written as $\delta = \frac{n}{\gamma^2}$). Yeong *et al.* [32] also derived the inverse cdf of $\hat{\gamma}$ (or the α quantile of $F_{\hat{\gamma}}$) as follows:

$$F_{\hat{\gamma}}^{-1}(\alpha|n,p,\delta) = \sqrt{\frac{n(n-p)}{(n-1)p}\left[\frac{1}{F_F^{-1}(1-\alpha|p,n-p,\delta)}\right]}. \tag{2.6}$$

Note that $F_F^{-1}(\cdot|p,n-p,\delta)$ is the inverse cdf of the non-central F distribution with p and $n-p$ degrees of freedom and non-centrality parameter δ.

2.2. MCV chart

The MCV chart is a Shewhart type chart where the statistic plotted on the chart is the sample MCV, $\hat{\gamma}$. To justify the use of the MCV chart, a check for the constant MCV assumption needs to be conducted. This check is conducted by plotting the rational group MCV, $\hat{\gamma}_t^2$ versus $\bar{\mathbf{X}}_t^\mathsf{T}\bar{\mathbf{X}}_t$, followed by a formal test of the regression slope [32].

Yeong *et al.* [32] suggested estimating the in-control sample MCV, $\hat{\gamma}_0$ using the root mean square method as this method has high statistical efficiency and the estimate can be easily computed. Consequently, $\hat{\gamma}_0$ is computed as

$$\hat{\gamma}_0 = \sqrt{\frac{1}{m}\sum_{t=1}^{m}\hat{\gamma}_t^2}\ , \tag{2.7}$$

where m is the number of Phase-I sample MCVs. As the distribution of $\hat{\gamma}$ is not symmetric, the use of two-sided limits will result in an average run length (ARL) biased chart. Therefore, Yeong *et al.* [32] suggested adopting two separate one-sided (an upward and a downward) charts to overcome this drawback. Using two separate one-sided charts allow the upper and lower limits of the respective charts to be determined independently based on the desired in-control ARL value.

For the downward MCV chart in detecting decreasing shifts in the process MCV, its lower control limit (LCL) is computed as

$$\mathrm{LCL} = F_{\hat{\gamma}}^{-1}(\alpha|n,p,\delta_0), \tag{2.8}$$

where α is the Type-I error probability and $\delta_0 = \frac{n}{\gamma_0^2}$ with γ_0 representing the in-control process MCV. The statistical performance of MCV chart can be measured using the ARL criterion. The corresponding value of the in-control average run length (ARL_0) computed using the LCL in Equation (2.8) is

$$\mathrm{ARL}_0 = \frac{1}{\alpha}. \tag{2.9}$$

In like manner, for the upward MCV chart in detecting increasing shifts in the process MCV, its upper control limit (UCL) is obtained as

$$\text{UCL} = F_{\hat{\gamma}}^{-1}(1-\alpha|n,p,\delta_0) \tag{2.10}$$

which gives the ARL_0 value in Equation (2.9). The process MCV is considered as out-of-control when $\hat{\gamma} < \text{LCL}$ (for the downward chart) or $\hat{\gamma} > \text{UCL}$ (for the upward chart).

The out-of-control process MCV is represented by $\gamma_1 = \tau\gamma_0$. Here, τ is the shift size in the process MCV, where $\tau < 1$ ($\gamma_1 < \gamma_0$) indicates process improvement, while $\tau > 1$ ($\gamma_1 > \gamma_0$) implies process deterioration. The probability of detecting a shift by the downward and upward MCV charts are

$$P = \Pr(\hat{\gamma} < \text{LCL}) = F_{\hat{\gamma}}(\text{LCL}|n,p,\delta_1) \tag{2.11}$$

and

$$P = \Pr(\hat{\gamma} > \text{UCL}) = 1 - F_{\hat{\gamma}}(\text{UCL}|n,p,\delta_1), \tag{2.12}$$

respectively, where $\delta_1 = \dfrac{n}{\gamma_1^2}$. The out-of-control average run length (ARL_1) is computed as

$$\text{ARL}_1 = \frac{1}{P}. \tag{2.13}$$

3. LORENZEN AND VANCE COST MODEL

The unified cost model proposed by Lorenzen and Vance [13] is adopted for the economic and economic-statistical designs of the MCV chart. The functional form of this model only requires the computation of ARL, sample size and control limit of the chart at hand. Thus, Lorenzen and Vance [13] cost model can be used on any type of control chart, regardless of the quality characteristics. Table 1 provides the list of notations for this cost model.

The total cost per hour as defined by this model includes the costs during the in-control and out-of-control states, cost of false alarms, cost of repair and cost of sampling. In Lorenzen and Vance [13] cost model, the assignable cause is assumed to occur randomly once in every λ hours. Another assumption is that the shift in the process MCV is due to only a single assignable cause. Lorenzen and Vance [13] cost function is defined as

$$C = \frac{\dfrac{C_0}{\lambda} + C_1 B + \dfrac{b+cn}{h}\left(\dfrac{1}{\lambda} + B\right) + \dfrac{sY}{\text{ARL}_0} + W}{\dfrac{1}{\lambda} + \dfrac{(1-\varphi_1)sT_0}{\text{ARL}_0} + EH}\,, \tag{3.1}$$

where

$$\begin{aligned} B &= (\text{ARL}_1 - 0.5)h + F, \\ F &= ne + \varphi_1 T_1 + \varphi_2 T_2, \\ EH &= (\text{ARL}_1 - 0.5)h + G, \\ G &= ne + T_1 + T_2, \end{aligned}$$

and

$$s = \frac{1}{\lambda h} - \frac{1}{2}.$$

Table 1: List of notations for Lorenzen and Vance (1986) cost model.

b	Fixed cost per sample
c	Variable cost per unit sampled
C	Cost per hour
C_0	Quality cost per hour while in-control
C_1	Quality cost per hour while out-of-control
e	Time to sample and interpret one unit
h	Sampling interval
n	Sample size
s	Expected number of samples taken while in-control
T_0	Expected search time during false alarm
T_1	Expected time to find the assignable cause
T_2	Expected time to repair the process
W	Cost to locate and remove the assignable cause
Y	Cost of false alarms
φ_1	$= 1$ if process continues during search $= 0$ if process stops during search
φ_2	$= 1$ if process continues during repair $= 0$ if process stops during repair
λ	Rate of occurrence of assignable cause

The objective of the economic design of the MCV chart is to obtain the optimal parameters n, h and α in minimizing the cost function, C in Equation (3.1), for specified values of p, τ and γ_0. Note that the parameters p, τ and γ_0 are not included in the optimization procedure because they are intrinsic properties of the process.

With the same objective, the economic-statistical design adds additional constraints on ARL_0 and ARL_1 while minimizing the cost function, C in Equation (3.1). Here, ARL_0 must be greater than a lower bound value while ARL_1 must be less than an upper bound value. The aim of these constraints is to ensure that the MCV chart gives acceptably high ARL_0 value when the process is in-control and low ARL_1 value when the process is out-of-control. In this research, the constraints $\mathrm{ARL}_0 \geq 250$ and $\mathrm{ARL}_1 \leq 20$, i.e. similar to those used by Yeong *et al.* [34] are adopted.

The optimal sampling interval, h can be computed as follows [33]:

$$h = \frac{-r_2 + \sqrt{r_2^2 - 4r_1r_3}}{2r_1}, \tag{3.2}$$

where

$$\begin{aligned} r_1 \;=\; & \frac{\mathrm{ARL}_1 - 0.5}{2\lambda \mathrm{ARL}_0} \Big\{ \lambda\big(Y + C_1T_0(-1+\varphi_1)\big) \\ & - 2\,\mathrm{ARL}_0 \Big[C_0 + \lambda\big((\mathrm{ARL}_1 - 0.5)b + (\mathrm{ARL}_1 - 0.5)cn + W\big) + C_1(-1 + F\lambda - G\lambda) \Big] \Big\}, \end{aligned}$$

$$r_2 \;=\; -\,\frac{2(\mathrm{ARL}_1 - 0.5)\Big[Y + C_1T_0(-1+\varphi_1) + \mathrm{ARL}_0(b+cn)(1+F\lambda) \Big]}{\lambda \mathrm{ARL}_0},$$

and

$$\begin{aligned} r_3 = & -\frac{1}{2\lambda^2 \mathrm{ARL}_0}\Big\{2Y + 2C_0T_0(-1+\varphi_1) - bT_0\lambda - 2(\mathrm{ARL}_1 - 0.5)bT_0\lambda - 2C_1FT_0\lambda \\ & - cnT_0\lambda - 2(\mathrm{ARL}_1 - 0.5)cnT_0\lambda - 2T_0W\lambda + 2GY\lambda + bT_0\varphi_1\lambda \\ & + 2(\mathrm{ARL}_1 - 0.5)bT_0\varphi_1\lambda + 2C_1FT_0\varphi_1\lambda + cnT_0\varphi_1\lambda \\ & + 2(\mathrm{ARL}_1 - 0.5)cnT_0\varphi_1\lambda + 2T_0W\varphi_1\lambda - bFT_0\lambda^2 - cFnT_0\lambda^2 \\ & + bFT_0\varphi_1\lambda^2 + cFnT_0\varphi_1\lambda^2 + 2\mathrm{ARL}_0(b+cn)(1+F\lambda)(1+G\lambda)\Big\}. \end{aligned}$$

From Equations (3.1) and (3.2), it is clear that both ARL_0 and ARL_1 need to be computed first before the computation of C and h can be made. The formulae for computing ARL_0 and ARL_1 are dependent on n, α, p, τ and γ_0. As the exact values of p, τ, γ_0 and the desired values of the thirteen input parameters in Table 2, i.e. λ, C_0, C_1, Y, W, b, c, e, T_0, T_1, T_2, φ_1 and φ_2 are specified, the parameters that control the cost minimization iteration in this case are n and α. The desired values of these thirteen input parameters are adopted from Lorenzen and Vance [13], where they are taken as the control case (Case 1) in Table 2.

Table 2: Input parameters for the cost function, C and the variations of each input parameters, labelled with case numbering.

Case	Changes	λ	C_0	C_1	Y	W	b	c	e	T_0	T_1	T_2	φ_1	φ_2
1	Control	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
2	$\lambda 2$	**0.01**	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
3	$\lambda 3$	**0.04**	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
4	$C_0 2$	0.02	**57.12**	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
5	$C_0 3$	0.02	**228.48**	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
6	$C_1 2$	0.02	114.24	**474.6**	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
7	$C_1 3$	0.02	114.24	**1898.4**	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
8	$Y2$	0.02	114.24	949.2	**488.7**	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
9	$Y3$	0.02	114.24	949.2	**1954.8**	977.4	0	4.22	0.083	0.083	0.083	0.75	1	0
10	$W2$	0.02	114.24	949.2	977.4	**488.7**	0	4.22	0.083	0.083	0.083	0.75	1	0
11	$W3$	0.02	114.24	949.2	977.4	**1954.8**	0	4.22	0.083	0.083	0.083	0.75	1	0
12	$b2$	0.02	114.24	949.2	977.4	977.4	**5**	4.22	0.083	0.083	0.083	0.75	1	0
13	$b3$	0.02	114.24	949.2	977.4	977.4	**10**	4.22	0.083	0.083	0.083	0.75	1	0
14	$c2$	0.02	114.24	949.2	977.4	977.4	0	**2.11**	0.083	0.083	0.083	0.75	1	0
15	$c3$	0.02	114.24	949.2	977.4	977.4	0	**8.44**	0.083	0.083	0.083	0.75	1	0
16	$e2$	0.02	114.24	949.2	977.4	977.4	0	4.22	**0.042**	0.083	0.083	0.75	1	0
17	$e3$	0.02	114.24	949.2	977.4	977.4	0	4.22	**0.166**	0.083	0.083	0.75	1	0
18	$T_0 2$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	**0.042**	0.083	0.75	1	0
19	$T_0 3$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	**0.166**	0.083	0.75	1	0
20	$T_1 2$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	**0.042**	0.75	1	0
21	$T_1 3$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	**0.166**	0.75	1	0
22	$T_2 2$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	**0.375**	1	0
23	$T_2 3$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	**1.5**	1	0
24	$\varphi_1\varphi_2 2$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	**0**	**0**
25	$\varphi_1\varphi_2 3$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	**0**	**1**
26	$\varphi_1\varphi_2 4$	0.02	114.24	949.2	977.4	977.4	0	4.22	0.083	0.083	0.083	0.75	**1**	**1**

The computations of the control values of these thirteen input parameters will be explained in detail in Section 4.

In order to impose changes to each of the thirteen input parameters of the control case (Case 1) in Table 2, each of these input parameters (except b, φ_1 and φ_2) is either increased (i.e. doubled) or decreased (i.e. halved). For example, $\lambda 2$ $(= 0.01)$ (Case 2) is half of its control value $(\lambda = 0.02)$ in Case 1, while $\lambda 3$ $(= 0.04)$ (Case 3) is twice of its control value in Case 1. The notations $\lambda 2$ and $\lambda 3$ are used to represent the second and third variations of the control value of λ, as not every input parameter (such as b, φ_1 and φ_2) is doubled or halved. For instance, the fixed cost per sample, b is set at \$0 for the control case (Case 1), while $b2$ involves a raise to \$5 (Case 12) and $b3$ to \$10 (Case 13).

In this research, the sample sizes, $n \in \{2, 3, ..., 30\}$ are considered. The upper limit of n $(= 30)$ is chosen because from a practical perspective, $n = 30$ is considered as a large sample size. In addition, the Type-I error probabilities $\alpha \in \{0.0010, 0.0011, ..., 0.05\}$ are adopted for the economic design, while $\alpha \in \{0.0010, 0.0011, ..., 0.004\}$ are adopted for the economic-statistical design. Note that the Type-I error rate for the economic-statistical design is kept at a maximum of $\alpha = 0.004$, in order to correspond to the constraint $\text{ARL}_0 \geq 250$ specified earlier. An optimization program is written in the MATLAB software to compute the optimal parameters n, α and h that minimize the cost function, C in Equation (3.1), based on the specified values of p, τ, γ_0 and thirteen input parameters in Table 2.

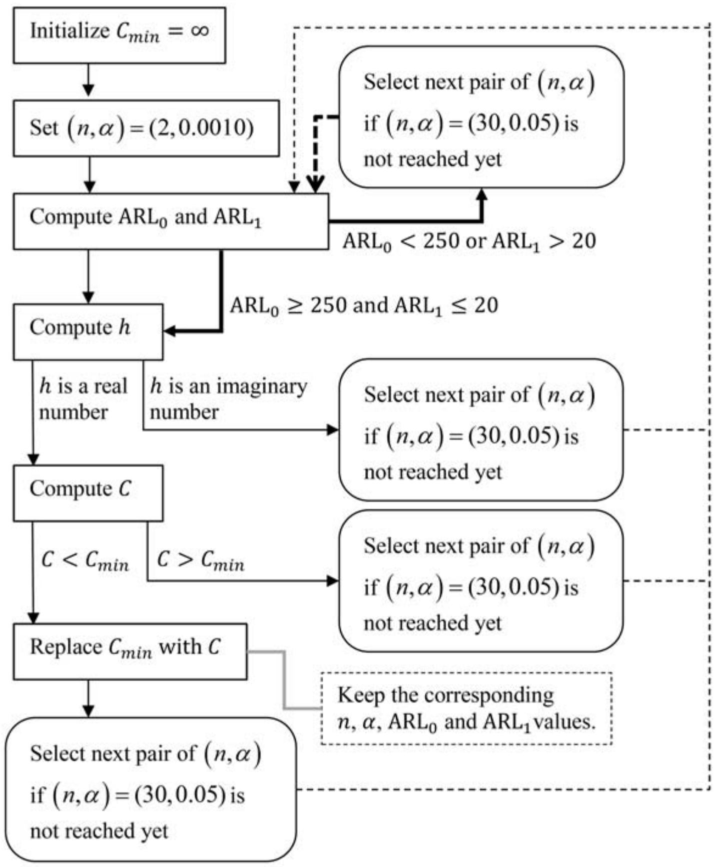

Figure 1: A flowchart explaining the minimization of the cost function, C in Equation (3.1), where thick arrows indicate additional steps for the economic-statistical design model.

The program starts with an assumingly large value of the cost per hour, C, which will be replaced by a new value of C each time a smaller one is obtained. For the controlled parameters, the first pair $(n, \alpha) = (2, 0.0010)$ is iteratively increased as $(2, 0.0011)$, $(2, 0.0012)$, ..., $(2, 0.05)$, $(3, 0.0010)$, $(3, 0.0011)$, ..., until it reaches (30, 0.05) for the economic design. However, for the economic-statistical design, the pair (n, α) is iteratively increased as $(2, 0.0010)$, $(2, 0.0011)$, ..., $(2, 0.004)$, $(3, 0.0010)$, $(3, 0.0011)$, ..., $(3, 0.004)$, ..., $(30, 0.0010)$, $(30, 0.0011)$, ..., $(30, 0.004)$. After the completion of all the iterations, the lowest cost per hour, C $(= C_{\min})$ is recorded, together with the corresponding optimal parameters n, α and h that produce the cost $C_{\min}$. The ARL_0 and ARL_1 values associated with these optimal parameter values are also recorded. Figure 1 shows a flowchart in minimizing C. In this flowchart, the statistical constraints imposed on the economic-statistical design of the MCV chart are shown as additional steps with thicker arrows.

4. NUMERICAL EXAMPLES

The thirteen input parameters and their values given in Lorenzen and Vance [13] for a real case problem of a casting operation process producing 84 castings per hour will be adopted in the numerical analyses in this section. These values are taken as the control values of the thirteen input parameters. In practice, the control values of these input parameters can be computed from historical data and prior knowledge of the process.

To demonstrate the computations of the control values of these thirteen input parameters in a real case problem, the following discussions adopted from Lorenzen and Vance [13] is provided. In this case study, the variable cost per unit sampled (c) is \$4.22 and it requires approximately 5 minutes to sample a single unit. The cost of each nonconforming unit produced is \$100. Historical data indicate that the process produces about 1.36% nonconforming units when it is in-control and about 11.3% nonconforming units when it is out-of-control, and the process stays in-control for an average of 50 hours. When an out-of-control signal is detected, a search for assignable cause is conducted. When one is found, the manufacturing system is stopped for repair, otherwise, the system is allowed to continue running. After repair is completed, the manufacturing system is restarted. The search for an assignable cause requires about 5 minutes, while repair requires 45 minutes. The repair cost is \$22.80 per hour and the downtime cost is \$21.34 per minute.

From the above paragraph, $\lambda = 1/50 = 0.02$ is the occurrence rate of assignable cause per hour. The time per unit sampled (e), expected search time during false alarm (T_0) and expected time to find the assignable cause (T_1) are $e = T_0 = T_1 = 5/60 = 0.083$ hour; while the expected time to repair the process is $T_2 = 45/60 = 0.75$ hour. During the search for the assignable cause, the process continues, thus $\varphi_1 = 1$, whereas the process is stopped during repair, hence, $\varphi_2 = 0$. The quality cost per hour while the process is in-control (C_0) is computed as follows: $C_0 = \$100$ (per nonconforming unit) $\times$ 84 (castings / units per hour) $\times$ 1.36% (nonconforming units) $= \$114.24$. Additionally, the quality cost per hour while the process is out-of-control (C_1) is calculated as follows: $C_1 = \$100 \times 84 \times 11.3\% = \$949.20$. Next, the cost of locating and removing the assignable cause (W) is obtained as the sum of the downtime cost and repair cost, i.e. $W = 45 \times \$21.34 + (45/60) \times \$22.80 = \$977.40$. It is assumed that the cost of false alarms (Y) is the same as the cost, W, hence, $Y = \$977.40$ is considered. Lastly, there is no fixed cost per sample, thus $b = \$0$.

Tables 3 and 4 provide the optimal parameters n, α and h of the MCV chart in minimizing the cost function, C in Equation (3.1), for the economic and economic-statistical designs of the aforementioned chart. The minimum cost, $C_{\min}$ and corresponding ARL_0 and ARL_1 values are also given in these tables. In Table 3, $p = 2$, $\gamma_0 = 0.1$ and $\tau = 0.5$ are considered for the downward MCV chart while in Table 4, $p = 2$, $\gamma_0 = 0.1$ and $\tau = 1.5$ are used for the upward MCV chart.

Table 3: Optimal parameters n, α and h in minimizing the cost function, C and the corresponding minimum cost ($C_{\min}$), ARL_0 and ARL_1 values computed for the downward MCV chart when $p = 2$, $\gamma_0 = 0.1$ and $\tau = 0.5$.

Case	Economic design						Economic-statistical design					
	n	α	h	$C_{\min}$	ARL_0	ARL_1	n	α	h	$C_{\min}$	ARL_0	ARL_1
1	13	0.0294	2.9112	206.7028	34.0136	1.1744	19	0.0039	2.8236	217.3567	256.4103	1.2426
2	14	0.0255	4.1072	**173.8845**	39.2157	1.1479	20	0.0039	4.1215	**180.1477**	256.4103	1.1865
3	12	0.0345	2.1108	*258.1688*	28.9855	1.2028	17	0.0040	1.8151	*275.8835*	250.0000	1.4009
4	13	0.0294	2.8124	**153.6580**	34.0136	1.1744	19	0.0039	2.7275	**164.9246**	256.4103	1.2426
5	13	0.0295	3.1474	*312.5772*	33.8983	1.1736	19	0.0039	3.0500	*321.9959*	256.4103	1.2426
6	14	0.0267	4.7588	**175.5130**	37.4532	1.1396	20	0.0039	4.7535	**180.7912**	256.4103	1.1865
7	12	0.0329	1.8891	*254.6568*	30.3951	1.2150	17	0.0040	1.6308	*274.8385*	250.0000	1.4009
8	11	0.0500	2.6535	**200.1941**	20.0000	1.1805	19	0.0039	2.7912	**216.7512**	256.4103	1.2426
9	15	0.0158	3.0772	*213.0151*	63.2911	1.1876	19	0.0039	2.8873	*218.5457*	256.4103	1.2426
10	13	0.0294	2.8935	**197.6308**	34.0136	1.1744	19	0.0039	2.8064	**208.3897**	256.4103	1.2426
11	13	0.0295	2.9507	*224.8406*	33.8983	1.1736	19	0.0039	2.8588	*235.2841*	256.4103	1.2426
12	13	0.0309	3.0492	*208.3568*	32.3625	1.1638	19	0.0039	2.9099	*219.0768*	256.4103	1.2426
13	13	0.0323	3.1805	*209.9396*	30.9598	1.1548	19	0.0039	2.9941	*220.7473*	256.4103	1.2426
14	14	0.0174	2.0716	**195.5321**	57.4713	1.2310	17	0.0040	1.7404	**200.6940**	250.0000	1.4009
15	11	0.0500	3.8064	*221.1789*	20.0000	1.1805	20	0.0039	4.2830	*240.5155*	256.4103	1.1865
16	14	0.0259	2.9664	**198.8200**	38.6100	1.1451	20	0.0039	2.9661	**206.0657**	256.4103	1.1865
17	11	0.0384	2.7917	*220.0642*	26.0417	1.2543	17	0.0040	2.4938	*236.8859*	250.0000	1.4009
18	13	0.0294	2.9112	206.7028	34.0136	1.1744	19	0.0039	2.8236	217.3567	256.4103	1.2426
19	13	0.0294	2.9112	206.7028	34.0136	1.1744	19	0.0039	2.8236	217.3567	256.4103	1.2426
20	13	0.0294	2.9092	**206.1229**	34.0136	1.1744	19	0.0039	2.8214	**216.7844**	256.4103	1.2426
21	13	0.0295	2.9183	*207.8738*	33.8983	1.1736	19	0.0039	2.8280	*218.5125*	256.4103	1.2426
22	13	0.0294	2.9140	*208.1516*	34.0136	1.1744	19	0.0039	2.8265	*218.8624*	256.4103	1.2426
23	13	0.0294	2.9056	**203.8647**	34.0136	1.1744	19	0.0039	2.8179	**214.4064**	256.4103	1.2426
24	13	0.0297	2.9074	**205.0555**	33.6700	1.1722	19	0.0039	2.8173	**215.8458**	256.4103	1.2426
25	13	0.0300	2.9575	*218.5185*	33.3333	1.1700	19	0.0039	2.8632	*229.2907*	256.4103	1.2426
26	13	0.0297	2.9620	*220.1732*	33.6700	1.1722	19	0.0039	2.8696	*230.8008*	256.4103	1.2426

In Tables 3 and 4, the italicized $C_{\min}$ values represent poorer performance (an increase in cost) while the boldfaced ones represent better performance (a decrease in cost) when the values of the input parameters are varied from the control values in case 1. The following discussions are based on the observations in Tables 3 and 4. It is found that the effects of changes in the input parameters on $C_{\min}$, ARL_0, ARL_1, n, α, and h for the economic design are almost similar to that for the economic-statistical design. In this section, the case number hereafter refers to the cases in Tables 3 and 4, unless stated otherwise.

Table 4: Optimal parameters n, α and h in minimizing the cost function, C and the corresponding minimum cost ($C_{\min}$), ARL_0 and ARL_1 values computed for the upward MCV chart when $p = 2$, $\gamma_0 = 0.1$ and $\tau = 1.5$.

Case	Economic design						Economic-statistical design					
	n	α	h	$C_{\min}$	ARL_0	ARL_1	n	α	h	$C_{\min}$	ARL_0	ARL_1
1	11	0.0286	1.8598	226.8698	34.9650	2.0070	13	0.0040	1.3199	240.2701	250.0000	2.9300
2	13	0.0287	2.9321	**188.9809**	34.8432	1.7783	15	0.0040	2.1137	**198.1568**	250.0000	2.5281
3	10	0.0294	1.2885	*283.8155*	34.0136	2.1403	10	0.0040	0.7368	*302.8548*	250.0000	3.8952
4	11	0.0285	1.7935	**174.4656**	35.0877	2.0088	13	0.0040	1.2741	**188.4030**	250.0000	2.9308
5	12	0.0295	2.1823	*331.3248*	33.8983	1.8689	13	0.0040	1.4282	*343.6606*	250.0000	2.9308
6	14	0.0316	3.7171	**188.3456**	31.6456	1.6559	17	0.0040	2.8399	**196.3893**	250.0000	2.2335
7	9	0.0260	1.0184	*283.2810*	38.4615	2.4111	10	0.0040	0.6580	*303.6778*	250.0000	3.8952
8	9	0.0500	1.7034	**217.5766**	20.0000	1.9849	12	0.0040	1.1850	**238.8530**	250.0000	3.1912
9	14	0.0154	2.0966	*235.8959*	64.9351	1.9456	15	0.0040	1.5939	*242.7479*	250.0000	2.5281
10	11	0.0286	1.8483	**217.9094**	34.9650	2.0070	13	0.0040	1.3117	**231.4024**	250.0000	2.9308
11	11	0.0286	1.8833	*244.7820*	34.9650	2.0070	13	0.0040	1.3367	*257.9956*	250.0000	2.9308
12	12	0.0322	2.1482	*229.2706*	31.0559	1.8291	15	0.0040	1.6083	*243.6569*	250.0000	2.5281
13	13	0.0353	2.4240	*231.4296*	28.3286	1.6959	17	0.0040	1.9016	*246.5326*	250.0000	2.2335
14	12	0.0144	1.2457	**211.4058**	69.4444	2.2442	12	0.0040	0.8688	**216.2331**	250.0000	3.1912
15	10	0.0500	2.6657	245.8762	20.0000	1.8487	15	0.0040	2.2106	*273.8320*	250.0000	2.5281
16	15	0.0300	2.4123	**219.5414**	33.3333	1.6013	20	0.0040	2.1005	**231.4886**	250.0000	1.9190
17	8	0.0258	1.3946	*237.4666*	38.7597	2.6647	9	0.0040	0.8917	*251.5718*	250.0000	4.3840
18	11	0.0286	1.8598	226.8698	34.9650	2.0070	13	0.0040	1.3199	240.2701	250.0000	2.9308
19	11	0.0286	1.8598	226.8698	34.9650	2.0070	13	0.0040	1.3199	240.2701	250.0000	2.9308
20	11	0.0286	1.8586	**226.3077**	34.9650	2.0070	13	0.0040	1.3189	**239.7114**	250.0000	2.9308
21	11	0.0286	1.8622	*228.0051*	34.9650	2.0070	13	0.0040	1.3220	*241.3983*	250.0000	2.9308
22	11	0.0286	1.8618	*228.4402*	34.9650	2.0070	19	0.0039	1.3214	**218.8624**	256.4103	2.9308
23	11	0.0286	1.8558	**223.7926**	34.9650	2.0070	19	0.0039	1.3169	**214.4064**	256.4103	2.9308
24	11	0.0291	1.8616	**225.1268**	34.3643	1.9978	19	0.0039	1.3166	**215.8458**	256.4103	2.9308
25	11	0.0296	1.8989	*238.5002*	33.7838	1.9889	19	0.0039	1.3383	**229.2907**	256.4103	2.9308
26	11	0.0290	1.8954	*240.2556*	34.4828	1.9996	19	0.0039	1.3417	**230.8008**	256.4103	2.9308

The thirteen input parameters of Lorenzen and Vance [13] cost model can be classified as expenses related parameters (C_0, C_1, Y, W, b, c), time related parameters (e, T_0, T_1, T_2) and process related parameters (λ, φ_1, φ_2). For a more effective and systematic way of discussing the effects of each input parameters on the minimum cost, ARLs and optimal parameters, this section is organized as follows: Firstly, the effects of expenses related parameters are discussed in Section 4.1, then those of time related parameters are enumerated in Section 4.2 and finally that of process related parameters are explained in Section 4.3. Additionally, the effects of the shift size τ in the process MCV is included in Section 4.3. Lastly, a comparison between economic and economic-statistical designs of the MCV chart is presented in Section 4.4.

4.1. Effects of expenses related parameters on $C_{\min}$, ARLs and optimal parameters

An increase in the quality cost (due to nonconformities produced) per hour while in-control, C_0 or out-of-control, C_1 results in an increase in the minimum cost, $C_{\min}$; and vice-versa (see cases 4–7). Although C_1 is larger than C_0 (Table 2, cases 4–7), C_0 has a more noticeable effect on the minimum cost ($C_{\min}$) as it results in a larger change in $C_{\min}$. It is also seen that an increase in C_0 (see case 5) or a decrease in C_1 (see case 6) leads to an increase in h, as compared to the control case (case 1). Note that a larger sampling interval, h is adopted when C_0 increases so that less frequent sampling is made when the process is in-control in order to offset the increase in quality cost per hour while the process is in-control. On a similar note, a decrease in C_1 indicates a lower quality cost per hour while the process is out-of-control, implying that sampling can be made less frequently (with an increase in h) so that the model remains economically viable. The same explanation applies for a decrease in h when C_0 decreases or C_1 increases.

Another cost parameter worthy of discussion is the cost of false alarm, Y. It is found that increasing (decreasing) Y only results in a slight increase (decrease) in the minimum cost, $C_{\min}$ but it substantially increases (decreases) the ARL_0 value for the economic design of the chart (see cases 8 and 9). An increased (decreased) ARL_0 value translates into a lower (higher) false alarm rate, hence a smaller (larger) α value (see case 9 for the economic design). A larger cost of false alarm (see case 9 in Table 2, where $Y = \$1954.8$ instead of the control value of \$977.4) will reduce the sampling frequency (larger h of 3.0772 instead of the control value of 2.9112 — see Table 3) for the economic design model. To compensate for the less frequent sampling, a larger sample size (larger n, increasing from 13 to 15) is adopted (see cases 1 and 9 for the economic design in Table 3). Note that the effect of changing Y on the optimal parameters, minimum cost and ARLs under the economic-statistical design model is less pronounced.

Comparing to Y, varying the cost to locate and remove the assignable cause, W poses no significant changes to the optimal parameters n, α and h. However, W has a greater influence on the minimum cost $C_{\min}$ than Y. As an example, increasing W from \$977.4 to \$1954.8 (see case 11 in Table 2) causes $C_{\min}$ to increase from \$226.8698 to \$244.7820 (see case 11 for economic design in Table 4) while the same amount of increment in Y (see case 9 in Table 2) results in a smaller increase in $C_{\min}$, i.e. from \$226.8698 to \$235.8959 (see case 9 for the economic design in Table 4). Likewise, $C_{\min}$ decreases at a quicker rate when W decreases compared to that for the same amount of a decrease in Y. Using another example based on the economic-statistical design in Table 3, decreasing W and Y to half of their original values causes $C_{\min}$ to decrease by \$8.9670 (i.e. \$217.3567 − \$208.3897 or the difference between $C_{\min}$ of cases 1 and 10) versus \$0.6055 (i.e. \$217.3567 − \$216.7512 or the difference between $C_{\min}$ of cases 1 and 8), respectively.

The sampling cost is affected by two different parameters, namely the fixed cost per sample, b and the variable cost per unit sampled, c. The control value of b is \$0. When b increases to \$5 and \$10, it is found that the minimum cost, $C_{\min}$ for case 13 is larger than that for case 12 but the $C_{\min}$ values for these two cases are larger than the control cost in case 1. In fact, increasing any cost parameter, including the variable cost per unit sampled, c will always result in an increase in $C_{\min}$, as expected. Increasing the cost b and (or) c (see cases 12, 13 and 15) results in a larger optimal sampling interval (larger h) for both economic and economic-statistical design models and a smaller ARL_0 value for the economic design model. The exact opposite results are observed by decreasing c (case 14 in Table 2), which results in smaller h, lower $C_{\min}$ and larger ARLs (see the economic design for both downward and upward charts in Tables 3 and 4). Note that the ARL_0 values in Tables 3 and 4 do not vary much in the economic-statistical design model in satisfying the constraint $\text{ARL}_0 \geq 250$.

4.2. Effects of time related parameters on $C_{\min}$, ARLs and optimal parameters

Besides the expenses related parameters, Lorenzen and Vance's [13] cost model also includes the time related parameters, namely e, T_0, T_1 and T_2. Other than the time to sample and interpret one unit (e), the remaining time related parameters have minimal effect on the optimal parameters, $C_{\min}$, ARL_0 and ARL_1 values (see cases 18–23). An increase (decrease) in e causes the minimum cost, $C_{\min}$ to increase (decrease) (see cases 16 and 17). As e increases (from 0.083 hours to 0.166 hours), smaller sample sizes (for example, see case 17, where $n = 11$ in Table 3 and $n = 8$ in Table 4 for the economic design) are adopted to offset the increase in $C_{\min}$. Consequently, shorter sampling intervals (see case 17, where $h = 2.7917$ hours in Table 3 and $h = 1.3946$ hours in Table 4 for the economic design) are adopted as more frequent samplings are needed to compensate for the smaller sample sizes used. In addition, increasing (decreasing) the value of e leads to a larger (smaller) ARL_1 value (see cases 16 and 17). Using an example from the economic-statistical design, increasing e causes ARL_1 to increase from 1.2426 to 1.4009 for the downward MCV chart (see case 17 in Table 3) and from 2.9300 to 4.3840 for the upward MCV chart (see case 17 in Table 4). In addition, decreasing e causes ARL_1 to decrease from 1.2426 to 1.1865 for the downward MCV chart (see case 16 in Table 3) and from 2.9300 to 1.9190 for the upward MCV chart (see case 16 in Table 4).

4.3. Effects of process related parameters on $C_{\min}$, ARLs and optimal parameters

The rate of occurrence of assignable cause, λ has a significant effect on the optimal sample size, n, optimal sampling interval, h and minimum cost, $C_{\min}$ (see cases 2 and 3). For example, when λ decreases from 0.02 to 0.01 (see cases 1 and 2), $C_{\min}$ decreases from \$206.7028 to \$173.8845 (see case 2 for the economic design in Table 3) because the process failure rate decreases. In contrast, when λ increases from 0.02 to 0.04 (see cases 1 and 3), $C_{\min}$ increases from \$206.7028 to \$258.1688 (see case 3 for the economic design in Table 3). To enable this undesirable condition (an increase in λ) to be detected quickly by the MCV chart, more frequent samplings (decreasing h) are needed while smaller sample sizes (decreasing n) are adopted in order to remain economically favourable (see cases 1 and 3 in Tables 3 and 4, for both economic and economic statistical designs).

The parameters φ_1 and φ_2 determine whether the process continues or stops during search and repair, respectively. As shown in Table 1, φ_1 (φ_2) has:

(i) the value 1 if the process continues while *searching for the assignable cause (repairing following the occurrence of an assignable cause)*;

(ii) the value 0 if the process stops during *search (repair)*.

By comparing cases 1, 24, 25 and 26, it is observed that case 24 (where $(\varphi_1, \varphi_2 = (0, 0))$ has the lowest minimum cost, $C_{\min}$ (see Tables 3 and 4). This is expected because when the process stops during both search and repair, the cost will be minimized. For example, for the economic design in Table 3, $C_{\min} \in \{\$205.0555, 206.7028, 218.5185, 220.1732\}$ for $(\varphi_1, \varphi_2) \in \{(0, 0), (1, 0), (0, 1), (1, 1)\}$, where the lowest $C_{\min}$ $(= \$205.0555)$ occurs at $(\varphi_1, \varphi_2) = (0, 0)$, i.e. when the process stops during both search and repair. On the contrary, case 26, i.e. the process continues during both search and repair ($(\varphi_1, \varphi_2) = (1, 1)$) undoubtedly results in the highest minimum cost, $C_{\min}$. Note that the effect of the same pair of (φ_1, φ_2) values on $C_{\min}$ is similar for both economic and economic-statistical designs of the downward and upward MCV charts.

Another interesting observation obtained is the influence of the shift, τ $(= \gamma_1/\gamma_0)$ on the minimum cost, $C_{\min}$. Table 3 deals with a 50% decreasing shift in the process MCV while Table 4 involves an increasing shift of 50%, hence, the size of shifts in both tables is the same. It is found that for the same size of shift in the process MCV, generally, the upward MCV chart incurs a higher $C_{\min}$ than that of the downward MCV chart. As an example, for the economic-statistical design in Table 3, $C_{\min} \in \{\$217.3567, 180.1477, 275.8835, 164.9246, 321.9959\}$ while in Table 4, $C_{\min} \in \{\$240.2701, 198.1568, 302.8548, 188.4030, 343.6606\}$ for cases 1, 2, 3, 4 and 5, respectively. This example clearly shows that $C_{\min}$ for the upward MCV chart is higher than the corresponding one for the downward MCV chart. It is noteworthy that a larger $C_{\min}$ for the upward MCV chart corresponds to detecting an increasing shift ($\tau = 1.5$) in the process MCV, which simply means process deterioration. In contrast, a smaller $C_{\min}$ incurred by the downward MCV chart is associated with the detection of a decreasing MCV shift ($\tau = 0.5$) or simply process improvement. As $C_{\min}$ incurred by the upward MCV chart is higher, smaller sample sizes, n must be adopted by this chart to offset the increase in cost. This is evident as n in Table 4 is generally lower than the corresponding one in Table 3. For example, based on the economic-statistical design in cases 1, 2, 3, 4 and 5, it is noticed that $n \in \{19, 20, 17, 19, 19\}$ and $n \in \{13, 15, 10, 13, 13\}$ in Tables 3 and 4, respectively, where it is obvious that the sample sizes in Table 4 are lower than the corresponding ones in Table 3. Consequently, to compensate for the smaller sample sizes adopted by the upward MCV chart in Table 4, samples must be taken more frequently, hence a smaller sampling interval, h is adopted. For the same example, $h \in \{1.3199, 2.1137, 0.7368, 1.2741, 1.4282\}$ are adopted for cases 1–5 in Table 4 while $h \in \{2.8236, 4.1215, 1.8151, 2.7275, 3.0500\}$ are employed for the same cases in Table 3. Evidently, the h values in Table 4 are smaller than that in Table 3.

4.4. Comparisons between economic and economic-statistical designs

It is shown in Tables 3 and 4 that imposing statistical constraints in the economic design of the MCV chart significantly improves the statistical performance of the chart as it results in larger ARL_0 values at the expense of slight increases in the minimum cost ($C_{\min}$) and ARL_1 values. For a better analysis, Table 5 shows the percentage of increase in the ARL_0 value for each of the 26 cases in Table 2 when the economic-statistical design is used in place of the economic design. Additionally, Table 5 shows the percentage of the slight increase in the minimum cost ($C_{\min}$) and ARL_1 values as a result of adding the statistical constraints (in the economic-statistical design). In Table 5, $p = 2$ and $\gamma_0 = 0.1$ are considered for the downward ($\tau = 0.5$) and upward ($\tau = 1.5$) MCV charts. It is found in Table 5 that by employing the economic-statistical design model, the ARL_0 value increases by at least 305.13% (case 9) and 260% (case 14), for the downward and upward MCV charts, respectively. In contrast, the chart's performances in terms of $C_{\min}$ and ARL_1 criteria only deteriorate slightly. For example, $C_{\min}$ increases by at most 8.79% (case 15) for the downward MCV chart and 11.42% (case 15) for the upward MCV chart. On similar lines, the ARL_1 increases by at most 16.47% (case 3) and 81.99% (case 3) for the downward and upward MCV charts, respectively.

Table 5: Percentages of increase in the minimum cost ($C_{\min}$), ARL_0 and ARL_1 values by using economic-statistical design in place of economic design for the downward and upward MCV charts when $p = 2$ and $\gamma_0 = 0.1$.

Case	Downward MCV chart			Upward MCV chart		
	% increase in $C_{\min}$	% increase in ARL_0	% increase in ARL_1	% increase in $C_{\min}$	% increase in ARL_0	% increase in ARL_1
1	5.15	653.85	5.81	5.91	615.00	45.99
2	3.60	553.85	3.36	4.86	617.50	42.16
3	6.86	762.50	16.47	6.71	635.00	81.99
4	7.33	653.85	5.81	7.99	612.50	45.90
5	3.01	656.41	5.88	3.72	637.50	56.82
6	3.01	584.62	4.12	4.27	690.00	34.88
7	7.93	722.50	15.30	7.20	550.00	61.55
8	8.31	1510.25	4.02	9.79	1255.00	64.56
9	2.60	305.13	4.63	2.90	285.00	29.94
10	5.44	653.85	5.81	6.19	615.00	46.03
11	4.64	656.41	5.88	5.40	615.00	46.03
12	5.15	692.31	6.77	6.27	705.00	38.22
13	5.15	728.20	7.60	6.53	782.50	31.70
14	2.64	335.00	13.80	2.28	260.00	42.20
15	8.79	1325.64	2.67	11.42	1342.50	42.08
16	3.64	564.10	3.62	5.44	650.00	19.84
17	7.64	860.00	11.69	5.94	545.00	64.52
18	5.15	653.85	5.81	5.91	615.00	46.03
19	5.15	653.85	5.81	5.91	615.00	46.03
20	5.17	653.85	5.81	5.92	615.00	46.03
21	5.12	656.41	5.88	5.87	615.00	46.03
22	5.15	653.85	5.81	5.90	615.00	46.03
23	5.17	653.85	5.81	5.92	615.00	46.03
24	5.26	661.54	6.01	6.04	627.50	46.70
25	4.93	669.23	6.21	5.74	640.00	47.36
26	4.83	661.54	6.01	5.61	625.00	46.57
Average	5.26	697.55	6.78	5.99	653.65	46.59

The last row in Table 5 shows the average percentages of increase in $C_{\min}$, ARL_0 and ARL_1 values when the economic-statistical design is used instead of the economic design. For the downward MCV chart, it is found that there is a huge average increase in the ARL_0 value, i.e. 697.55% as compared to significantly smaller average increase in $C_{\min}$ and ARL_1 values, i.e. at only 5.26% and 6.78%, respectively. Similarly, for the upward MCV chart, a large average increase in ARL_0, i.e. 653.65% is obtained at the expense of enormously smaller average increases in $C_{\min}$ (5.99%) and ARL_1 (46.59%) values. It is obviously seen in Table 5 that when the economic-statistical design is adopted in lieu of the economic design, the downward MCV chart (average increase of 6.78%) results in a smaller increase in the value compared to the upward MCV chart (average increase of 46.59%).

Additional analyses are conducted for the number of correlated variables, $p \geq 3$, where the same trends as that for $p = 2$ are observed. Thus, the results for $p \geq 3$ are not given here so as not to increase the length of this manuscript unnecessarily.

5. CONCLUSIONS

The MCV chart is used in the monitoring of the process MCV. The use of the MCV chart in process monitoring requires not only the statistical consideration in assessing its performance but also from a cost point of view. In line with this requirement, this research studies the economic and economic-statistical designs of the MCV chart. The economic design takes into account of minimizing the cost, but it ignores the statistical evaluation of the chart. Therefore, the economic design exposes the MCV chart to a poor statistical performance, resulting in an undesirable Type-I error rate. To circumvent this setback, statistical constraints, in terms of the ARL_0 and ARL_1 considerations, are imposed on the cost minimization model, resulting in the economic-statistical design of the chart. The effects of changes in the input parameters on the minimum cost and the corresponding optimal parameters of the MCV chart, as well as the char's ARL_0 and ARL_1 values are enumerated. Additionally, this work also compares the impact of adding statistical constraints on the performance of the MCV chart. It is found that the economic-statistical design significantly improves the ARL_0 performance of the MCV chart at the expense of slight increases in minimum cost ($C_{\min}$) and ARL_1 values.

ACKNOWLEDGMENTS

This research is supported by the Kementerian Pendidikan Malaysia, Fundamental Research Grant Scheme, number 203.PMATHS.6711603. This research was conducted while the corresponding author is on sabbatical leave at Universiti Tunku Abdul Rahman, Kampar.

REFERENCES

[1] Abbasi, S.A. and Adegoke, N.A. (2018). Multivariate coefficient of variation control charts in phase I of SPC, *The International Journal of Advanced Manufacturing Technology*, **99**(5–8), 1903–1916.

[2] Albert, A. and Zhang, L. (2010). A novel definition of the multivariate coefficient of variation, *Biometrical Journal*, **52**(5), 667–675.

[3] Carstensen, B. (2010). *Comparing Clinical Measurement Methods: A Practical Guide*, John Wiley & Sons, Ltd, Chapter **23**, 107–114.

[4] Castagliola, P.; Celano, G. and Psarakis, S. (2011). Monitoring the coefficient of variation using EWMA charts, *Journal of Quality Technology*, **43**(3), 249–265.

[5] Duncan, A.J. (1956). The economic design of $\bar{X}$ charts used to maintain current control of a process, *Journal of the American Statistical Association*, **51**(274), 228–242.

[6] Giner-Bosch, V.; Tran, K.P.; Castagliola, P. and Khoo, M.B.C. (2019). An EWMA control for the multivariate coefficient of variation, *Quality and Reliability Engineering International*, **35**(6), 1515–1541.

[7] Golub, T.R.; Slonim, D.K.; Tomayo, P.; Huard, C.; Gaasenbeek, M.; Mesirov, J.P.; Coller, H.; Loh, M.L.; Downing, J.R.; Caligiuri, M.A.; Bloomfield, C.D. and Lander, E.S. (1999). Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, *Science*, **286**(5439), 531–537.

[8] Hua, Y.; Qiu, W.; Xiao, Q. and Wu, Q. (2018). Precision (repeatability and reproducibility) of ocular parameters obtained by the tomey OA-2000 biometer compared to the IOLMaster in healthy eyes, *PLoS ONE*, **13**(2), e0193023.

[9] Kang, C.W.; Lee, M.S.; Seong, Y.J. and Hawkins, D.M. (2007). A control chart for the coefficient of variation, *Journal of Quality Technology*, **39**(2), 151–158.

[10] Khatun, M.; Khoo, M.B.C.; Lee, M.H. and Castagliola, P. (2019). One-sided control charts for monitoring the multivariate coefficient of variation in short production runs, *Transactions of the Institute of Measurement and Control*, **41**(6), 1712–1728.

[11] Khaw, K.W.; Khoo, M.B.C.; Castagliola, P. and Rahim, M.A. (2018). New adaptive control charts for monitoring the multivariate coefficient of variation, *Computers & Industrial Engineering*, **126**, 595–610.

[12] Linderman, K. and Love, T.E. (2000). Economic and economic statistical designs for MEWMA control charts, *Journal of Quality Technology*, **32**(4), 410–417.

[13] Lorenzen, T.J. and Vance, L.C. (1986). The economic design of control chart: a unified approach, *Technometrics*, **28**(1), 3–10.

[14] Molnau, W.E.; Montgomery, D.C. and Runger, G.C. (2011). Statistically constrained economic design of the multivariate exponentially weighted moving average control chart, *Quality and Reliability Engineering International*, **17**(1), 39–49.

[15] Ng, P.S.; Khoo, M.B.C.; Saha, S. and Yeong, W.C. (2019). Economic and economic-statistical designs of auxiliary information based $\bar{X}$, synthetic and EWMA charts, *Communications in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2019.1664575.

[16] Nguyen, Q.T.; Tran, K.P.; Castagliola, P.; Celano, G. and Lardjane, S. (2019). One-sided synthetic control charts for monitoring the multivariate coefficient of variation, *Journal of Statistical Computation and Simulation*, **89**(10), 1841–1862.

[17] Onkelinx, S.; Cornelissen, V.; Goetschalckx, K.; Thomaes, T.; Verhamme, P. and Vanhees, L. (2012). Reproducibility of different methods to measure the endothelial function, *Vascular Medicine*, **17**(2), 79–84.

[18] Peruchi, R.S.; Balestrassi, P.P.; de Paiva, A.P.; Ferreira, J.R. and de Santana Carmelossi, M. (2013). A new multivariate gage R&R method for correlated characteristics, *International Journal of Production Economics*, **144**(1), 301–315.

[19] Reed, G.F.; Lynn, F. and Meade, B.D. (2002). Use of coefficient of variation in assessing variability of quantitative assays, *Clinical and Diagnostic Laboratory Immunology*, **9**(6), 1235–1239.

[20] Reinstein, D.Z.; Archer, T.J.; Silverman, R.H. and Coleman, D.J. (2006). Accuracy, repeatability, and reproducibility of artemis very high-frequency digital ultrasound arc-scan lateral dimension measurements, *Journal of Cataract & Refractive Surgery*, **32**(11), 1799–1802.

[21] Reyment, R.A. (1960). Studies on nigerian upper cretaceous and lower tertiary ostracoda: part 1, senonian and maastrichtian ostracoda, *Stockholm Contributions in Geology*, **7**, 1–238.

[22] Safe, H.; Kazemzadeh, R.B. and Gholipour Kanami, Y. (2018). A Markov chain approach for double-objective economic statistical design of the variable sampling interval control chart, *Communications in Statistics – Theory and Methods*, **47**(2), 277–288.

[23] Saniga, E.M. (1989). Economic statistical control-chart designs with an application to $\bar{X}$ and R charts, *Technometrics*, **31**(3), 313–320.

[24] Sharpe, W.F. (1994). The sharpe ratio, *The Journal of Portfolio Management*, **21**(1), 49–58.

[25] Soulé, M.E. and Zegers, G.P. (1996). Phenetics of natural populations. V. Genetic correlates of phenotypic variation in the pocket gopher (Thomomys bottae) in California, *The Journal of Heredity*, **87**(5), 341–350.

[26] Surtihadi, J. and Raghavachar, M. (1994). Exact economic design of $\bar{X}$ charts for general time in-control distributions, *International Journal of Production Research*, **32**(7), 2287–2302.

[27] Sweeney, S. (2007). Analysis of two-dimensional gage repeatability and reproducibility, *Quality Engineering*, **19**(1), 29–37.

[28] Teoh, W.L.; Khoo, M.B.C.; Castagliola, P.; Yeong, W.C. and Teh, S.Y. (2017). Run-sum control charts for monitoring the coefficient of variation, *European Journal of Operational Research*, **257**(1), 144–158.

[29] Van Valen, L. (1974). Multivariate structural statistics in natural history, *Journal of Theoretical Biology*, **45**(1), 235–247.

[30] Voinov, V.G. and Nikulin, M.S. (1996). *Unbiased Estimators and Their Applications, Multivariate Case*, Vol. 2, Kluwer, Dordrecht.

[31] Wan, Q. and Zhu, M. (2021). Detecting a shift in variance using economically designed VSI control chart with combined attribute-variable inspection, *Communications in Statistics – Simulation and Computation*, **50**(11), 3483–3497.

[32] Yeong, W.C.; Khoo, M.B.C.; Teoh, W.L. and Castagliola, P. (2016). A control chart for the multivariate coefficient of variation, *Quality and Reliability Engineering International*, **32**(3), 1213–1225.

[33] Yeong, W.C.; Khoo, M.B.C.; Wu, Z. and Castagliola, P. (2012). Economically optimum design of a synthetic $\bar{X}$ chart, *Quality and Reliability Engineering International*, **28**(7), 725–741.

[34] Yeong, W.C.; Lim, S.L.; Khoo, M.B.C.; Chuah, M.H. and Lim, A.J.X. (2018). The economic and economic-statistical designs of the synthetic chart for the coefficient of variation, *Journal of Testing and Evaluation*, **46**(3), 1175–1195.

[35] Zhang, L.; Albarède, S.; Dumont, G.; Campenhout, C.V.; Libeer, J.C. and Albert, A. (2010). The multivariate coefficient of variation for comparing serum protein electrophoresis techniques in external quality assessment schemes, *Accreditation and Quality Assurance*, **15**(6), 351–357.

[36] Zhang, L.; Campenhout, C.V.; Devleeschouwer, N.; Libeer, J.C. and Albert, A. (2008). Statistical analysis of serum protein electrophoresis results in external quality assessment schemes, *Accreditation and Quality Assurance*, **13**(3), 149–155.

REVSTAT-Statistical journal

Aims and Scope

The aim of REVSTAT-Statistical Journal is to publish articles of high scientific content, developing Statistical Science focused on innovative theory, methods, and applications in different areas of knowledge. Important survey/review contributing to Probability and Statistics advancement is also welcome.

Background

Statistics Portugal started in 1996 the publication of the scientific statistical journal *Revista de Estatística*, in Portuguese, a quarterly publication whose goal was the publication of papers containing original research results, and application studies, namely in the economic, social and demographic fields. Statistics Portugal was aware of how vital statistical culture is in understanding most phenomena in the present-day world, and of its responsibilities in disseminating statistical knowledge.

In 1998 it was decided to publish papers in English. This step has been taken to achieve a larger diffusion, and to encourage foreign contributors to submit their work. At the time, the editorial board was mainly composed by Portuguese university professors, and this has been the first step aimed at changing the character of *Revista de Estatística* from a national to an international scientific journal. In 2001, the *Revista de Estatística* published a three volumes special issue containing extended abstracts of the invited and contributed papers presented at the 23rd European Meeting of Statisticians (EMS). During the EMS 2001, its editor-in-chief invited several international participants to join the editorial staff.

In 2003 the name changed to REVSTAT-Statistical Journal, published in English, with a prestigious international editorial board, hoping to become one more place where scientists may feel proud of publishing their research results.

Editorial policy

REVSTAT-Statistical Journal is an open access peer-reviewed journal published quarterly, in English, by Statistics Portugal.

The editorial policy of REVSTAT is mainly placed on the originality and importance of the research. The whole submission and review processes for REVSTAT are conducted exclusively online on the journal's webpage revstat.ine.pt based in Open Journal System (OJS). The only working language allowed is English. Authors intending to submit any work must register, login and follow the guidelines.

There are no fees for publishing accepted manuscripts that will be made available in open access.

All articles consistent with REVSTAT aims and scope will undergo scientific evaluation by at least two reviewers, one from the Editorial Board and another external. Authors can suggest an editor or reviewer who is expert on the paper subject providing her/his complete information, namely: name, affiliation, email and, if possible, personal URL or ORCID number.

All published works are Open Access (CC BY 4.0) which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. Also, in the context of archiving policy, REVSTAT is a *blue* journal welcoming authors to deposit their works in other scientific repositories regarding the use of the published edition and providing its source.

Journal prints may be ordered at expenses of the author(s), and prior to publication.

Abstract and Indexing services

REVSTAT-Statistical Journal is covered *by Journal Citation Reports - JCR (Clarivate); Current Index to Statistics; Google Scholar; Mathematical Reviews® (MathSciNet®); Zentralblatt für Mathematic; Scimago Journal & Country Rank; Scopus*

Author guidelines

The whole submission and review processes for REVSTAT are conducted exclusively online on the journal's webpage https://revstat.ine.pt/ based in Open Journal System (OJS). Authors intending to submit any work must ***register, login*** and follow the indications choosing ***Submissions***.

REVSTAT - Statistical Journal adopts the COPE guidelines on publication ethics.

Work presentation

- the only working language is English;
- the first page should include the name, ORCID iD (optional), Institution, country, and mail-address of the author(s);
- a summary of fewer than one hundred words, followed by a maximum of six keywords and the MSC 2020 subject classification should be included also in the first page;
- manuscripts should be typed only in black, in double-spacing, with a left margin of at least 3 cm, with numbered lines, and a maximum of 25 pages;
- the title should be with no more than 120 characters (with spaces);
- figures must be a minimum of 300dpi and will be reproduced online as in the original work, however, authors should take into account that the printed version is always in black and grey tones;
- authors are encouraged to submit articles using LaTeX which macros are available at *REVSTAT style*;
- citations in text should be included in the text by name and year in parentheses, as in the following examples: § article title in lowercase (Author 1980); § This theorem was proved later by AuthorB and AuthorC (1990); § This

subject has been widely addressed (AuthorA 1990; AuthorB et al. 1995; AuthorA and AuthorB 1998).

- references should be listed in alphabetical order of the author's scientific surname at the end of the article;
- acknowledgments of people, grants or funds should be placed in a short section before the References title page. Note that religious beliefs, ethnic background, citizenship and political orientations of the author(s) are not allowed in the text;
- authors are welcome to suggest one of the Editors or Associate Editors or yet other reviewer expert on the subject providing a complete information, namely: name, affiliation, email and personal URL or ORCID number in the Comments for the Editor (submission form).

Accepted papers

After final revision and acceptance of an article for publication, authors are requested to provide the corresponding LaTeX file, as in REVSTAT style.

Supplementary files may be included and submitted separately in .tiff, .gif, .jpg, .png .eps, .ps or .pdf format. These supplementary files may be published online along with an article, containing data, programming code, extra figures, or extra proofs, etc; however, REVSTAT is not responsible for any supporting information supplied by the author(s).

Copyright Notice

Upon acceptance of an article, the author(s) will be asked to transfer copyright of the article to the publisher, Statistics Portugal, in order to ensure the widest possible dissemination of information.

According to REVSTAT's *archiving policy*, after assigning the copyright form, authors may cite and use limited excerpts (figures, tables, etc.) of their works accepted/published in REVSTAT in other publications and may deposit only the published edition in scientific repositories providing its source as REVSTAT while the original place of publication. The Executive Editor of the Journal must be notified in writing in advance.

EDITORIAL BOARD 2019-2023

Editor-in-Chief
Isabel FRAGA ALVES, University of Lisbon, Portugal

Co-Editor
Giovani L. SILVA, University of Lisbon, Portugal

Associate Editors
Marília ANTUNES, University of Lisbon, Portugal
Barry ARNOLD, University of California, USA
Narayanaswamy BALAKRISHNAN, McMaster University, Canada
Jan BEIRLANT, Katholieke Universiteit Leuven, Belgium
Graciela BOENTE, University of Buenos Aires, Argentina
Paula BRITO, University of Porto, Portugal
Valérie CHAVEZ-DEMOULIN, University of Lausanne, Switzerland
David CONESA, University of Valencia, Spain
Charmaine DEAN, University of Waterloo, Canada
Fernanda FIGUEIREDO, University of Porto, Portugal
Jorge Milhazes FREITAS, University of Porto, Portugal
Alan GELFAND, Duke University, USA
Stéphane GIRARD, Inria Grenoble Rhône-Alpes, France
Marie KRATZ, ESSEC Business School, France
Victor LEIVA, Pontificia Universidad Católica de Valparaíso, Chile
Artur LEMONTE, Federal University of Rio Grande do Norte, Brazil
Shuangzhe LIU, University of Canberra, Australia
Maria Nazaré MENDES-LOPES, University of Coimbra, Portugal
Fernando MOURA, Federal University of Rio de Janeiro, Brazil
John NOLAN, American University, USA
Paulo Eduardo OLIVEIRA, University of Coimbra, Portugal
Pedro OLIVEIRA, University of Porto, Portugal
Carlos Daniel PAULINO, University of Lisbon, Portugal
Arthur PEWSEY, University of Extremadura, Spain
Gilbert SAPORTA, Conservatoire National des Arts et Métiers, France
Alexandra M. SCHMIDT, McGill University, Canada
Julio SINGER, University of Sao Paulo, Brazil
Manuel SCOTTO, University of Lisbon, Portugal
Lisete SOUSA, University of Lisbon, Portugal
Milan STEHLÍK, University of Valparaiso, Chile and LIT-JK University Linz, Austria
María Dolores UGARTE, Public University of Navarre, Spain

Executive Editor
José A. PINTO MARTINS, Statistics Portugal

Assistant Editors
José CORDEIRO, Statistics Portugal
OLGA Bessa MENDES, Statistics Portugal