ISSN 1645-6726

REVSTAT

Statistical Journal

vol. 22 - n. 3 - July 2024

Instituto Nacional de Estatística
Statistics Portugal

REVSTAT — Statistical Journal, vol.22, n. 3 (July 2024)

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

Continues: Revista de Estatística = ISSN 0873-4275.

ISSN 1645-6726 ; e-ISSN 2183-0371

Publisher – Statistics Portugal

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

Edition - 130 copies | **Legal Deposit Registration** - 191915/03 | **Price** [VAT included] - € 9,00

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

INDEX

Abstracted/indexed in: *Journal Citation Reports™ – JCR (Clarivate); DOAJ – Directory of Open Access Journals; Current Index to Statistics; Google Scholar; Mathematical Reviews® (MathSciNet®); Zentralblatt für Mathematic; SCImago Journal & Country Rank; Scopus®.*

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 277–292
https://doi.org/10.57805/revstat.v22i3.508

Estimation of Population Variance for a Sensitive Variable in Stratified Sampling using Randomized Response Technique

Authors: Badr Aloraini
– Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro, North Carolina, USA
boalorai@uncg.edu

Sadia Khalil
– Department of Statistics, Lahore College for Women University, Lahore, Pakistan
sadia_khalil@hotmail.com

Muhammad Nouman Qureshi
– Department of Statistics, National College of Business Administration and Economics, Lahore, Pakistan
nqureshi633@gmail.com

Sat Gupta ✉
– Department of Mathematics and Statistics, University of North Carolina at Greensboro, Greensboro, North Carolina, USA
sngupta@uncg.edu

Received: May 2021 Revised: June 2022 Accepted: June 2022

Abstract:

- In this paper, Randomized Response technique (RRT) is used to propose some separate and combined variance estimators for a sensitive variable using stratified random sampling. The performances of the proposed estimators are examined using a unified measure of respondent privacy and estimator efficiency.

Keywords:

- *auxiliary information; mean squared error; stratified random sampling; respondent privacy; variance estimation.*

AMS Subject Classification:

- 62D05.

✉ Corresponding author

1. INTRODUCTION

Our main focus in this study is on variance estimation for sensitive variables in stratified sampling. Many researchers have dealt with the problem of mean and variance estimation under simple random sampling and stratified random sampling when the study variable is non-sensitive and is directly observable. Zahid and Shabbir (2018) [17] and many other authors have investigated the problem of mean estimation in stratified random sampling when the study variable is non-sensitive. Important contributions in the area of variance estimation in stratified random sampling for non-sensitive random variables have been made by Kadilar and Cingi (2006) [8], Sidelel *et al.* (2014) [12], Özel *et al.* (2014) [10], Clement (2018) [2], Sanaullah *et al.* (2017) [11], Younis and Shabbir (2019) [16], and Asghar *et al.* (2019) [1]. In all of these studies, the study variable is directly observed and an auxiliary variable is used to increase the efficiency of estimation.

In research involving sensitive survey questions, standard estimation techniques are unreliable. Warner(1965) [14] introduced the Randomized Response Technique (RRT) as a research method to reduce response Bias in estimation of a sensitive study variable and at the same time improve the respondent cooperation. Many authors, including Kalucha *et al.* (2017) [9] and Zhang *et al.* (2021) [18], have estimated the mean of a sensitive study variable under stratified sampling. However, not much work exists for variance estimation under RRT. Gupta *et al.* (2020) [5] introduced several variance estimators under RRT in simple random sampling. The primary goal of this study is to re-examine the Gupta *et al.* (2020) [5] study in the context of stratified random sampling.

Let us consider Y and X to be the observed and auxiliary variables defined on a finite population $U = \{U_1, U_2, ..., U_N\}$. We assume that Y is sensitive in nature and we observe a scrambled version of it given by $Z = TY + S$, where T, S, Y and X are mutually uncorrelated. Let the population be divided into L homogeneous strata with N_h unites $(h = 1, 2, ..., L)$ in the h^{th} stratum such that $\sum_{h=1}^{L} N_h = N$. From h^{th} stratum, a simple random sample of size n_h is drawn without replacement such that $\sum_{h=1}^{L} n_h = n$. Let (x_{hi}, y_{hi}, z_{hi}) be the observed values on the variables X, Y, and Z in the h^{th} stratum. Let $\bar{x}_{st} = \sum_{h=1}^{L} W_h \bar{x}_h$, $\bar{y}_{st} = \sum_{h=1}^{L} W_h \bar{y}_h$, $\bar{z}_{st} = \sum_{h=1}^{L} W_h \bar{z}_h$ be the stratified sample means where $\bar{y}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} y_{hi}$, $\bar{x}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} x_{hi}$, $\bar{z}_h = \frac{1}{n_h}\sum_{i=1}^{n_h} z_{hi}$ are the stratum sample means and $\bar{Y}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} y_{hi}$, $\bar{X}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} x_{hi}$, $\bar{Z}_h = \frac{1}{N_h}\sum_{i=1}^{N_h} z_{hi}$ are corresponding population stratum means. Let $W_h = \frac{N_h}{N}$ $(h = 1, 2, ..., L)$ be the known stratum weights.

The population variance of the study variable in stratified sampling is given by Kadilar and Cingi (2006) [8] as

$$\sigma_{c0}^2 = \sum_{h=1}^{L} W_h \sigma_{yh}^2 + \sum_{h=1}^{L} W_h (\bar{Y}_h - \bar{Y})^2. \tag{1.1}$$

The combined ordinary and combined ratio estimators of population variance given by Kadilar and Cingi (2006) [8] in stratified sampling are given, respectively, by

$$t_{c0} = \sum_{h=1}^{L} W_h s_{yh}^2 + \sum_{h=1}^{L} W_h (\bar{y}_h - \bar{y}_{st})^2, \tag{1.2}$$

and

$$t_{c1} = t_{c0}\left(\frac{\sigma_x^2}{s_{xst}^2}\right), \text{ where } s_{xst}^2 = \sum_{h=1}^{L} W_h s_{xh}^2 + \sum_{h=1}^{L} W_h(\bar{x}_h - \bar{x}_{st})^2. \tag{1.3}$$

Some authors including Özel *et al.* (2014) [10] have suggested the separate ordinary and separate ratio estimators of population variance in stratified sampling which are given respectively by

$$t_{s0} = \sum_{h=1}^{L} W_h s_{yh}^2, \tag{1.4}$$

and

$$t_{s1} = \sum_{h=1}^{L} W_h \left(\frac{s_{yh}^2}{s_{xh}^2}\right)\sigma_{xh}^2. \tag{1.5}$$

In this paper, we have considered the problem of estimating population variance using auxiliary information by adapting Kadilar and Cingi (2006) [8], Özel *et al.* (2014) [10], and Gupta *et al.* (2020) [5] under RRT. We will discuss the proposed combined variance estimators in detail in Section 2. Separate variance estimators will be discussed in detail in Section 3. We also examine the effect of ignoring the term $\sum_{h=1}^{L} W_h(\bar{y}_h - \bar{y}_{st})^2$ in (1.2) on the estimates of the variance in stratified random sampling. Section 4 presents the results of a simulation study; Section 5 presents a real data example; and Section 6 provides some concluding remarks.

2. SOME COMBINED VARIANCE ESTIMATORS IN STRATIFIED RANDOM SAMPLING

In this study, the respondent is asked to provide a scrambled response for the sensitive study Y by using the generalized RRT model given by $Z = TY + S$, as in Diana and Perri (2011) [3], where S and T are uncorrelated scrambling variables such that $E(S) = 0$ and $E(T) = 1$. Gupta *et al.* (2020) [5] used this RRT model for estimating the population variance in simple random sampling. They proposed the following estimators:

$$t_0(R) = \frac{s_z^2 - \sigma_S^2 - \sigma_T^2 * \bar{z}^2}{\sigma_T^2 + 1}, \tag{2.1}$$

$$t_1(R) = t_0(R) * \left(\frac{\sigma_x^2}{s_x^2}\right), \tag{2.2}$$

and

$$t_p(R) = \left[t_0(R) + (\sigma_x^2 - s_x^2)\right] * \left(\frac{(\alpha\sigma_x^2 + \beta)}{\omega(\alpha s_x^2 + \beta) + (1-\omega)(\alpha\sigma_x^2 + \beta)}\right)^g, \tag{2.3}$$

where α and β are suitably chosen constants associated with the auxiliary variable X. With $g = 1$, one can obtain various ratio estimators, and with $g = -1$ one can obtain various product estimators. ω is an unknown whose optimal value will be used.

Motivated by Gupta *et al.* (2020) [5] and Kadilar and Cingi (2006) [8], we propose the following combined variance estimators in the stratified random sampling.

2.1. The Combined Basic Variance Estimator

Based on the RRT model $Z = TY + S$, we have σ^2_{zh} as

$$\sigma^2_{zh} = \sigma^2_{Th}(\sigma^2_{yh} + \mu^2_{yh}) + \sigma^2_{yh} + \sigma^2_{Sh}.$$

Rearranging, we get

$$\sigma^2_{yh} = \frac{\sigma^2_{zh} - \sigma^2_{Sh} - (\sigma^2_{Th} * \bar{Z}^2_h)}{\sigma^2_{Th} + 1}.$$

The population variance of the study variable in stratified sampling is given by

$$\sigma^2_{c0}(R) = \sum_{h=1}^{L} W_h \left(\frac{\sigma^2_{zh} - \sigma^2_{Sh} - \sigma^2_{Th} * \bar{Z}^2_h}{\sigma^2_{Th} + 1} \right) + \sum_{h=1}^{L} W_h (\bar{Z}_h - \bar{Z})^2. \tag{2.4}$$

Let

$$\sigma^2_{c0}(R) = A_1 + B_1,$$

where

$$A_1 = \sum_{h=1}^{L} W_h \left(\frac{\sigma^2_{zh} - \sigma^2_{Sh} - \sigma^2_{Th} * \bar{Z}^2_h}{\sigma^2_{Th} + 1} \right) \text{ and } B_1 = \sum_{h=1}^{L} W_h (\bar{Z}_h - \bar{Z})^2.$$

We have our first proposed combined estimator given by

$$t_{c0}(R) = \sum_{h=1}^{L} W_h \left(\frac{s^2_{zh} - \sigma^2_{Sh} - \sigma^2_{Th} * \bar{z}^2_h}{\sigma^2_{Th} + 1} \right) + \sum_{h=1}^{L} W_h (\bar{z}_h - \bar{z}_{st})^2. \tag{2.5}$$

Let

$$t_{c0}(R) = \hat{A}_1 + \hat{B}_1,$$

where

$$\hat{A}_1 = \sum_{h=1}^{L} W_h \left(\frac{s^2_{zh} - \sigma^2_{Sh} - \sigma^2_{Th} * \bar{z}^2_h}{\sigma^2_{Th} + 1} \right) \text{ and } \hat{B}_1 = \sum_{h=1}^{L} W_h (\bar{z}_h - \bar{z}_{st})^2.$$

To obtain the Bias and MSE expressions for the proposed estimators in the stratified random sampling, we define the following error terms

$$\delta_{zh} = \frac{s^2_{zh} - \sigma^2_{zh}}{\sigma^2_{zh}}, \quad e_{zh} = \frac{\bar{z_h} - \bar{Z}_h}{\bar{Z}_h}, \quad e_{zst} = \frac{\bar{z_{st}} - \bar{Z}}{\bar{Z}}, \quad e_{xst} = \frac{\bar{x_{st}} - \bar{X}}{\bar{X}},$$

such that

$$E(\delta_{zh}) = E(e_{zh}) = E(e_{zst}) = E(e_{xst}) = 0,$$

$$E(\delta^2_{zh}) = \theta_h(\lambda_{40h} - 1), \quad E(\delta^2_{xh}) = \theta_h(\lambda_{04h} - 1), \quad E(\delta_{zh}\delta_{xh}) = \theta_h(\lambda_{22h} - 1),$$

$$E(\delta_{zh}e_{zh}) = \theta_h \lambda_{30h} C_{zh}, \quad E(\delta_{xh}e_{zh}) = \theta_h \lambda_{12h} C_{zh}, \quad E(e^2_{zh}) = \theta_h C^2_{zh},$$

$$E(e_{zst}e_{zh}) = \sum_{h=1}^{L} W_h \theta_h \sigma^2_{zh}, \quad E(e^2_{zst}) = \frac{1}{\bar{Z}^2} \sum_{h=1}^{L} W^2_h \theta_h \sigma^2_{zh}, \quad E(e^2_{xst}) = \frac{1}{\bar{X}^2} \sum_{h=1}^{L} W^2_h \theta_h \sigma^2_{xh},$$

$$E(e_{xst}e_{zh}) = \sum_{h=1}^{L} W_h \theta_h \sigma_{zxh}, \quad E(e_{xst}e_{zst}) = \frac{1}{\bar{Z}\bar{X}} \sum_{h=1}^{L} W^2_h \theta_h \sigma_{zxh},$$

where

$$\sigma_{zxh} = \rho_{zxh}\sigma_{zh}\sigma_{xh}, \quad \rho_{zxh} = \frac{\rho_{yxh}}{\sqrt{1 + \frac{\sigma^2_{Th}(\sigma^2_{yh}+\mu^2_{yh})+\sigma^2_{Sh}}{\sigma^2_{yh}}}}, \quad \lambda_{rsh} = \frac{\mu_{rsh}}{\mu_{20h}^{\frac{r}{2}}\mu_{02h}^{\frac{s}{2}}},$$

$$\mu_{rsh} = \frac{1}{N_h - 1}\sum_{i=1}^{N_h}(Z_{hi} - \bar{Z}_h)^r(X_{hi} - \bar{X}_h)^s \quad \text{and} \quad C^2_{zh} = C^2_{yh}\sigma^2_{Th} + \left(\frac{\sigma^2_{Sh}}{\bar{Y_h}^2}\right).$$

Consider the first term

$$\hat{A}_1 = \sum_{h=1}^{L} W_h\left(\frac{s^2_{zh} - \sigma^2_{Sh} - \sigma^2_{Th} * \bar{z}^2_h}{\sigma^2_{Th} + 1}\right). \tag{2.6}$$

Rewriting (2.6), we have

$$\hat{A}_1 = \sum_{h=1}^{L} W_h\left(\frac{\sigma^2_{zh}(1 + \delta_{zh}) - \sigma^2_{Sh} - \sigma^2_{Th}[\bar{Z}_h(1 + e_{zh})]^2}{\sigma^2_{Th} + 1}\right).$$

Subtracting A_1 on both sides, we obtain

$$(\hat{A}_1 - A_1) = \sum_{h=1}^{L} W_h\left(\frac{\sigma^2_{zh}\delta_{zh} - 2\sigma^2_{Th}\bar{Z}^2_h e_{zh} - \sigma^2_{Th}\bar{Z}^2_h e^2_{zh}}{\sigma^2_{Th} + 1}\right). \tag{2.7}$$

Taking the expectation on both sides of (2.7), the Bias of $\hat{A}_1$ is obtained as

$$\text{Bias}(\hat{A}_1) \approx -\sum_{h=1}^{L} \theta_h W_h\left(\frac{\sigma^2_{Th}\bar{Z}^2_h C^2_{zh}}{\sigma^2_{Th} + 1}\right). \tag{2.8}$$

By squaring both sides of (2.7) and using the first order approximation, the MSE is obtained as

(2.9)

$$\text{MSE}(\hat{A}_1) \approx \sum_{h=1}^{L} \theta_h W^2_h\left(\frac{1}{(\sigma^2_{Th}+1)^2}\right)\left(\sigma^4_{zh}(\lambda_{40h} - 1) + 4\sigma^4_{Th}\bar{Z}^4_h C^2_{zh} - 4\sigma^2_{zh}\sigma^2_{Th}\bar{Z}^2_h\lambda_{30h}C_{zh}\right).$$

Consider the second term

$$\hat{B}_1 = \sum_{h=1}^{L} W_h(\bar{z}_h - \bar{z}_{st})^2. \tag{2.10}$$

Rewriting (2.10), we have

$$\hat{B}_1 = \sum_{h=1}^{L} W_h[\bar{Z}_h(1 + e_{zh}) - \bar{Z}(1 + e_{zst})]^2. \tag{2.11}$$

Expanding (2.11), and restricting to terms up to order 2, we have

$$\hat{B}_1 = \sum_{h=1}^{L} W_h[(\bar{Z}_h - \bar{Z})^2 + (\bar{Z}_h e_{zh} - \bar{Z}e_{zst})^2 + 2(\bar{Z}^2_h e_{zh} - \bar{Z}_h\bar{Z}e_{zst} - \bar{Z}_h\bar{Z}e_{zh} + \bar{Z}^2_h e_{zst})]. \tag{2.12}$$

Subtracting B_1 on both sides, we obtain

(2.13)

$$(\hat{B}_1 - B_1) = \bar{Z}^2 e^2_{zst} + \sum_{h=1}^{L} W_h[\bar{Z}^2_h e^2_{zh} - 2\bar{Z}_h\bar{Z}e_{zh}e_{zst} + 2(\bar{Z}^2_h e_{zh} - \bar{Z}_h\bar{Z}e_{zst} - \bar{Z}_h\bar{Z}e_{zh} + \bar{Z}^2_h e_{zst})].$$

Taking the expectation on both sides of (2.13), the Bias of $\hat{B}_1$ is obtained as

$$\text{Bias}(\hat{B}_1) \approx \bar{Z}^2 \sum_{h=1}^{L} W_h^2 \theta_h C_{zh}^2 + \sum_{h=1}^{L} W_h \theta_h [\bar{Z}_h^2 C_{zh}^2 - 2\bar{Z}_h \bar{Z} \sigma_{zh}^2]. \tag{2.14}$$

By squaring both sides of (2.13), using the first order approximation and simplifying, the MSE is obtained as

$$\begin{aligned} \text{MSE}(\hat{B}_1) \approx\ & 4\bar{Z}^4 \sum_{h=1}^{L} W_h^2 \theta_h C_{zh}^2 \\ & + \sum_{h=1}^{L} W_h^2 \theta_h \Bigg[4\bar{Z}_h^2 C_{zh}^2 - 8\bar{Z}_h \bar{Z} \sum_{h=1}^{L} W_h \sigma_{zh}^2 (\bar{Z}_h - \bar{Z})^2 + 4\bar{Z}_h \bar{Z}^2 \sum_{h=1}^{L} W_h^2 C_{zh}^2 (\bar{Z}_h - 2\bar{Z}) \Bigg]. \end{aligned} \tag{2.15}$$

The expressions for Bias and MSE of $t_{c0}(R)$ are given by

$$\text{Bias}(t_{c0}(R)) = \text{Bias}(\hat{A}_1) + \text{Bias}(\hat{B}_1), \tag{2.16}$$

and

$$\text{MSE}(t_{c0}(R)) \approx \text{MSE}(\hat{A}_1) + \text{MSE}(\hat{B}_1). \tag{2.17}$$

In (2.17), we assume that $\hat{A}_1$ and $\hat{B}_1$ are uncorrelated. This is not an unreasonable assumption since the sample mean and the sample variance are uncorrelated for normal data. This is also confirmed by large number of simulated values of $\hat{A}_1$ and $\hat{B}_1$ that we generated.

2.2. The Combined Ratio Variance Estimator

$$t_{c1}(R) = \sum_{h=1}^{L} W_h \left[\left(\frac{s_{zh}^2 - \sigma_{Sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1} \right) * \left(\frac{\sigma_{xh}^2}{s_{xh}^2} \right) \right] + \sum_{h=1}^{L} W_h \left(\bar{z}_h - \frac{\bar{z}_{st}}{\bar{x}_{st}} \bar{X} \right)^2, \tag{2.18}$$

$$t_{c1}(R) = \hat{A}_2 + \hat{B}_2.$$

Consider the first term:

$$\hat{A}_2 = \sum_{h=1}^{L} W_h \left[\left(\frac{s_{zh}^2 - \sigma_{Sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1} \right) * \left(\frac{\sigma_{xh}^2}{s_{xh}^2} \right) \right]. \tag{2.19}$$

Rewriting (2.19), we have

$$\hat{A}_2 = \sum_{h=1}^{L} W_h \left[\frac{\sigma_{zh}^2 - \sigma_{Sh}^2 - \sigma_{Th}^2 \bar{Z}_h^2}{\sigma_{Th}^2 + 1} + \frac{2\sigma_{Th}^2 \bar{Z}_h^2 e_{zh} \delta_{xh} - \sigma_{zh}^2 \delta_{zh} \delta_{xh} - \sigma_{Th}^2 \bar{Z}_h^2 e_{zh}^2}{\sigma_{Th}^2 + 1} \right].$$

Subtracting A_1 and taking the expectation on both sides, the Bias of $\hat{A}_2$ is obtained as

$$\text{Bias}(\hat{A}_2) \approx \sum_{h=1}^{L} \theta_h W_h \left[\left(\frac{2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh} - \sigma_{zh}^2 (\lambda_{22h} - 1) - \sigma_{Th}^2 \bar{Z}_h^2 C_{zh}^2}{\sigma_{Th}^2 + 1} \right) \right]. \tag{2.20}$$

For MSE, we have

$$\hat{A}_2 = \sum_{h=1}^{L} W_h \Big[\frac{\sigma_{zh}^2 + \sigma_{zh}^2\delta_{zh} - \sigma_{Sh}^2 - \sigma_{Th}^2\bar{Z}_h^2 - 2\sigma_{Th}^2\bar{Z}_h^2 e_{zh} - \sigma_{Th}^2\bar{Z}_h^2 e_{zh}^2}{\sigma_{Th}^2+1}$$
$$\frac{-\sigma_{zh}^2\delta_{xh} - \sigma_{zh}^2\delta_{zh}\delta_{xh} + \sigma_{Sh}^2\delta_{xh} + \sigma_{Th}^2\bar{Z}_h^2\delta_{xh} + 2\sigma_{Th}^2\bar{Z}_h^2 e_{zh}\delta_{xh} + \sigma_{Th}^2\bar{Z}_h^2 e_{zh}^2\delta_{xh}}{\sigma_{Th}^2+1}\Big].$$

Simplifying and ignoring second and higher order terms,

$$\hat{A}_2 = \sum_{h=1}^{L} W_h \Big[\frac{\sigma_{zh}^2 - \sigma_{Sh}^2 - \sigma_{Th}^2\bar{Z}_h^2}{\sigma_{Th}^2+1} + \frac{\sigma_{zh}^2\delta_{zh} - 2\sigma_{Th}^2\bar{Z}_h^2 e_{zh} - \sigma_{zh}^2\delta_{xh} + \sigma_{Sh}^2\delta_{xh} + \sigma_{Th}^2\bar{Z}_h^2\delta_{xh}}{\sigma_{Th}^2+1}\Big].$$

Squaring and taking the expectation on both sides, we have

$$\hat{A}_2 = \sum_{h=1}^{L} W_h^2 E\Big(\frac{\sigma_{zh}^2\delta_{zh}}{\sigma_{Th}^2+1} - \frac{2\sigma_{Th}^2\bar{Z}_h^2 e_{zh}}{\sigma_{Th}^2+1} - \sigma_{yh}^2\delta_{xh}\Big)^2.$$

After some simplifications, the MSE of $\hat{A}_2$ is obtained as

$$\text{MSE}(\hat{A}_2) \approx \sum_{h=1}^{L} \frac{W_h^2\theta_h}{(\sigma_{Th}^2+1)^2} \Big[\sigma_{zh}^4(\lambda_{40h}-1) - 2\sigma_{zh}^2\sigma_{yh}^2(\lambda_{22h}-1)(\sigma_{Th}^2+1) + \sigma_{yh}^4(\lambda_{04h}-1)(\sigma_{Th}^2+1)^2$$
$$+ 4C_{zh}\Big(\sigma_{Th}^4\bar{Z}_h^4 C_{zh} - \sigma_{zh}^2\sigma_{Th}^2\bar{Z}_h^2\lambda_{30h} + \sigma_{Th}^2\sigma_{yh}^2\bar{Z}_h^2\lambda_{12h}(\sigma_{Th}^2+1)\Big)\Big]. \tag{2.21}$$

Consider the second term:

$$\hat{B}_2 = \sum_{h=1}^{L} W_h \Big(\bar{z}_h - \frac{\bar{z}_{st}}{\bar{x}_{st}}\bar{X}\Big)^2. \tag{2.22}$$

Repeating the procedure outlined in steps (2.10)–(2.15) for the estimator (2.22), yields definitions of Bias and MSE for $\hat{B}_2$ as

$$\text{Bias}(\hat{B}_2) \approx \bar{Z}^2 \sum_{h=1}^{L} W_h^2\theta_h(C_{zh}^2 + C_{xh}^2) + \sum_{h=1}^{L} W_h\theta_h \Big[\bar{Z}_h^2 C_{zh}^2 - 2\bar{Z}_h\bar{Z}\sum_{h=1}^{L} W_h\sigma_{zh}^2$$
$$+ 2\Big(\frac{\bar{Z}_h}{\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh} + \frac{\bar{Z}_h}{\bar{X}}\sum_{h=1}^{L} W_h\sigma_{zxh} - 2\Big(\frac{\bar{Z}}{\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh}\Big)\Big)\Big], \tag{2.23}$$

$$\text{MSE}(\hat{B}_2) \approx 4\bar{Z}^4 \sum_{h=1}^{L} W_h^2\theta_h(C_{zh}^2 + C_{xh}^2) + \sum_{h=1}^{L} W_h^2\theta_h \Big[4\bar{Z}_h^2 C_{zh}^2(\bar{Z}_h - \bar{Z})^2$$
$$+ 4\bar{Z}_h^2\bar{Z}^2 \sum_{h=1}^{L} W_h^2(C_{zh}^2 + C_{xh}^2) + 8\bar{Z}_h^3\bar{Z}\sum_{h=1}^{L} W_h\Big(\frac{\sigma_z xh}{\bar{Z}\bar{X}} - \sigma_{zh}^2\Big)$$
$$- 8\bar{Z}_h^2\bar{Z}^2\Big(\frac{2}{\bar{Z}\bar{X}}\sum_{h=1}^{L} W_h\sigma_{zxh} - \sum_{h=1}^{L} 2W_h\sigma_{zh}^2 + \frac{1}{\bar{Z}\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh}\Big)$$
$$+ 8\bar{Z}_h\bar{Z}^3\Big(\frac{1}{\bar{Z}\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh} - \sum_{h=1}^{L} W_h^2 C_{zh}^2 + \frac{1}{\bar{Z}\bar{X}}\sum_{h=1}^{L} W_h\sigma_{zxh} - \sum_{h=1}^{L} W_h\sigma_{zh}^2\Big)$$
$$- 8\bar{Z}^2\frac{1}{\bar{Z}\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh}\Big]. \tag{2.24}$$

The expressions for Bias and MSE of $t_{c1}(R)$ are given by

$$\text{Bias}(t_{c1}(R)) = \text{Bias}(\hat{A}_2) + \text{Bias}(\hat{B}_2), \tag{2.25}$$

and

$$\text{MSE}(t_{c1}(R)) \approx \text{MSE}(\hat{A}_2) + \text{MSE}(\hat{B}_2). \tag{2.26}$$

2.3. A Combined Generalized Variance Estimator

We now propose the following class of generalized population variance estimators:

$$\begin{aligned} t_{cp}(R) &= \sum_{h=1}^{L} W_h \left[\left(\frac{s_{zh}^2 - \sigma_{Sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1} \right) + (\sigma_{xh}^2 - s_{xh}^2) \right] \\ &\quad * \left(\frac{(\alpha\sigma_{xh}^2 + \beta)}{\omega(\alpha s_{xh}^2 + \beta) + (1-\omega)(\alpha\sigma_{xh}^2 + \beta)} \right)^g \\ &\quad + \sum_{h=1}^{L} W_h \left[\left(\bar{z}_h - \left[\bar{z}_{st} + (\bar{X} - \bar{x}_{st}) \right] \right) * \left(\frac{(\alpha\bar{X} + \beta)}{\lambda(\alpha\bar{x}_{st} + \beta) + (1-\lambda)(\alpha\bar{X} + \beta)} \right)^g \right]^2, \end{aligned} \tag{2.27}$$

$$t_{cp}(R) = \hat{A}_3 + \hat{B}_3.$$

Consider the first term:

$$\hat{A}_3 = \sum_{h=1}^{L} W_h \left[\left(\frac{s_{zh}^2 - \sigma_{Sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1} \right) + (\sigma_{xh}^2 - s_{xh}^2) \right] * \left(\frac{(\alpha\sigma_{xh}^2 + \beta)}{\omega(\alpha s_{xh}^2 + \beta) + (1-\omega)(\alpha\sigma_{xh}^2 + \beta)} \right)^g.$$

Using Taylor series approximation, we obtain the Bias in $\hat{A}_3$ as

(2.28)

$$\text{Bias}(\hat{A}_3) = \sum_{h=1}^{L} -W_h\theta_h \left[\frac{\sigma_{Th}^2 \bar{Z}_h^2 C_{zh}^2}{\sigma_{Th}^2 + 1} - (g\omega\psi_h) \left(\frac{\sigma_{zh}^2(\lambda_{22h} - 1) - 2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh}}{\sigma_{Th}^2 + 1} - \sigma_{xh}^2(\lambda_{04h} - 1) \right) \right],$$

where $\psi_h = \sum_{h=1}^{L} \frac{\alpha\sigma_{xh}^2}{\alpha\sigma_{xh}^2 + \beta}$.

The mean square error is given by

$$\begin{aligned} \text{MSE}(\hat{A}_3) &= \sum_{h=1}^{L} W_h^2 \theta_h \left[\left(\frac{\sigma_{zh}^4(\lambda_{40h} - 1) + 4\sigma_{Th}^4 \bar{Z}_h^4 C_{zh}^2 - 4\sigma_{zh}^2 \sigma_{Th}^2 \bar{Z}_h^2 \lambda_{30h} C_{zh}}{(\sigma_{Th}^2 + 1)^2} \right) \right. \\ &\quad \left. + \left((\sigma_{xh}^2 + Q_h\sigma_{yh}^2)^2 (\lambda_{04h} - 1) \right) - 2 \left(\frac{\sigma_{zh}^2(\lambda_{22h} - 1) - 2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh}}{\sigma_{Th}^2 + 1} \right) (\sigma_{xh}^2 + Q_h\sigma_{yh}^2) \right], \end{aligned} \tag{2.29}$$

where $Q_h = g\omega\psi_h$.

Differentiate (2.29) w.r.t Q_h:

$$2\sigma_{yh}^2(\sigma_{xh}^2 + Q_h\sigma_{yh}^2)(\lambda_{04h} - 1) = 2\sigma_{yh}^2 \left(\frac{\sigma_{zh}^2(\lambda_{22h} - 1) - 2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh}}{\sigma_{Th}^2 + 1} \right),$$

$$Q_{h\text{opt}} = \sum_{h=1}^{L} \frac{1}{\sigma_{yh}^2} \left[\left(\frac{\sigma_{zh}^2(\lambda_{22h} - 1) - 2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh}}{\sigma_{Th}^2 + 1} \right) \left(\frac{1}{(\lambda_{04h} - 1)} \right) - \sigma_{xh}^2 \right].$$

The MSE at this optimum value is given by

$$\begin{aligned}\mathrm{MSE}(\hat{A}_3)_{\text{opt}} = \sum_{h=1}^{L}\frac{W_h^2\theta_h}{(\sigma_{Th}^2+1)^2}\Bigg[\Big(\sigma_{zh}^4(\lambda_{40h}-1)+4\sigma_{Th}^4\bar{Z}_h^4C_{zh}^2-4\sigma_{zh}^2\sigma_{Th}^2\bar{Z}_h^2\lambda_{30h}C_{zh}\Big)\\ -\frac{1}{(\lambda_{04h}-1)}\Big(\sigma_{zh}^2(\lambda_{22h}-1)-2\sigma_{Th}^2\bar{Z}_h^2\lambda_{12h}C_{zh}\Big)^2\Bigg]. \qquad (2.30)\end{aligned}$$

Consider the second term:

$$\hat{B}_3=\sum_{h=1}^{L}W_h\Bigg[\Big(\bar{z}_h-\big[\bar{z}_{st}+(\bar{X}-\bar{x}_{st})\big]\Big)*\Bigg(\frac{(\alpha\bar{X}+\beta)}{\lambda(\alpha\bar{x}_{st}+\beta)+(1-\lambda)(\alpha\bar{X}+\beta)}\Bigg)^g\Bigg]^2. \qquad (2.31)$$

Repeating the procedure outlined in steps (2.10)–(2.15) for the estimator (2.31), yields definitions of Bias and MSE for $\hat{B}_3$ as

$$\begin{aligned}\mathrm{Bias}(\hat{B}_3) \approx{}& \bar{Z}^2\sum_{h=1}^{L}W_h^2\theta_h\Big(C_{zh}^2+D^2C_{xh}^2\Big)+\bar{X}^2\sum_{h=1}^{L}W_h^2C_{xh}^2+\sum_{h=1}^{L}W_h\theta_h\Bigg[\bar{Z}_h^2C_{zh}^2\\ &+\frac{2\bar{Z}_h}{\bar{Z}}\sum_{h=1}^{L}W_h\sigma_{zxh}+2D\Bigg(\frac{\bar{Z}_h}{\bar{X}}\sum_{h=1}^{L}W_h^2\sigma_{zxh}-\bar{Z}_h\bar{X}\sum_{h=1}^{L}W_h^2C_{xh}^2+\frac{\bar{Z}_h}{\bar{X}}\sum_{h=1}^{L}W_h\sigma_{zxh}-\frac{2\bar{Z}}{\bar{X}}\sum_{h=1}^{L}W_h^2\sigma_{zxh}\\ &+\bar{Z}\bar{X}\sum_{h=1}^{L}W_h^2C_{xh}^2+\sum_{h=1}^{L}W_h^2\sigma_{zxh}\Bigg)-2\bar{Z}_h\bar{Z}\sum_{h=1}^{L}W_h\sigma_{zh}^2-2\sum_{h=1}^{L}W_h^2\sigma_{zxh}\Bigg], \qquad (2.32)\end{aligned}$$

where $D=(g\lambda\phi)$ and $\phi=\frac{\alpha\bar{X}}{\alpha\bar{X}+\beta}$;

$$\begin{aligned}\mathrm{MSE}(\hat{B}_3)_{\text{opt}} \approx{}& \theta\Bigg\{\bar{Z}^2\bar{X}^2\sum_{h=1}^{L}W_h^2C_{xh}^2+\bar{Z}^4\sum_{h=1}^{L}W_h^2C_{zh}^2-\bar{Z}^2\sum_{h=1}^{L}W_h^2\sigma_{zxh}\\ &+D_{\text{opt}}\Bigg[D_{\text{opt}}\bar{Z}^4\sum_{h=1}^{L}W_h^2C_{xh}^2+\bar{Z}^3\bar{X}\sum_{h=1}^{L}W_h^2C_{xh}^2-\frac{\bar{Z}^3}{\bar{X}}\sum_{h=1}^{L}W_h^2\sigma_{zxh}\Bigg]\\ &+\sum_{h=1}^{L}W_h^2\Bigg[4\Bigg(\big(\bar{Z}_h^4-\bar{Z}_h^3\bar{Z}\big)C_{zh}^2+\bar{Z}_h^2\bar{Z}^2\Big(\sum_{h=1}^{L}W_h^2C_{zh}^2+C_{zh}^2+\sum_{h=1}^{L}W_h\sigma_{zh}^2\Big)\Bigg)\\ &+\big(\bar{Z}_h^2\bar{X}^2-\bar{Z}_h\bar{Z}\bar{X}^2\big)\sum_{h=1}^{L}W_h^2C_{xh}^2-\big(\bar{Z}_h^3\bar{Z}^2+\bar{Z}_h\bar{Z}^3\big)\sum_{h=1}^{L}W_h\sigma_{zh}^2\\ &+\frac{\big(\bar{Z}_h^3\bar{X}+\bar{Z}_h\bar{Z}^2\bar{X}\big)}{\bar{Z}\bar{X}}\sum_{h=1}^{L}W_h\sigma_{zxh}-2\bar{Z}_h^2\sum_{h=1}^{L}W_h\sigma_{zxh}+\bar{Z}_h\bar{Z}\sum_{h=1}^{L}W_h^2\sigma_{zxh}\\ &+D_{\text{opt}}\Bigg(D_{\text{opt}}\bar{Z}_h^2\bar{Z}^2\sum_{h=1}^{L}W_h^2C_{xh}^2+\frac{\big(2\bar{Z}_h^3-2\bar{Z}_h^2\bar{Z}\big)}{\bar{X}}\sum_{h=1}^{L}W_h\sigma_{zxh}\\ &+\big(\bar{Z}_h^2\bar{X}\bar{Z}-2\bar{Z}_h\bar{X}\bar{Z}^2-\bar{Z}_h\bar{Z}^3\big)\sum_{h=1}^{L}W_h^2C_{xh}^2+\frac{\bar{Z}_h\bar{Z}^2}{\bar{X}}\sum_{h=1}^{L}W_h^2\sigma_{zxh}\Bigg)\Bigg]\Bigg\} \qquad (2.33)\end{aligned}$$

where

$$D_{\text{opt}} = \frac{-\sum_{h=1}^{L}\left[\frac{\left(2\bar{Z}_h^3-2\bar{Z}_h^2\bar{Z}\right)}{\bar{X}}\sum_{h=1}^{L} W_h\sigma_{zxh} + \bar{Z}^2\left(\bar{Z}_h\bar{X}-2\bar{X}\bar{Z}-\bar{Z}^2\right)\sum_{h=1}^{L} W_h^2C_{xh}^2 + \frac{\bar{Z}_h\bar{Z}^2}{\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh}\right]}{2\left\{\bar{Z}_h^4\sum_{h=1}^{L} W_h^2C_{xh}^2 + \sum_{h=1}^{L} W_h^2\left[\bar{Z}_h^2\bar{Z}^2\sum_{h=1}^{L} W_h^2C_{xh}^2\right] + \left[\bar{Z}^3\bar{X}\sum_{h=1}^{L} W_h^2C_{xh}^2 - \frac{\bar{Z}^4}{\bar{X}}\sum_{h=1}^{L} W_h^2\sigma_{zxh}\right]\right\}}.$$

The expressions for Bias and MSE of $t_{cp}(R)$ are given by

$$\text{Bias}(t_{cp}(R)) = \text{Bias}(\hat{A}_3) + \text{Bias}(\hat{B}_3), \tag{2.34}$$

and

$$\text{MSE}(t_{cp}(R))_{\text{opt}} \approx \text{MSE}(\hat{A}_3)_{\text{opt}} + \text{MSE}(\hat{B}_3)_{\text{opt}}. \tag{2.35}$$

3. SOME SEPARATE VARIANCE ESTIMATORS IN STRATIFIED RANDOM SAMPLING

Some authors, including Özel *et al.* (2014) [10], Clement (2018) [2] and Younis and Shabbir (2019) [16], have presented separate variance estimators. In doing so, they have ignored the B_1 term introduced in (1.2). We examine the following separate variance estimators in stratified random sampling mainly to show that ignoring the B_1 term can give misleadingly low MSE values.

3.1. The Separate Basic Variance Estimator

Following the authors listed above, the separate population variance of the study variable in stratified sampling is given by

$$\sigma_{s0}^2(R) = \sum_{h=1}^{L} W_h\left(\frac{\sigma_{zh}^2 - \sigma_{sh}^2 - \sigma_{Th}^2 * \bar{Z}_h^2}{\sigma_{Th}^2 + 1}\right). \tag{3.1}$$

This leads to the following estimator:

$$t_{s0}(R) = \sum_{h=1}^{L} W_h\left(\frac{s_{zh}^2 - \sigma_{sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1}\right). \tag{3.2}$$

The Bias and MSE of $t_{s0}(R)$ are given respectively as

$$\text{Bias}(t_{s0}(R)) \approx -\sum_{h=1}^{L}\theta_h W_h\left(\frac{\sigma_{Th}^2\bar{Z}_h^2C_{zh}^2}{\sigma_{Th}^2 + 1}\right), \tag{3.3}$$

and

(3.4)

$$\text{MSE}(t_{s0}(R)) \approx \sum_{h=1}^{L}\theta_h W_h^2\left(\frac{1}{(\sigma_{Th}^2+1)^2}\right)\left(\sigma_{zh}^4(\lambda_{40h}-1) + 4\sigma_{Th}^4\bar{Z}_h^4C_{zh}^2 - 4\sigma_{zh}^2\sigma_{Th}^2\bar{Z}_h^2\lambda_{30h}C_{zh}\right).$$

3.2. The Separate Ratio Variance Estimator

$$t_{s1}(R) = \sum_{h=1}^{L} W_h \left[\left(\frac{s_{zh}^2 - \sigma_{sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1} \right) * \left(\frac{\sigma_{xh}^2}{s_{xh}^2} \right) \right]. \tag{3.5}$$

The Bias and MSE of $t_{s1}(R)$ are given respectively as

$$\text{Bias}(t_{s1}(R)) \approx \sum_{h=1}^{L} \theta_h W_h \left[\left(\frac{2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh} - \sigma_{zh}^2(\lambda_{22h} - 1) - \sigma_{Th}^2 \bar{Z}_h^2 C_{zh}^2}{\sigma_{Th}^2 + 1} \right) \right], \tag{3.6}$$

and

$$\begin{aligned} \text{MSE}(t_{s1}(R)) \approx \sum_{h=1}^{L} \frac{W_h^2 \theta_h}{(\sigma_{Th}^2 + 1)^2} \Big[& \sigma_{zh}^4(\lambda_{40h} - 1) - 2\sigma_{zh}^2 \sigma_{yh}^2 (\lambda_{22h} - 1)(\sigma_{Th}^2 + 1) \\ & + \sigma_{yh}^4(\lambda_{04h} - 1)(\sigma_{Th}^2 + 1)^2 \\ & + 4C_{zh} \Big(\sigma_{Th}^4 \bar{Z}_h^4 C_{zh} - \sigma_{zh}^2 \sigma_{Th}^2 \bar{Z}_h^2 \lambda_{30h} + \sigma_{Th}^2 \sigma_{yh}^2 \bar{Z}_h^2 \lambda_{12h} (\sigma_{Th}^2 + 1) \Big) \Big] \end{aligned} \tag{3.7}$$

3.3. A Separate Generalized Variance Estimator

The generalized separate population variance estimators can be written as

$$\begin{aligned} t_{sp}(R) = & \sum_{h=1}^{L} W_h \left[\left(\frac{s_{zh}^2 - \sigma_{sh}^2 - \sigma_{Th}^2 * \bar{z}_h^2}{\sigma_{Th}^2 + 1} \right) + (\sigma_{xh}^2 - s_{xh}^2) \right] \\ & * \left(\frac{(\alpha \sigma_{xh}^2 + \beta)}{\omega(\alpha s_{xh}^2 + \beta) + (1 - \omega)(\alpha \sigma_{xh}^2 + \beta)} \right)^g . \end{aligned} \tag{3.8}$$

The Bias and MSE of $t_{sp}(R)$ are given respectively as

$$\begin{aligned} \text{Bias}(t_{sp}(R)) = & \sum_{h=1}^{L} - W_h \theta_h \left[\frac{\sigma_{Th}^2 \bar{Z}_h^2 C_{zh}^2}{\sigma_{Th}^2 + 1} - (g \omega \psi_h) \right. \\ & \left. * \left(\frac{\sigma_{zh}^2(\lambda_{22h} - 1) - 2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh}}{\sigma_{Th}^2 + 1} - \sigma_{xh}^2(\lambda_{04h} - 1) \right) \right], \end{aligned} \tag{3.9}$$

and

$$\begin{aligned} \text{MSE}(t_{sp}(R))_{\text{opt}} = & \sum_{h=1}^{L} \frac{W_h^2 \theta_h}{(\sigma_{Th}^2 + 1)^2} \left[\left(\sigma_{zh}^4(\lambda_{40h} - 1) + 4\sigma_{Th}^4 \bar{Z}_h^4 C_{zh}^2 - 4\sigma_{zh}^2 \sigma_{Th}^2 \bar{Z}_h^2 \lambda_{30h} C_{zh} \right) \right. \\ & \left. - \frac{1}{(\lambda_{04h} - 1)} \left(\sigma_{zh}^2(\lambda_{22h} - 1) - 2\sigma_{Th}^2 \bar{Z}_h^2 \lambda_{12h} C_{zh} \right)^2 \right]. \end{aligned} \tag{3.10}$$

4. SIMULATION STUDY

We consider a sample of size N = 2000 from two bivariate normal populations for $\begin{bmatrix} X \\ Y \end{bmatrix}$ determined by the following means and covariance matrices with $N_1 = 1200$ and $N_2 = 800$:

$$\text{Stratum 1:} \quad \mu = \begin{bmatrix} 4 \\ 2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2 & 2.7 \\ 2.7 & 6 \end{bmatrix}, \quad \rho_{yx} = 0.80,$$

$$\text{Stratum 2:} \quad \mu = \begin{bmatrix} 6 \\ 4 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 2 & 2.2 \\ 2.2 & 5 \end{bmatrix}, \quad \rho_{yx} = 0.70.$$

These 2000 observations are treated as our finite populations. For the 2000 values generated from these distributions, the means, variances, and correlations are given by:

$$\text{Stratum 1:} \quad \mu_{x1} = 4.021, \ \mu_{y1} = 2.010, \ \sigma^2_{x1} = 1.975, \ \sigma^2_{y1} = 5.987, \ \rho_{yx1} = 0.797,$$

$$\text{Stratum 2:} \quad \mu_{x2} = 6.070, \ \mu_{y2} = 4.006, \ \sigma^2_{x2} = 1.982, \ \sigma^2_{y2} = 4.977, \ \rho_{yx2} = 0.702.$$

Overall parameter values are given by

$$\mu_x = 4.8413, \quad \mu_y = 2.8791, \quad \sigma^2_x = 2.9671, \quad \sigma^2_y = 6.4644, \quad \rho_{yx} = 0.7596.$$

We consider a sample of size $n = 600$, where $n_1 = 360$ and $n_2 = 240$. The stratum sample size n_h $(h = 1, 2)$ is based on the proportional allocation, that is, $n_h = W_h \times n$. The scrambling variable S and T are assumed to have normal distributions with $E(S) = 0$, $E(T) = 1$, $\text{Var}(S) = 0.5$ and different values for $\text{Var}(T)$. In the combined and separate generalized variance estimators, we choose $\alpha = 1$, $\beta = 0$ and $g = 1$. Other choices of α and β in our simulations had minimal impact.

The Percent Relative Efficiency (PRE) with respect to the stratified sampling is defined as

$$\text{PRE} = \frac{\text{MSE}(t_{c0}(R))}{\text{MSE}(t_{ci}(R))} \times 100, \ \text{where } i = 0, 1 \text{ and } p.$$

Since we are developing the proposed estimators based on randomized data, it is important to consider the privacy level as well. Gupta *et al.* (2018) [6] introduced a unified measure of estimator quality (δ) given by

$$\delta = \frac{\text{Theoretical MSE}}{\Delta_{DP}}, \quad \text{where } \Delta_{DP} = \sum_{h=1}^{L} W_h \Delta_{DPh}$$

is the privacy level for the model $Z{=}TY{+}S$ as given by Yan *et al.* (2009) [15].

Theoretical and empirical MSEs and PREs for both the separate variance estimators and combined variance estimators are reported in Table 1. For either separate or combined estimators, the generalized estimator is clearly more efficient than the basic estimator and the ratio estimator. One can note that the MSEs increase as the variances of T increase, which is on expected lines due to extra noise in the data. However, this loss in efficiency is off-set by the gain in privacy as shown by the δ-column. For example, the MSEs of the combined generalized variance estimator $t_{cp}(R)$ increases from 0.2842 to 0.5024 when $\text{Var}(T)$ increases from 0.5 to 1, but δ value decreases from 0.0364 to 0.0323. In general, the proposed variance estimators under the additive model ($Z = Y + S$) where $\text{Var}(T) = 0$ are more efficient compared to the generalized model ($Z = TY + S$) where $\text{Var}(T) > 0$ by providing smaller MSEs. However, the proposed variance estimators under the generalized model ($Z = TY + S$) are better by providing smaller δ values if we consider the efficiency and the privacy simultaneously.

Table 1: Theoretical (in **bold**) and empirical MSEs and PREs of the variance estimators with $\sigma_y^2 = 6.4644$.

Var(S)	**Var(T)**	**Estimator**	$\hat{\sigma_y^2}$	**MSE**	**PRE**	$\boldsymbol{\delta}$	**Estimator**	$\hat{\sigma_y^2}$	**MSE**	**PRE**	$\boldsymbol{\delta}$
0.5	0	$t_{s0}(R)$	5.4368	**0.0857** 0.0858	**100** 100	0.1664	$t_{c0}(R)$	6.4141	**0.1119** 0.1131	**100** 100	0.2144
		$t_{s1}(R)$	5.4542	**0.0752** 0.0765	**113.9627** 112.1568	0.1441	$t_{c1}(R)$	6.4322	**0.0985** 0.1014	**113.6040** 111.5384	0.1887
		$t_{sp}(R)$	5.4402	**0.0635** 0.0647	**134.9606** 132.6120	0.1368	$t_{cp}(R)$	6.4182	**0.0858** 0.0863	**130.4195** 131.0544	0.1644
	0.3	$t_{s0}(R)$	5.5456	**0.2142** 0.2151	**100** 100	0.0442	$t_{c0}(R)$	6.4434	**0.2588** 0.2606	**100** 100	0.0530
		$t_{s1}(R)$	5.5547	**0.1951** 0.1964	**109.8923** 109.5213	0.0400	$t_{c1}(R)$	6.4534	**0.2376** 0.2399	**108.9225** 108.6285	0.0486
		$t_{sp}(R)$	5.5443	**0.1757** 0.1769	**122.0261** 121.5941	0.0374	$t_{cp}(R)$	6.4429	**0.2169** 0.2174	**119.3176** 119.8712	0.0444
	0.5	$t_{s0}(R)$	5.5357	**0.2788** 0.2781	**100** 100	0.0357	$t_{c0}(R)$	6.4141	**0.3361** 0.3370	**100** 100	0.0431
		$t_{s1}(R)$	5.5446	**0.2579** 0.2596	**108.1039** 107.1263	0.0330	$t_{c1}(R)$	6.4239	**0.3147** 0.3158	**106.8001** 106.7131	0.0404
		$t_{sp}(R)$	5.5344	**0.2351** 0.2491	**118.5878** 111.6419	0.0306	$t_{cp}(R)$	6.4137	**0.2842** 0.2922	**118.2617** 114.9897	0.0364
	0.8	$t_{s0}(R)$	5.5128	**0.3801** 0.3766	**100** 100	0.0310	$t_{c0}(R)$	6.3931	**0.4560** 0.4528	**100** 100	0.0375
		$t_{s1}(R)$	5.5216	**0.3621** 0.3589	**104.9710** 104.9317	0.0294	$t_{c1}(R)$	6.4041	**0.4355** 0.4331	**104.7072** 104.5486	0.0358
		$t_{sp}(R)$	5.5115	**0.3406** 0.3399	**111.5971** 110.7972	0.0279	$t_{cp}(R)$	6.3981	**0.4047** 0.4141	**112.6760** 109.3455	0.0333
	1	$t_{s0}(R)$	5.4970	**0.4550** 0.4512	**100** 100	0.0299	$t_{c0}(R)$	6.3897	**0.5431** 0.5382	**100** 100	0.0360
		$t_{s1}(R)$	5.5057	**0.4399** 0.4372	**103.4325** 103.2021	0.0287	$t_{c1}(R)$	6.3929	**0.5235** 0.5202	**103.7440** 103.4602	0.0347
		$t_{sp}(R)$	5.4957	**0.4207** 0.4169	**108.1530** 108.2273	0.0262	$t_{cp}(R)$	6.3928	**0.5024** 0.4950	**108.1011** 108.7272	0.0323

Comparing the proposed separate variance estimators to the proposed combined variance estimators, it may appear that the separate estimators are better since they have smaller MSE. However, one should note that these estimators degrade accuracy in comparison to the combined estimators. For example, as the true variance of Y is 6.4644, the estimated variance of Y is 6.4137 for the combined variance estimator when $\text{Var}(T) = 0.5$. However, the estimated variance of Y is 5.5344 for the separate variance estimator when $\text{Var}(T) = 0.5$. The same is true for the other cases. This indicates that the proposed combined variance estimators are more accurate than the separate variance estimators.

5. APPLICATION

In this section a real data set is used to compare the performances of the combined variance estimators. The data is obtained from Eurostat (2008) [4], and the sampling details are provided in Sousa *et al.* (2014) [13], a paper that was co-authored by one of the co-authors of the current paper. There are 1698 records in the population. The volume of purchase orders reported by the Information and Communication Technologies for 2010 is taken as the study variable Y. Turnover for the individual enterprises is the auxiliary variable X. The study variable Y is scrambled using the additive scrambling variable S assumed to be a normally distributed random variable with mean 0 and variance 0.5, and the multiplicative scrambling variable T assumed to be a normally distributed random variable with mean 1 and four different choices for its variance (0, 0.3, 0.5, and 1). Data summary is provided in Table 2.

Table 2: Population Characteristics and Sampling Information.

Stratum	N	ρ_{yx}	μ_y	σ_y	μ_y	σ_x	Population
1	979	0.7802	2.15	2.46	3.12	2.68	$N = 1698,\ \rho_y x = 0.9368$
2	362	0.7952	16.67	6.86	20.31	6.02	$\mu_y = 14.44,\ \sigma_y^2 = 501.31$
3	357	0.8408	45.88	30.21	56.33	30.18	$\mu_x = 17.97,\ \sigma_x^2 = 640.59$

Table 3: Theoretical (in **bold**) and empirical MSEs and PREs of the variance estimators.

Var(S)	Var(T)	Estimator	$\hat{\sigma_y^2}$	MSE	PRE	δ
0.5	0	$t_{c0}(R)$	502.1499	**1230.678** 1228.833	**100** 100	2485.8154
		$t_{c1}(R)$	502.6576	**1079.764** 1086.812	**113.9765** 113.0676	2180.9880
		$t_{cp}(R)$	501.8341	**955.227** 947.223	**128.8361** 129.7300	1929.4389
	0.3	$t_{c0}(R)$	500.7685	**3615.078** 3656.872	**100** 100	15.8680
		$t_{c1}(R)$	501.557	**3483.887** 3457.51	**103.7656** 105.7660	15.2921
		$t_{cp}(R)$	503.1153	**3353.681** 3337.229	**111.1073** 112.9630	14.2817
	0.5	$t_{c0}(R)$	503.8318	**5389.539** 5353.592	**100** 100	14.2064
		$t_{c1}(R)$	501.3672	**5092.119** 5004.276	**105.8407** 106.9803	13.4224
		$t_{cp}(R)$	501.8327	**4878.576** 4841.631	**110.4736** 110.5741	12.8596
	1	$t_{c0}(R)$	501.3862	**9792.979** 9719.734	**100** 100	12.9152
		$t_{c1}(R)$	503.1442	**9540.721** 9574.217	**102.6440** 101.5198	12.5825
		$t_{cp}(R)$	500.9520	**9240.276** 9209.5	**105.9814** 105.5403	12.1863

Theoretical and empirical MSEs and PREs are provided in Table 3 for each of the proposed combined estimators. We used only the combined estimators in this numerical application because of the inherent drawback in the separate estimators as pointed out at the beginning of Section 3. The combined generalized variance estimator is clearly more efficient than both the combined basic variance estimator and the combined ratio variance estimator. Furthermore, the MSE increases as the variance of T is increased, meanwhile the unified measure (δ) value decreases. For example, for the combined generalized variance estimator, theoretical MSE is 3353.681 for $\sigma_T^2 = 0.3$ but increases to 9240.276 for $\sigma_T^2 = 1$. In contrast, the (δ) value decreases from 14.2817 to 12.1863 indicating that using the multiplicative noise T lowers the efficiency but the added privacy because of this more than compensates this loss.

6. CONCLUSION

Separate and combined variance estimators are considered under RRT in stratified random sampling. The simulation study shows that the generalized variance estimator is more efficient than the other estimators. Also, the proposed combined variance estimators are more accurate than the separate variance estimators. Furthermore, if one considers efficiency and privacy simultaneously, the linear combination model $Z = TY + S$, where $\text{Var}(T) > 0$, produces better variance estimators compared to the additive model $Z = Y + S$ where $\text{Var}(T) = 0$. This can be attributed to the fact that proposed variance estimators under $Z = TY + S$ have higher privacy level and hence smaller δ values. The real data application in Section 5 shows the same improvement with the generalized estimator as was seen in the simulation results of Section 4. We would like to mention that this work can be extended in several directions in new studies. For example, one can work with the case when the mean of the auxiliary variable is unknown. Also, the generalized estimator we suggest is not the only option. One can use other forms of generalizations.

ACKNOWLEDGMENTS

The authors appreciate the helpful comments by the two reviewers which helped improve the presentation of the paper.

REFERENCES

[1] Asghar, A.; Sanaullah, A. and Hanif, M. (2019). Generalized class of variance estimators under two-phase sampling for partial information case, *Electronic Journal of Applied Statistical Analysis*, **12**(1), 44–54.

[2] Clement, E.P. (2018). Improved Family of Ratio Estimators of Finite Population Variance in Stratified Random Sampling, *Biostat. Biometrics Open Acc. J.*, **5**(2), 555–659.

[3] Diana, G. and Perri, P.F. (2011). A class of estimators for quantitative sensitive data, *Statistical Papers*, **52**(3), 633–650.

[4] Eurostat (2008). NACE Rev.2 — statistical classification of economic activities in the european community, *Official Publications of the European Communities*, 112–285 and 306–311.

[5] Gupta, S.; Aloraini, B.; Qureshi, M. and Khalil, S. (2020). Variance Estimation using Randomized Response Technique, *REVSTAT*, **18**(2), 165–176.

[6] Gupta, S.; Samridhi, M.; Shabbir, J. and Khalil, S. (2018). A class of estimators for quantitative sensitive data, *Statistical Papers*, **52**(3), 633–650.

[7] Isaki, C.T. (1983). Variance estimation using auxiliary information, *Journal of the American Statistical Association*, **78**(381), 117–123.

[8] Kadilar, C. and Cingi, H. (2006). Ratio estimators for the population variance in simple and stratified random sampling, *Applied Mathematics and Computation*, **173**(2), 1047–1059.

[9] Kalucha, G.; Sousa, R.; Gupta, S. and Shabbir, J. (2017). Improved Ratio and Regression Estimators of the Mean of a Sensitive Variable in Stratified Sampling, *Statistics and Applications*, **15**(1-2), 63–78.

[10] Özel, G.; Çingi, H. and Oğuz, M. (2014). Separate Ratio Estimators for the Population Variance in Stratified Random Sampling, *Communications in Statistics – Theory and Methods*, **43**(22), 4766–4779. DOI: 10.1080/03610926.2012.729642

[11] Sanaullah, A.; Asghar, A. and Hanif, M. (2017). General class of exponential estimator for estimating finite population variance, *Journal of Reliability and Statistical Studies*, 1–16.

[12] Sidelel, E.B.; Orwa, G.O. and Otieno, R.O. (2014). Variance estimation in stratified random sampling in the presence of two auxiliary random variables, *International Journal of Science and Research*, **3**(9), 2453–2459.

[13] Sousa, R.; Gupta, S.; Shabbir, J. and Corte-Real, P. (2014). Improved mean estimation of a sensitive variable using auxiliary information in stratified sampling, *Journal of Statistics and Management Systems*, **17**(5-6), 503–518.

[14] Warner, S.L. (1965). Randomized response: A survey technique for eliminating evasive answer bias, *Journal of the American Statistical Association*, **60**(309), 63–69.

[15] Yan, Z.; Wang, J. and Lai, J. (2009). An efficiency and protection degree-based comparison among the quantitative randomized response strategies, *Communications in Statistics – Theory and Methods*, **38**, 404–408.

[16] Younis, F. and Shabbir, J. (2019). Estimation of General Parameter Under Stratified Adaptive Cluster Sampling Based on Dual Use of Auxiliary Information, *Scientia Iranica*. DOI: 10.24200/sci.2019.52515.2753

[17] Zahid, E. and Shabbir, J. (2018). Estimation of population mean in the presence of measurement error and non response under stratified random sampling, *PloS one*, **13**(2), e0191572.

[18] Zhang, Q.; Khalil, S. and Gupta, S. (2021). Estimation in the Simultaneous Presence of Measurement Errors and Non-response using Optional RRT Models under Stratified Sampling, *Journal of Statistical Computation and Simulation* (to appear).

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 293–308
https://doi.org/10.57805/revstat.v22i3.504

The Marshall and Olkin-G and Gamma-G Family of Distributions: Properties and Applications

Authors: Maria do Carmo S. Lima
– Department of Statistics, Federal University of Pernambuco, Pernambuco, Recife, Brazil
maria@de.ufpe.br

Gauss M. Cordeiro
– Department of Statistics, Federal University of Pernambuco, Pernambuco, Recife, Brazil
gauss@de.ufpe.br

Pedro Rafael D. Marinho ✉
– Department of Statistics, Federal University of Paraíba, Paraíba, João Pessoa, Brazil
pedro.rafael.marinho@gmail.com

Received: December 2020 Revised: June 2022 Accepted: June 2022

Abstract:

- The article introduces a new family by combining the Marshall and Olkin-G and Gamma-G classes. It has only two extra shape parameters and can be a better model than other existing classes of distributions. Simulations are performed to verify the consistency of the estimators. Its flexibility is shown by means of two real data sets.

Keywords:

- *applications; distribution family; mathematical properties; simulations.*

AMS Subject Classification:

- 60E05, 62E10, 62E15.

✉ Corresponding author

1. INTRODUCTION

The advances in the field of Data Science requires the search for new families of distributions that adequately model real data has been increasing steadily in the last years. The construction of different generation methods and even generators of families has made it difficult to compare new proposals. In the midst of a huge set of existing families in the literature of new distributions, to find a proposal that is in fact an excellent competitor when compared to other existing ones, in terms of adjustment to real data sets and also that does not present estimation problems, is a major challenge.

The classes of distributions in the early 1980s were based on the simple idea of adding parameters to a baseline distribution. The mechanism by adding shape parameters to a baseline distribution has proved to be useful to make the generated distributions more flexible especially for studying tail properties than existing ones and for improving their goodness-of-fit statistics to real data. Many special distributions in these families are discussed by Tahir and Nadarajah (2015) [1].

The addition of parameters in the construction of new distributions/families was improved by the inclusion of mathematical functions known in the literature, such as beta and gamma functions, for example, which produce new generators with more flexible properties than their baselines. Two well-known examples are the beta-G (Eugene *et al.*, 2002) [2] and gamma-G (Zografos and Balakrishnan, 2009) [3] generators.

However, the inclusion of such functions for generating new families brought, in some cases, problems for parameter estimation. So, despite the fit being more suitable for some types of data and, therefore, having a superior performance when compared to other generators, the estimation process can often be a problem.

In this context, this work presents a new family obtained by composing a very competitive class in the literature with another class that has the gamma function in its structure.

Let $G(x)$ be the cumulative distribution function (CDF) of a baseline distribution and $g(x) = dG(x)/dx$ be the corresponding probability density function (PDF) depending on a parameter vector η. A generalized family is presented with two extra shape parameters by transforming the CDF $G(x)$ according to two sequential important classes. These classes, called Marshall and Olkin-G and Gamma-G, are important for modeling data in several areas, and they are reviewed below.

The CDF of the Marshall and Olkin's (1997) [4] (MO-G) class (for $\theta > 0$) is

$$F_{\text{MO-G}}(x) = \frac{G(x)}{\theta + (1-\theta)G(x)} = \frac{G(x)}{1-(1-\theta)[1-G(x)]}, \quad x \in \mathbb{R}. \tag{1.1}$$

The density function corresponding to (1.1) has the form

$$f_{\text{MO-G}}(x) = \frac{\theta\, g(x)}{[\theta + (1-\theta)G(x)]^2}. \tag{1.2}$$

For $\theta = 1$, $f_{\text{MO-G}}(x)$ is equal to $g(x)$. Equation (1.2) represents the PDF of the minimum of n iid random variables having density $g(x)$, say $T_1, ..., T_N$, where N has a geometric distribution with probability parameters θ and θ^{-1} if $0 < \theta < 1$ and $\theta > 1$, respectively.

Tahir and Nadarajah (2015, Table 2) [1] presented thirty distributions belonging to this family. It is easily generated from the baseline quantile function (QF) by $Q_{\text{MO-G}}(u) = Q_G(\theta u/[\theta u + 1 - u])$ for $u \in (0, 1)$.

Marshall and Olkin considered the exponential and Weibull distributions for the baseline G and derived some structural properties of the generated distributions. If G is an exponential distribution, the special case refers to a two-parameter competitive model to the Weibull and gamma distributions.

The CDF of the gamma-G (Γ-G) class (Zografos and Balakrishnan, 2009) [3] is

$$F_{\Gamma\text{-G}}(x) = \gamma_1(a, -\log[1 - G(x)]), \quad x \in \mathbb{R}, \tag{1.3}$$

where $a > 0$ is an extra shape parameter, $\gamma_1(a, z) = \gamma(a, z)/\Gamma(a)$ is the incomplete gamma function ratio, and $\gamma(a, z) = \int_0^z t^{a-1}\, \mathrm{e}^{-\mathrm{t}} \mathrm{dt}$.

Then, the PDF of the Γ-G class can be expressed as

$$f_{\Gamma\text{-G}}(x) = \frac{1}{\Gamma(a)} \{-\log[1 - G(x)]\}^{a-1}\, g(x). \tag{1.4}$$

Each new Γ-G distribution follows from a given baseline G. For $a = 1$, the Γ-G class reduces to G. If Z is a gamma random variable with unit scale and shape $a > 0$, then $W = Q_G(1 - \mathrm{e}^{-\mathrm{Z}})$ has density (1.4). So, the Γ-G distribution is easily generated from the gamma distribution and the QF of G.

The remaining of the paper is addressed as follows. Section 2 introduces the *Marshall and Olkin-Gamma-G* (MOGa-G) family, and provides some special models. The maximum likelihood estimates (MLEs) of its parameters is addressed in Section 3. Some simulations are performed in Section 4 to estimate the biases of the MLEs. Two empirical applications illustrate the potentiality of the proposed family in Section 5. A variety of theoretical properties are obtained in Section 6. Some conclusions remarks are offered in Section 7.

2. THE NEW FAMILY

By combining Equations (1.1) and (1.3), the CDF of the random variable $X \sim$MOGa-G representing the new family has the form

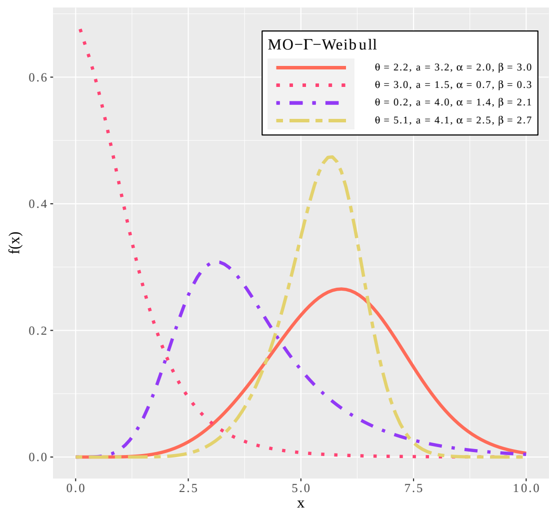
$$F_X(x) = \frac{\gamma_1(a, -\log[1 - G(x)])}{\theta + (1-\theta)\gamma_1(a, -\log[1 - G(x)])}, \quad x \in \mathbb{R}. \tag{2.1}$$

By differentiating (2.1), the PDF of X follows as

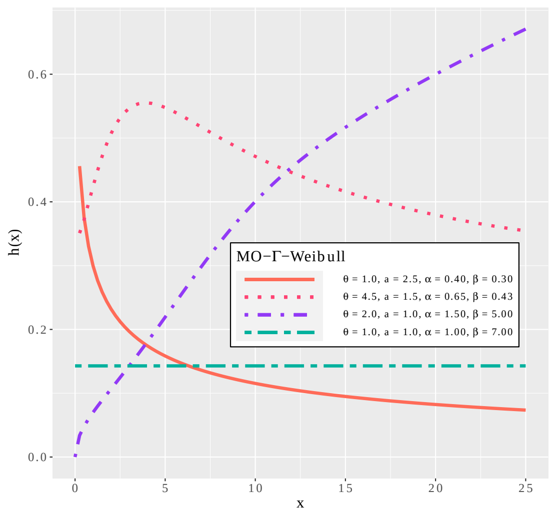
$$f_X(x) = \frac{\theta\{-\log[1 - G(x)]\}^{a-1}\, g(x)}{\Gamma(a)\, \{\theta + (1-\theta)\gamma_1(a, -\log[1 - G(x)])\}^2}. \tag{2.2}$$

The density (2.2) can be interpreted from a sequence of N iid random variables, say $Z_1, ..., Z_N$, each one having a gamma density with unit scale and shape $a > 0$, assuming that N (is not fixed) has a geometric distribution with probabilities θ and θ^{-1} for $0 < \theta < 1$ and $\theta > 1$, respectively. By transforming the Z_i's via the baseline QF by $W_i = Q_G(1 - \mathrm{e}^{-\mathrm{Z_i}})$ (for $i = 1, ..., N$), Equation (1.2) defines the PDF of the minimum $W_1, ..., W_n$. The proposed family from the double composition of the two classes absorbs the impacts of their different flexibilities on real applications.

Table 1 provides some special cases of (2.2), where $\Phi(x)$ and $\phi(x)$ are the CDF and PDF of the standard normal distribution. The density and hazard ($h(x) = f(x)/[1 - F(x)]$) functions of the MOGa-Weibull (MOGa-W) model are displayed in Figure 1, which provide more flexibility to both functions for both classes applied separately to the Weibull model.



(a) Density function.



(b) Hazard function.

Figure 1: The density and hazard functions of the MOGa-W model.

The CDF (2.1) can be easily inverted to calculate the QF of the MOGa-G distribution, say $x = Q_X(u) = F_X^{-1}(u)$ (for $u \in (0, 1)$), in terms of the baseline QF $Q_G(\cdot)$. The inverse of $F_X(x) = u$, where u is a uniform number in $(0, 1)$, follows by combining the inverses of Equations (1.1) and (2.1). So, $F_X(x) = u$ gives $z = z(u) = \theta u/[1 - (1 - \theta)u]$ and $\gamma_1(a, -\log[1 - G(x)]) = z(u)$. Hence, the QF of X can be expressed as

$$x = Q_G(v(u)),$$

where

$$v(u) = 1 - \exp\left[-\gamma_1^{-1}\left(a, z(u)\right)\right],$$

and $\gamma_1^{-1}(a,w) = Q^{-1}(a, 1-w)$ is the inverse function of $\gamma_1(a,w)$. Some formulae for $Q^{-1}(a, 1-w)$ are given in `http://functions.wolfram.com/GammaBetaErf/InverseGammaRegularized/`.

Table 1: Special Distributions in the MOGa-G family.

Distribution	**Baseline CDF**	**Generated PDF**
Normal	$G(x) = \Phi(x)$	$f_X(x) = \frac{\theta\{-\log[1-\Phi(x)]\}^{a-1}\,\phi(x)}{\Gamma(a)\,\{\theta+(1-\theta)\gamma_1(a,-\log[1-\Phi(x)])\}^2}$
Logistic	$G(x) = \frac{1}{1+\mathrm{e}^{-\mathrm{x}}}$	$f_X(x) = \frac{\theta\,\mathrm{e}^{-\mathrm{x}}\,\left\{-\log[1-(1+\mathrm{e}^{-\mathrm{x}})^{-1}]\right\}^{\mathrm{a}-1}}{\Gamma(a)\,(1+\mathrm{e}^{-\mathrm{x}})^2\,\left\{\theta+(1-\theta)\gamma_1\left(\mathrm{a},-\log\left[1-(1+\mathrm{e}^{-\mathrm{x}})^{-1}\right]\right)\right\}^2}$
Gumbel	$G(x) = 1 - \exp(-\mathrm{e}^{\mathrm{x}})$	$f_X(x) = \frac{\theta\,\exp(a\,x-\mathrm{e}^{\mathrm{x}})}{\Gamma(a)\,\{\theta+(1-\theta)\gamma_1(a,\mathrm{e}^{\mathrm{x}})\}^2}$
Log-Normal	$G(x) = \Phi(\log x)$	$f_X(x) = \frac{\theta\,\phi(\log x)\,\{-\log[1-\Phi(\log x)]\}^{a-1}}{\Gamma(a)\,x\,\{\theta+(1-\theta)\gamma_1(a,-\log[1-\Phi(\log x)])\}^2}$
Exponential	$G(x) = 1 - \exp(-\lambda x)$, $\lambda > 0$	$f_X(x) = \frac{\theta\,\lambda^a\,x^{(a-1)}}{\Gamma(a)\,\{\theta+(1-\theta)\gamma_1(a,\lambda x)\}^2}$
Weibull	$G(x) = 1 - \exp(-(\lambda x)^\gamma)$, $\lambda, \gamma > 0$	$f_X(x) = \frac{\theta\,\gamma\lambda^{a\,\gamma}\,x^{a\,\gamma-1}\,\exp\{-(\lambda\,\gamma)^\gamma\}}{\Gamma(a)\{\theta+(1-\theta)\gamma_1[a,(\lambda\,x)^\gamma]\}^2}$
Gamma	$G(x) = \gamma_1(\alpha, \beta x)$, α, $\beta > 0$	$f_X(x) = \frac{\theta\,\beta^\alpha\,x^{\alpha-1}\,\mathrm{e}^{-\beta\mathrm{x}}\,\{-\log[1-\gamma_1(\alpha,\beta\mathrm{x})]\}^{\mathrm{a}-1}}{\Gamma(a)\,\{\theta+(1-\theta)\gamma_1(a,-\log[1-\gamma_1(\alpha,\beta x)])\}^2}$
Pareto	$G(x) = 1 - \frac{1}{(1+x)^\nu}$, $\nu > 0$	$f_X(x) = \frac{\theta\,\mathrm{e}^{-\mathrm{x}}\,[\nu\log(1+\mathrm{x})]^{\mathrm{a}-1}\,\mathrm{g(x)}}{\Gamma(a)\,(1+\mathrm{e}^{-\mathrm{x}})^2\,\{\theta+(1-\theta)\gamma_1(\mathrm{a},\nu\log[1+\mathrm{x}])\}^2}$
Dagum	$G(x) = [1 + (x/\beta)^{-\alpha}]^{-p}$, $\alpha, \beta, p > 0$	$f_X(x) = \frac{\theta\left\{-\log[1-[1+(x/\beta)^{-\alpha}]^{-p}]\right\}^{a-1}\,g(x)}{\Gamma(a)\{\theta+(1-\theta)\gamma_1[a,-\log(1-((x/\beta)^{-\alpha}+1)^{-p})]\}}$

3. ESTIMATION

The MOGa-G family can be fitted to real data using the **AdequacyModel** package (Marinho *et al.*, 2019) [5] in the R software. This package does not require to define the log-likelihood function, and it computes the MLEs, their standard errors (SEs), and the formal statistics defined in Section 5. It is only necessary to provide the PDF and CDF of the distribution to be fitted to a data set.

For example, if x_i is one observation from (2.2) and $\boldsymbol{\eta}$ is a q-parameter vector specifying $G(\cdot)$ as the Weibull CDF, the log-likelihood function for $\boldsymbol{\theta}^\top = (a, \theta, \boldsymbol{\eta}^\top)$ from n observations can be expressed as

$$\begin{aligned} \ell(\boldsymbol{\theta}) \;=\; & n\log(\theta) + n\log(\gamma) + n\,a\,\gamma\log(\lambda) + (a\gamma-1)\sum_{i=1}^{n}\log(x_i) - \lambda^\gamma\sum_{i=1}^{n}x_i^\gamma \\ & - n\log[\Gamma(a)] - 2\sum_{i=1}^{n}\log\{\theta + (1-\theta)\gamma_1[a,(\lambda\,x_i)^\gamma]\}. \end{aligned} \tag{3.1}$$

Due to the impossibility of obtaining the MLEs in closed form, numerical methods to calculate the estimates that maximize $\ell(\cdot)$ are necessary. Several programming languages and statistical software provides functions and routines that make it easy to obtain numerical estimates by various interactive methods. In practice, these estimates are commonly found in this way, since the Newton and quasi-Newton methods produce satisfactory results under reasonable conditions of the object function, i.e., when they do not impose restrictions that disturb the convergence of the algorithms.

The **AdequacyModel** package of the programming language R is used to obtain the MLEs, see R Core Team (2020) [6]. This library, created and maintained by one of the authors of this paper, is widely cited by several works in statistics, and serves as a basis for other library implementations available on the Comprehensive R Archive Network (CRAN). By using the goodness.fit function, it is possible to provide an implementation of (2.2), and obtain $\ell(\cdot)$ by returning the MLEs, and some measures of adequacy of fit. Further details regarding this package can be obtained from Marinho *et al.* (2019) [5].

4. SIMULATIONS

Due to the probable absence of MLEs in closed-form for distributions belonging to the MOGa-G family, it is necessary to examine the precision of the estimates calculated numerically.

In order to do that, the biases of the estimators of the parameters of the MOGa-Dagum$(\theta, a, \alpha, \beta, p)$ and MOGa-Weibull$(\theta, a, \lambda, \gamma)$ distributions are determined, where $G \sim$ Dagum(α, β, p) and Weibull(γ, λ) are the baseline models, respectively. All parameters are taken equal to one for different sample sizes reported in Tables 2 and 3.

The numbers in Tables 2 and 3 indicate that the estimation method behaves well when the sample size increases. This is theoretically expected. However, in practice, difficulties can be faced to other families due to the flatness of the log-likelihood function.

All Monte Carlo simulations can be reproduced using the script in `https://github.com/prdm0/MOGG`. The simulations are parallelized and able to use all threads available by a multicore processor, thus making them more computationally efficient, and consequently requiring less time to complete.

The simulations are performed on a computer with an Intel(R) Core(TM) i5-9500 CPU processor with 6 threads working at a maximum frequency of 3.00GHz, requiring, on this hardware, a time of 15.4828 hours to perform all simulations, 7.7414 hours for the MOGa-Dagum$(\theta, a, \alpha, \beta, p)$ distribution, and 4.9688 hours for the MOGa-Weibull$(\theta, a, \lambda, \gamma)$ distribution. Tables 2 and 3 reveal that the average biases of the MLEs could be very small for $n > 2,000$.

To generate observations from the random variable X with density f, the well-known Acceptance-Rejection Algorithm for continuous random variables is very useful when the QF involves complex functions that can lead to some numerical inaccuracies. For doing this, another random variable Y is chosen such that it can generate observations from a PDF h with the same support as f. Then, the acceptance and rejection algorithm is defined by the following steps:

1. Generate an outcome y from Y;

2. Generate an observation u from a random variable $U \sim \mathcal{U}(0,1)$;

3. If $u < \frac{f(y)}{c\,g(y)}$, where c is a real constant, accept $x = y$; otherwise reject y as an outcome from X and return to 1.

The constant c must be chosen in such a way that $\frac{f(y)}{c\,g(y)} \leq 1$. Thus, to minimize the computational cost of generating observations from X through the generated observations from Y, c is chosen as the lowest possible value to maximize the likelihood of acceptance. Further details of this method can be found in Rizzo (2019) [7].

Table 2: Average biases of the MLEs of the MOGa-Dagum$(\theta, a, \alpha, \beta, p)$ distribution calculated by the BFGS method from simulations.

n	$B(\hat{\theta})$	$B(\hat{a})$	$B(\hat{\alpha})$	$B(\hat{\beta})$	$B(\hat{p})$	Time (mins)
10	0.2213	2.1944	2.6971	1.5803	1.3190	0.6960
20	0.4240	2.4793	1.5591	1.8083	0.7414	0.9819
60	0.7458	2.2661	0.5598	1.8495	0.2812	1.9417
100	0.6194	1.9438	0.3312	1.6142	0.2935	2.7208
200	0.3950	1.4262	0.1856	1.1556	0.3611	4.4534
400	0.2077	0.9599	0.1082	0.6157	0.4076	7.4698
600	0.1200	0.7213	0.0767	0.4024	0.3572	9.4975
1000	0.0629	0.4791	0.0503	0.2123	0.2584	12.4221
2000	0.0362	0.2958	0.0298	0.1145	0.1878	20.7251
5000	−0.0040	0.1325	0.0159	0.0144	0.0167	28.3380
10000	−0.0133	0.0815	0.0096	0.0081	0.0039	50.9298
20000	−0.0111	0.0349	0.0037	0.0006	−0.0109	68.6320
30000	−0.0036	0.0191	0.0006	−0.0041	−0.0034	97.3046
50000	−0.0057	0.0129	0.0016	0.0015	−0.0026	158.3737

Table 3: Average biases of the MLEs of the MOGa-Weibull$(\theta, a, \lambda, \gamma)$ distribution calculated by the BFGS method from simulations.

n	$B(\hat{\theta})$	$B(\hat{a})$	$B(\hat{\lambda})$	$B(\hat{\gamma})$	Time (mins)
10	0.0818	0.1362	4.9274	1.2407	0.6716
20	0.3404	−0.0177	3.4160	1.4117	0.8077
60	0.7037	−0.0677	1.8806	1.3385	1.1773
100	0.6698	−0.0535	1.3684	1.1796	1.2643
200	0.5371	−0.0299	0.8265	0.9110	1.7886
400	0.3371	−0.0047	0.4205	0.5967	2.8386
600	0.2457	0.0076	0.2685	0.4306	3.6867
1000	0.1476	0.0093	0.1553	0.2818	5.0944
2000	0.0731	0.0035	0.0758	0.1530	8.8577
5000	0.0264	0.0007	0.0283	0.0618	15.8586
10000	0.0128	−0.0007	0.0142	0.0318	29.8629
20000	0.0053	−0.0012	0.0071	0.0160	48.7417
30000	0.0023	−0.0014	0.0053	0.0119	64.7387
50000	0.0023	−0.0004	0.0028	0.0063	112.7422

5. APPLICATIONS

Two applications compare the MOGa-Weibull (MOGa-W for short) model with seven extended Weibull distributions: the beta-Weibull (β-W) (Famoye *et al.*, 2005) [8], Kumaraswamy Weibull (Kw-W) (Cordeiro *et al.*, 2010) [9], Marshall-Olkin Weibull (MO-W) (Ahmed *et al.*, 2017) [10], Marshall-Olkin Extended Weibull (MOE-W) (Cordeiro *et al.*,

2019) [11], exponentiated Weibull (exp-W) (Mudholkar and Srivastava, 1993) [12], gamma Weibull (Γ-W) (Cordeiro *et al.*, 2016) [13], and exponentiated generalized Weibull (EG-W) (Oguntunde *et al.*, 2015) [14] (with $a = 1$). Some of these distributions are widely used in practice.

The log-likelihood for $\boldsymbol{\theta}$ from the MOGa-W distribution from one observation can be expressed as

$$\begin{aligned}\ell(\boldsymbol{\theta}) \;=\;& \log(\theta) + \log(\gamma) + (a\,\gamma)\log(\lambda) + (a\,\gamma - 1)\log(x) - (\gamma\,x)^{\gamma} - \log[\Gamma(a)] \\ & - 2\log\{\theta + (1-\theta)\gamma_1[a, (\lambda\,x)^{\gamma}]\},\end{aligned} \tag{5.1}$$

where $\boldsymbol{\theta} = (a, \theta, \lambda, \gamma)^{\top}$. The components of the score function are

$$U_a(\boldsymbol{\theta}) \;=\; \gamma\log(\lambda) + \gamma\log(x) - \psi^{(0)}(a) - \frac{2\big\{(1-\theta)A - (1-\theta)\psi^{(0)}(a)\gamma_1[a, (x\lambda)^{\gamma}]\big\}}{\theta\,\Gamma(a) + (1-\theta)\gamma_1[a, (\lambda x)^{\gamma}]},$$

$$U_\theta(\boldsymbol{\theta}) \;=\; \frac{1}{\theta} - \frac{2\{\Gamma(a) - \gamma_1[a, (\lambda\,x)^{\gamma}]\}}{\theta\,\Gamma(a) + (1-\theta)\gamma_1[a, (\lambda x)^{\gamma}]},$$

$$U_\lambda(\boldsymbol{\theta}) \;=\; \frac{\gamma}{\lambda}[a - (\lambda\,x)^{\gamma}] + \frac{2\gamma\,\lambda^{-1}(\lambda\,x)^{a\,\gamma}(1-\theta)\exp\{-(\lambda\,x)^{\gamma}\}}{\theta\,\Gamma(a) + (1-\theta)\gamma_1[a, (\lambda x)^{\gamma}]}$$

and

$$U_\gamma(\boldsymbol{\theta}) = \frac{1}{\gamma} + a\log(\lambda) + a\log(x) - (\lambda x)^{\gamma}\log(\lambda x) + \frac{2(1-\theta)(\lambda x)^{\gamma\,a}\log(\lambda x)\exp\{-(\lambda x)^{\gamma}\}}{\theta\,\Gamma(a) + (1-\theta)\,\gamma_1[a, (\lambda x)^{\gamma}]},$$

where $\psi^{(n)}(x)$ is the n-th derivative of the digamma function,

$$A = \log[(\lambda x)^{\gamma}]\,\gamma_1[a, (\lambda x)^{\gamma}] + G_{2,3}^{3,0}\left((\lambda x)^{\gamma}\,\middle|\, \begin{matrix} 1, 1 \\ 0, 0, a \end{matrix}\right),$$

and $G_{p,q}^{m,n}\left(z\,\middle|\, \begin{matrix} a_1, \ldots, a_p \\ b_1, \ldots, b_q \end{matrix}\right)$ is the Meijer G function.

The **AdequacyModel** package is used to fit the previous distributions to two real data sets. The SANN method, which is a variant of simulated annealing algorithm (Belisle, 1992) [15], is used here. The distributions are compared via the Anderson Darling (A^*) and Cramér von Mises (W^*) statistics reported in the `goodness.fit` function.

For the first case, the **betareg** package is applied to a modification of the "FoodExpenditure" data, which refer to the proportions of income spent on food for 38 households in a large US city (according to the package information). The household expenditures for food are given by

$$data = FoodExpenditure_{food}/\#(FoodExpenditure_{food}),$$

where $FoodExpenditure_{food}$ is the random variable corresponding to the household expenditures for food, and $\#(\cdot)$ indicates the number of observations on this variable. The observations for the first data set are given bellow:

0.4210000, 0.4382105, 0.5721316, 0.1955526, 0.2758158, 0.3565263, 0.6120000, 0.4730526, 0.3726579, 0.2322368, 0.3732632, 0.5158947, 0.3612632, 0.5563421, 0.4591053, 0.2533947, 0.3685526, 0.2410526, 0.4955526, 0.2010789, 0.3653158, 0.2544737, 0.5685263, 0.2859474, 0.7626316, 0.2863684, 0.4884474, 0.3060263, 0.4754474, 0.3826053, 0.5050526, 0.6820526, 0.7587632, 0.4176053, 0.3923684, 0.2513158, 0.6070000, 0.3881842.

Some descriptive statistics are reported in Table 4. The minimum value refers to a family of 3 people and income of 39,151, which does not represent the lowest family income for the current data, as expected, occupying only the fifth position among those with the lowest income. The maximum value corresponds to a family of 6 people with an income of 69,929, the second-largest number among the number of people per family in the group in question. Furthermore, we can note that the current data present positive asymmetry and negative kurtosis.

Table 4: Descriptive statistics for the food data.

Minimum	0.1956
1st Qu.	0.2913
Median	0.3903
Mean	0.4198
3rd Qu.	0.5027
Maximum	0.7626
Standard Deviation	0.1480
Skewness	0.5250
Kurtosis	−0.4440

In addition, the standard deviation is relatively low. Figure 2 displays the total time on test (TTT) plot for the first data set, which shows that the failure function is decreasing. So, the MOGa-W distribution is appropriate to fit these data, since its hazard function presents this shape (see, Figure 1(b)).

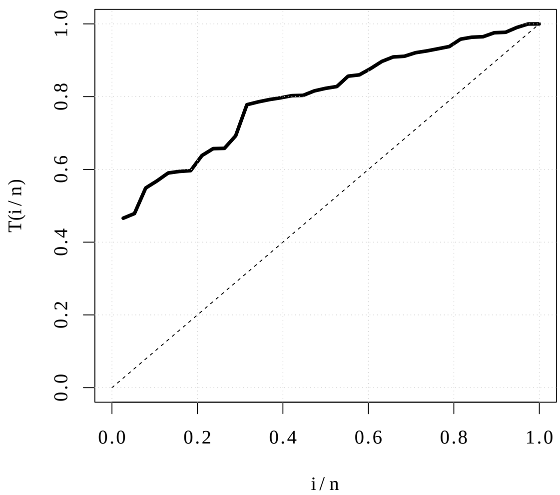


Figure 2: TTT plot for the food data.

The MLEs, their standard errors (SEs) (in parentheses), and the statistics W^* and A^* for the fitted models to the current data are listed in Table 5. The results indicate that the proposed model has better performance than the other seven fitted models.

Table 5: Estimation results for food data.

Model	$\hat{a}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	W^*	A^*
MOGa-W$(a,\theta,\lambda,\gamma)$	0.9261 (0.0262)	1.3796 (0.2238)	33.3230 (0.2853)	25.3988 (0.0825)	0.0339	0.2376
$\beta-$W$(a,\theta,\lambda,\gamma)$	9.9288 (0.0290)	0.1700 (0.0204)	9.7594 (<0.0001)	1.5305 (<0.0001)	0.0435	0.2594
KW-W$(a,\theta,\lambda,\gamma)$	0.0498 (0.0080)	99.9998 (16.2259)	1.0760 (0.0031)	23.4028 (0.0146)	1.3309	6.7426
MOE-W$(a,\theta,\lambda,\gamma)$	0.1366 (0.1599)	2.0204 (<0.0001)	62.7220 (<0.0001)	4.2956 (0.7365)	0.0354	0.2579
EGW(a,b,λ,γ)	5.6189 (0.0028)	6.1833 (0.0009)	1.2870 (0.1159)	1.3798 (0.1480)	0.0371	0.2518
MO-W(a,λ,γ)	0.1592 (0.0717)	— (—)	1.5860 (0.1335)	4.2671 (0.1665)	0.0345	0.2573
exp-W(a,λ,γ)	6.1102 (0.4222)	— (—)	4.4680 (0.3175)	1.3858 (0.1677)	0.0372	0.2523
γ-W(a,λ,γ)	5.7515 (0.0015)	— (—)	10.0000 (0.0001)	1.2087 (0.0082)	0.0879	0.6599

A data set collected in a pilot study about hypertension in the Dominican Republic in 1997 refers to the second application. The observations are the systolic blood pressure of persons who came to medical clinics in several villages for a variety of complaints. The observations for the data set in question are:

150, 120, 120, 180, 138, 115, 130, 150, 200, 120, 190, 90, 130, 120, 200, 140, 110, 134, 160, 140, 105, 126, 129, 120, 100, 130, 118, 144, 180, 138, 110, 140, 120, 118, 110, 110, 130, 140, 130, 165, 180, 130, 140, 112, 130, 158, 112, 150, 140, 142, 110, 140, 130, 132, 140, 140, 122, 128, 90, 118, 120, 110, 122, 200, 110, 140, 150, 120, 150, 120, 164, 122, 112, 130, 140, 102, 122, 130, 102, 130, 122, 200, 140, 180, 124, 110, 124, 90, 120, 159, 142, 140, 118, 122, 108, 170, 120, 140, 100, 118, 110, 114, 150, 160, 140, 190, 118, 120, 150, 120, 200, 150, 168, 110, 142, 150, 160, 142, 160, 150, 110, 128, 122, 150, 140, 122, 120, 130, 100, 130, 150, 130, 100, 120, 105, 100, 150, 196, 130, 110, 140, 122, 110, 164, 120, 120, 150, 160, 150, 135, 124, 110, 100, 95, 130, 120, 108, 118, 170, 105, 120, 95, 95, 120, 140, 142, 160, 110, 190, 180, 130, 130, 120, 204, 150, 150, 120, 122, 120, 130, 140, 148, 118, 126, 136, 140, 130, 102, 110, 110, 130, 126, 142, 140, 128, 130, 124, 162, 130, 130, 110, 80, 166, 140, 160, 160, 140, 98, 138, 120, 112, 112, 134, 140, 115, 140, 98, 115, 120, 80, 160, 126, 110, 130, 104, 236, 118, 120, 140, 120, 98, 164, 150, 110, 120, 130, 170, 180, 110, 120, 130, 118, 130, 190, 158, 90, 99, 210, 180, 140, 184, 105, 120, 150, 140, 130, 160, 118, 210, 100, 170, 150, 130, 170, 150, 120, 134, 90, 125, 170, 140, 150, 110, 105, 140, 120, 100, 124, 112, 160, 140, 118, 190, 110, 118, 160, 150, 124, 128, 150, 120, 125, 118, 132, 110, 143, 170, 98, 124, 180, 178, 110, 98, 159, 110, 140, 130, 122, 110, 98, 180, 90, 118, 165, 138, 138, 170, 106, 170, 140, 90, 118, 110, 102, 102, 180, 100, 110, 162, 140, 110, 98, 140, 140, 110, 170, 112, 90, 102, 106, 124, 110, 180, 138, 90, 150, 126, 110, 130, 150, 145, 140, 156, 110, 150, 160, 120, 140, 120, 110, 120, 140, 160, 160, 110, 150, 118, 110, 120, 120, 146, 124, 170, 124, 170, 159, 120, 120, 118, 152, 190.

Table 6 provides the descriptive statistics for the second data set. Considering that the systolic blood pressure represents the highest number presented in the pressure measuring equipment, the maximum (236) found in this table should represent an individual with serious heart problems. This is due to the fact that the normal systolic pressure is 120.

Table 6: Descriptive statistics for the clinic data.

Minimum	80
1st Qu.	118
Median	130
Mean	133
3rd Qu.	150
Maximum	236
Standard Deviation	25.7157
Skewness	0.7893
Kurtosis	0.5908

The minimum value (80) should be for an individual who probably suffers from low blood pressure. Also note that the data set has positive skewness. This indicates that few people have high pressure values and, in this case, the mode is 120, which represents a good value for systolic pressure.

Figure 3 provides the TTT plot for the second clinic data, which shows that the hazard function is decreasing, thus supporting the MOGa-W model.

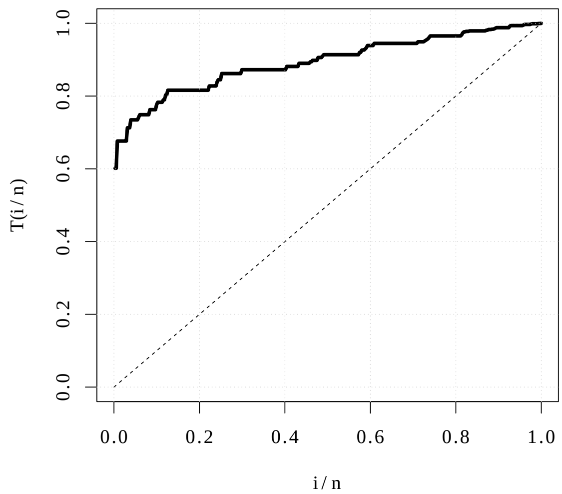


Figure 3: TTT plot for the clinic data.

The MLEs of the parameters, their SEs and the values of the adequacy measures for the fitted models to the clinic data are reported in Table 7. By comparing the measure values, the proposed distribution outperforms all other fitted models.

Table 7: Estimation results for clinic data.

Model	$\hat{a}$	$\hat{\theta}$	$\hat{\lambda}$	$\hat{\gamma}$	W^*	A^*
MOGa-W$(a, \theta, \lambda, \gamma)$	9.6293 (0.0062)	3.6407 (0.1824)	6.2608 (0.0244)	12.8234 (0.0079)	0.5093	2.8076
β−W$(a, \theta, \lambda, \gamma)$	31.0847 (0.0127)	47.1463 (<0.0001)	0.0169 (0.0001)	2.0540 (<0.0001)	0.7540	4.2794
KW-W$(a, \theta, \lambda, \gamma)$	7363.2810 (0.0419)	0.0392 (0.0004)	1.4676 (<0.0001)	0.6146 (<0.0001)	0.5351	2.9617
MOE-W$(a, \theta, \lambda, \gamma)$	101.1347 (46.7828)	0.4238 (0.1008)	0.0309 (0.0036)	1.7024 (0.1681)	1.1441	6.6446
EGW(a, b, λ, γ)	0.2351 (0.0025)	140.0000 (3.2785)	0.4576 (<0.0001)	0.7425 (<0.0001)	0.8925	5.3338
MO-W(a, λ, γ)	173.2139 (0.0001)	— (—)	0.0212 (0.0002)	1.6019 (0.0001)	1.4760	8.5870
exp-W(a, λ, γ)	69.0291 (0.0838)	— (—)	0.0240 (0.0002)	1.3455 (<0.0001)	0.8899	5.0941
Γ-W(a, λ, γ)	9.1122 (0.9633)	— (—)	0.0261 (0.0029)	1.7464 (0.0803)	0.6227	3.4882

6. MATHEMATICAL PROPERTIES

Here, some mathematical properties for the MOGa-G family are presented based on a linear representation for its density function in terms of “exponentiated-G” (exp-G) densities.

6.1. Linear Representation

For an arbitrary CDF $G(x)$, the CDF and PDF of the exp-G distribution with power parameter $a > 0$ are

$$\Pi_a(x) = G(x)^a \qquad \text{and} \qquad \pi_a(x) = a\, g(x)\, G(x)^{a-1},$$

respectively. This class of distributions is quite useful in several applications. In fact, Tahir and Nadarajah (2015) [1] cited more than seventy papers on exponentiated distributions in their Table 1.

First, the CDF of the MO-G distribution (1.2) admits the linear representation (Barreto-Souza *et al.*, 2013) [16]

$$F_{\text{MO}-\Gamma}(x) = \sum_{i=0}^{\infty} w_i^{\text{MO-G}}\, \Pi_{i+1}(x) = \sum_{i=0}^{\infty} w_i^{\text{MO-G}}\, G(x)^{i+1}, \tag{6.1}$$

where the coefficients are (for $i = 0, 1, ...$)

$$w_i^{\text{MO}-\Gamma} = w_i^{\text{MO}-\Gamma}(\theta) = \begin{cases} \dfrac{(-1)^i\, \theta}{(i+1)} \sum\limits_{j=i}^{\infty} (j+1) \dbinom{j}{i} \bar{\theta}^j, & \theta \in (0,1), \\ \theta^{-1}(1-\theta^{-1})^i, & \theta > 1, \end{cases}$$

and $\bar{\theta} = 1 - \theta$.

Second, the linear combination for the Γ-G cumulative distribution (1.4) follows from Castellares and Lemonte (2015) [17] as

$$F_{\Gamma\text{-G}}(x) = \sum_{j=0}^{\infty} w_j^{\Gamma\text{-G}}\, \Pi_{a+j}(x). \tag{6.2}$$

Here,

$$w_j^{\Gamma\text{-G}} = w_j^{\Gamma\text{-G}}(a) = \frac{\varphi_j(a)}{(a+j)},$$

$$\varphi_0(a) = \frac{1}{\Gamma(a)}, \quad \varphi_j(a) = \frac{(a-1)}{\Gamma(a)}\,\psi_{j-1}(j+a-2), \quad j \geq 1,$$

and

$$\psi_{n-1}(x) = \frac{(-1)^{n-1}}{(n+1)!}\Bigg[H_n^{n-1} - \frac{x+2}{n+2}H_n^{n-2} + \frac{(x+2)(x+3)}{(n+2)(n+3)}H_n^{n-3} - \cdots + (-1)^{n-1}\frac{(x+2)(x+3)\cdots(x+n)}{(n+2)(n+3)\cdots(2n)}H_n^0\Bigg]$$

is the Stirling polynomial, $H_{n+1}^m = (2n+1-m)H_n^m + (n-m+1)H_n^{m-1}$ is a positive integer, $H_0^0 = 1$, $H_{n+1}^0 = 1 \times 3 \times 5 \times \cdots \times (2n+1)$ and $H_{n+1}^n = 1$.

By inserting (6.2) in Equation (6.1) and via a result for a power series raised to a positive integer (Gradshteyn and Ryzhik, 2000) [18], the expansion for the cdf of the MOGa-G distribution reduces to

$$\begin{aligned} F_{\mathrm{MO}-\Gamma\text{-G}}(x) &= \sum_{i=0}^{\infty} w_i^{\mathrm{MO}\text{-}\Gamma}\, G(x)^{(i+1)a}\left[\sum_{j=0}^{\infty} w_j^{\Gamma\text{-G}}\, G(x)^j\right]^{i+1} \\ &= \sum_{i=0}^{\infty} w_i^{\mathrm{MO}\text{-}\Gamma}\, G(x)^{(i+1)a} \sum_{j=0}^{\infty} c_{i+1,j}\, G(x)^j = \sum_{i,j=0}^{\infty} d_{i,j}\, \Pi_{(i+1)a+j}(x), \end{aligned}$$

where $d_{i,j} = d_{i,j}(a,\theta) = w_i^{\mathrm{MO\text{-}G}}\, c_{i+1,j}(a)$, $c_{i+1,0}(a) = (w_0^{\Gamma\text{-G}})^{i+1}$ and, for $m \geq 1$, $c_{i+1,m}(a) = \frac{1}{m w_0^{\Gamma\text{-G}}} \sum_{r=1}^{m} [r(i+2)-m]\, w_r^{\Gamma\text{-G}}\, c_{i+1,m-r}(a)$.

By differentiating the last equation, the linear representation for the MOGa-G density holds

$$f_{\mathrm{MO}-\Gamma\text{-G}}(x) = \sum_{i,j=0}^{\infty} d_{i,j}\, \pi_{(i+1)a+j}(x). \tag{6.3}$$

So, some structural properties of the proposed family can be determined from the double linear combination (6.3) and those properties of the exp-G distribution. In most applications, the indices i and j can vary up to a small number of terms.

6.2. Some quantities

Hereafter, let $T_{i,j} \sim \text{exp-G}[(i+1)a+j]$. The n-th moment of X can be determined from (6.3) as

$$\mu'_n = E(X^n) = \sum_{i,j=0}^{\infty} d_{i,j}\, E(T_{i,j}) = \sum_{i,j=0}^{\infty} [(i+1)a+j-1]\, d_{i,j}\, \tau[n,(i+1)a+j-1], \tag{6.4}$$

where

$$\tau(n,a) = \int_{-\infty}^{\infty} x^n\, G(x)^a\, g(x)dx = \int_0^1 Q_G(u)^n\, u^a du.$$

Expressions for moments of several exponentiated distributions can be found in the papers cited in Tahir and Nadarajah (2015, Table 1). We give just one example from Equation (6.4) by taking the exponential distribution with rate $\lambda > 0$ for the baseline G. It follows easily as

$$\mu'_n = n!\,\lambda^n \sum_{i,j,m=0}^{\infty} \frac{(-1)^{n+m}\,[(i+1)a+j]\,d_{i,j}}{(m+1)^{n+1}} \binom{(i+1)a+j-1}{m}.$$

For empirical purposes, the shape of many distributions can be usefully described by the incomplete moments. These moments play an important role for measuring inequality. For example, the mean deviations and Lorenz and Bonferroni curves depend upon the first incomplete moment of the distribution. The n-th incomplete moment of X can be expressed as

$$m_n(y) = \int_{-\infty}^{y} x^n\, f_X(x)dx = \sum_{i,j=0}^{\infty} [(i+1)a+j]\, d_{i,j} \int_0^{G(y)} Q_G(u)^n\, u^{(i+1)a+j-1}\, du. \tag{6.5}$$

The definite integral in (6.5) can be evaluated for most baseline G distributions.

The moment generating function (MGF) $M(t) = E(\mathrm{e}^{\mathrm{t\,X}})$ of X can be expressed from (6.3)

$$M(t) = \sum_{i,j=0}^{\infty} d_{i,j}\, M_{i,j}(t) = \sum_{i,j=0}^{\infty} [(i+1)a+j]\, d_{i,j}\, \rho(t,(i+1)a+j-1), \tag{6.6}$$

where $M_{i,j}(t)$ is the MGF of $Y_{i,j}$ and

$$\rho(t,a) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{tx}}\, \mathrm{G(x)^a\, g(x)dx} = \int_0^1 \exp\{\mathrm{t\, Q_G(u)}\}\, \mathrm{u^a du}.$$

The MGFs of several MOGa-G distributions can be determined from Equation (6.6). For example, the generating function of the MOGa-exponential with parameter λ (if $t < \lambda^{-1}$) is

$$M(t) = \sum_{i,j=0}^{\infty} [(i+1)a+j]\, d_{i,j}\; B((i+1)a+j, 1-\lambda t).$$

7. CONCLUSIONS

A new family of distributions called the Marshall and Olkin-Gamma-G family with two shape parameters is introduced. The estimation of the unknown parameters is done via the maximum likelihood method and a simulation study is conducted to verify its adequacy. Additionally, the usefulness of the proposed family is shown empirically by means of two applications to real data. In fact, the new family can generate very competitive distributions with the same number of parameters than others constructed by existing classes.

REFERENCES

[1] Tahir, M.H. and Nadarajah, S. (2015). Parameter induction in continuous univariate distributions: Well-established G families, *Anais da Academia Brasileira de Ciências*, **87**, 539–568.

[2] Eugene, N.; Lee, C. and Famoye, F. (2002). Beta-normal distribution and its applications, *Communications in Statistics-Theory and Methods*, **31**(4), 497–512.

[3] Zografos, K. and Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference, *Statistical Methodology*, **6**(4), 344–362.

[4] Marshall, A.W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families, *Biometrika*, **84**(3), 641–652.

[5] Marinho, P.R.D.; Silva, R.B.; Bourguignon, M.; Cordeiro, G.M. and Nadarajah, S. (2019). AdequacyModel: An R package for probability distributions and general purpose optimization, *PloS one*, **14**(8), e0221487.

[6] Team, R.C. (2022). R: A language and environment for statistical computing, *R Foundation for Statistical Computing*, Vienna, Austria. `http://www.R-project.org/`

[7] Rizzo, Maria L. (2019). Statistical computing with R, *CRC Press*.

[8] Famoye, F.; Lee, C. and Olumolade, O. (2005). The beta-Weibull distribution, *Journal of Statistical Theory and Applications*, **4**(2), 121–136.

[9] Cordeiro, G.M.; Ortega, E.M. and Nadarajah, S. (2010). The Kumaraswamy Weibull distribution with application to failure data, *Journal of the Franklin Institute*, **347**(8), 1399–1429.

[10] Ahmad, H.H.; Bdair, O.M.; Ahsanullah, M. and Ahsanullah, M. (2017). On Marshall-Olkin Extended Weibull Distribution, *J. Stat. Theory Appl.*, **16**(1), 1–17.

[11] Cordeiro, G.M.; Prataviera, F.; do Carmo S. Lima, M. and Ortega, E.M. (2019). The Marshall-Olkin extended flexible Weibull regression model for censored lifetime data, *Model Assisted Statistics and Applications*, **14**(1), 1–17.

[12] 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.

[13] Cordeiro, G.M.; do Carmo S. Lima, M.; Gomes, A.E.; da-Silva, C.Q. and Ortega, E.M. (2016). The gamma extended Weibull distribution, *Journal of Statistical Distributions and Applications*, **3**(1), 1–19.

[14] Oguntunde, P.E.; Odetunmibi, O.A. and Adejumo, A.O. (2015). On the exponentiated generalized Weibull distribution: A generalization of the Weibull distribution, *Indian Journal of Science and Technology*, **8**(35), 1–7.

[15] Bélisle, C.J. (1992). Convergence theorems for a class of simulated annealing algorithms on Rd, *Journal of Applied Probability*, **29**(4), 885–895.

[16] Barreto-Souza, W.; Lemonte, A.J. and Cordeiro, G.M. (2013). General results for the Marshall and Olkin's family of distributions, *Anais da Academia Brasileira de Ciências*, **85**(1), 3–21.

[17] Castellares, F. and Lemonte, A.J. (2015). A new generalized Weibull distribution generated by gamma random variables, *Journal of the Egyptian Mathematical Society*, **23**(2), 382–390.

[18] Gradshteyn, I.S. and Ryzhik, I.M. (2014). Table of integrals, series, and products, *Academic Press*.

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 309–319
https://doi.org/10.57805/revstat.v22i3.506

Trend Resistant General Efficiency Balanced Block Designs for Two Disjoint Sets of Treatments

Authors: Kader Ali Sarkar
– Visva-Bharati University, Santiniketan, West Bengal, India
kader804@gmail.com

Seema Jaggi
– Indian Council of Agricultural Research, New Delhi, India
seema.jaggi@icar.gov.in

Arpan Bhowmik ✉
– ICAR-Indian Agricultural Research Institute, Gogamukh, Assam
arpan.stat@gmail.com; arpan.bhowmik@icar.gov.in

Eldho Varghese
– ICAR-Central Marine Fisheries Research Institute, Kochi, Kerala, India
eldho.varghese@icar.gov.in

Cini Varghese
– ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
cini.varghese@icar.gov.in

Anindita Datta
– ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
anindita.datta@icar.gov.in

Ashutosh Dalal
– ICAR-Indian Agricultural Statistics Research Institute, New Delhi, India
ashutosh.dalal97@gmail.com

Received: December 2020 Revised: June 2022 Accepted: June 2022

Abstract:

- General Efficiency Balance (GEB) is an important property of designs. Variance balance and efficiency balance are special cases of GEB. Here, GEB block designs for comparing treatments belonging to two disjoint sets in the presence of systematic trend have been discussed. Methods of constructing Trend Resistant General Efficiency Balanced Bipartite Block Designs (TR-GEBBPB) designs have been presented. The block designs so obtained are trend resistant, general efficiency balanced and are more efficient for estimating the contrasts pertaining to two treatments from two different sets.

Keywords:

- *block design; balanced bipartite; general efficiency; trend.*

AMS Subject Classification:

- 62Kxx, 62K10, 05B05.

✉ Corresponding author

1. INTRODUCTION

Balancing is an essential and desirable statistical property of block design. The concept of balance has been used in several senses in the literature, viz., variance balance, efficiency balance, pairwise balance, general efficiency balance etc. The concept of general efficiency balance was given by [9]. When an incomplete block design is compared against any other design i.e. either a completely randomized design (CRD) or randomized complete block design (RCBD) both having same number of treatments, but not necessarily the same number of replications such that the ratio of variances of the estimates of any treatment contrast for two designs is constant, then such an incomplete block design has been called GEB design.

Definition 1.1. A connected block design is called a General Efficiency Balanced (GEB) design, if for some $\theta, s_1, s_2, ..., s_v (> 0)$, the information matrix ($\mathbf{C}$) can be expressed as

$$\mathbf{C} = \theta\left[\mathbf{S} - \frac{1}{g}\mathbf{s}\mathbf{s}'\right], \tag{1.1}$$

where $\theta = \{n - \text{trace}(\mathbf{N}\mathbf{K}^{-1}\mathbf{N}')\}/(g - \frac{1}{g}\mathbf{s}'\mathbf{s})$, $\mathbf{S} = \text{diag}(s_1, s_2, ..., s_v)$, $\mathbf{s} = (s_1, s_2, ..., s_v)'$ and $\mathbf{s}'\mathbf{1} = g$. $\mathbf{N}$ is the $v \times b$ incidence matrix of treatments vs blocks, $\mathbf{K}$ is the diagonal matrix of block sizes and n is the total number of observations.

Several series of variance balanced and efficiency balanced designs as subclasses of GEB designs through the technique of reinforcement were constructed by [9]. It was pointed out that a variance balanced design or an efficiency balanced design cannot be constructed in $(v+2)$ or more treatments through reinforcement. A method of constructing GEB designs through method of reinforcement of a Balanced Incomplete Block (BIB) design was given by [24]. They found that if one new treatment is added to each block of the BIB design, then the resultant design will be a GEB design with $(v+1)$ treatments. Different aspects of efficiency-balanced designs have been studied in [30] and [34].

Definition 1.2. A connected block design with $v^* = v + 1$ treatments, b^* blocks, block sizes k, replication numbers $\mathbf{r} = (r\mathbf{1}'_v \;\; r_0)'$ and $\mathbf{C}$ of the form

$$\mathbf{C} = \begin{bmatrix} (a+b)\mathbf{I}_v - b\mathbf{1}_v\mathbf{1}'_v & -c\mathbf{1}_v \\ -c\mathbf{1}'_v & d \end{bmatrix} \tag{1.2}$$

is a GEB design with $\mathbf{s} = [b\mathbf{1}'_v \;\; c]'$ and $g = vb+c$, where a, b, c and d are positive integers satisfying $a - b(v-1) = c$, $cv = d$, r is the replication of v treatments and r_0 is the replication of $(v+1)$-th treatment.

The above $\mathbf{C}$-matrix is identical to the structure of the $\mathbf{C}$-matrix of a **Balanced Test Treatment Incomplete Block (BTIB)** design given by [2]. This equivalence shows that GEB designs are identical to BTIB designs and can be useful in making test treatments control comparisons. In [32], it has been shown that all the BTIB designs are also GEB designs and vice-versa for a single control case. However for many controls, this result does not hold good.

Example 1.1. The block contents of a GEB design (see in [32]) with parameters $v^* = 7$ (= 6+1), $b^* = 11$, $r = 7$, $r_0 = 2$ and $k = 4$ are (1, 2, 3, 4), (5, 6, 1, 2), (3, 4, 5, 6), (5, 2, 1, 4), (3, 6, 5, 2), (1, 6, 3, 2), (5, 4, 1, 6), (3, 2, 5, 4), (1, 4, 3, 6), (1, 3, 5, 7) and (2, 4, 6, 7). The information matrix for this design is

$$\mathbf{C} = \frac{1}{4}\begin{bmatrix} 25\mathbf{I}_6 - 4\mathbf{1}_6\mathbf{1}_6 & -\mathbf{1}_6' \\ -\mathbf{1}_6' & 6 \end{bmatrix}.$$

It is seen that this is a GEB design with $\mathbf{s} = \frac{1}{4}[4\mathbf{1}_6' \;\; 1]'$ and $g = 6.25$. Some methods of constructing GEB design with equal and unequal block sizes were given by [10] along with a catalogue of GEB designs with efficiencies.

Example 1.2. The block contents of a GEB design (see in [10]) with parameters $v^* = 7$ (= 6+1), $b^* = 11$, $r = 4$, $r_0 = 9$ and $k = 3$ are (1, 2, 7), (3, 4, 7), (5, 6, 7), (1, 6, 7), (3, 2, 7) (5, 4, 7), (1, 4, 7), (3, 6, 7), (5, 2, 7), (1, 3, 5) and (2, 4, 6). The information matrix for the given design is

$$\mathbf{C} = \frac{1}{3}\begin{bmatrix} 8\mathbf{I}_6 - 4\mathbf{1}_6\mathbf{1}_6' & -\mathbf{1}_6 \\ -\mathbf{1}_6' & 18 \end{bmatrix}.$$

This is a GEB design with $\mathbf{s} = \frac{1}{3}[4\mathbf{1}_6' \;\; 1]'$ and $g = 8.33$.

A definition of GEB design for the case when there are treatments belonging to two disjoint sets is given below.

Definition 1.3. Consider a design d with $v = v_1 + v_2$ treatments (where v_1 is the number of treatments belonging to 1-st set and v_2 is the number of treatments belonging to 2-nd set; v_1, $v_2 \geq 2$) having a $\mathbf{C}$ matrix of the form

$$\mathbf{C} = \begin{bmatrix} (f_1 - f_2)\mathbf{I}_{v_1} + f_2\mathbf{1}_{v_1}\mathbf{1}_{v_1}' & -f_3\mathbf{1}_{v_1}\mathbf{1}_{v_2}' \\ -f_3\mathbf{1}_{v_2}\mathbf{1}_{v_1}' & (f_4 - f_5)\mathbf{I}_{v_2} + f_5\mathbf{1}_{v_2}\mathbf{1}_{v_2}' \end{bmatrix}, \tag{1.3}$$

where $f_1, f_2, f_3, f_4, f_5 > 0$ and $f_1 = f_2v_1 + f_3v_2$, and $f_4 = f_3v_1 + f_5v_2$. The design d is said to be a GEB design if and only if $f_2f_5 = f_3^2$. It can be shown that the $\mathbf{C}$-matrix of a GEB design given in (1.3) can be expressed in the form of $\mathbf{C}$-matrix of (1.1) with

$$\mathbf{s} = [f_2\mathbf{1}_{v_1}' \;\; f_3\mathbf{1}_{v_2}']', \quad \mathbf{S} = \begin{bmatrix} f_2\mathbf{I}_{v_1} & 0 \\ 0 & f_3\mathbf{I}_{v_2} \end{bmatrix}, \quad g = f_2v_1 + f_3v_2 \quad \text{and} \quad \theta = \frac{f_1}{f_2}.$$

The above $\mathbf{C}$-matrix in (1.3) is identical to the structure of the $\mathbf{C}$-matrix of the balanced block design obtained for comparing two disjoint sets of treatments called Balanced Bipartite Block (BBPB) design. The interest here is to estimate the contrasts of the type $(\tau_i - \tau_j)$ with as high precision as possible, where τ_i and τ_j belong to 1-st and 2-nd set of treatments, respectively.

Example 1.3. The block contents of a GEB design (see in [18]) with parameters $v_1 = 8$, $v_2 = 2$, $b = 18$, $r_1 = 5$, $r_2 = 16$ and $k = 4$ are (1, 2, 9, 10), (3,4, 9, 10), (5, 6, 9, 10), (7, 8, 9, 10), (1, 4, 9, 10), (3, 2, 9, 10), (5, 8, 9, 10), (7, 6, 9, 10), (6, 1, 9, 10), (8, 3, 9, 10), (2, 5, 9, 10), (4, 7, 9, 10), (8, 1, 9, 10), (6, 3, 9, 10), (4, 5, 9, 10), (2, 7, 9, 10), (1, 3, 5, 7) and (2, 4, 6, 8). The information matrix of this design is

$$\mathbf{C} = \frac{1}{4}\begin{bmatrix} 16\mathbf{I}_8 - \mathbf{1}_8\mathbf{1}_8' & -4\mathbf{1}_8\mathbf{1}_2' \\ -4\mathbf{1}_2\mathbf{1}_8' & 64\mathbf{I}_2 - 16\mathbf{1}_2\mathbf{1}_2' \end{bmatrix}.$$

Here $f_3^2 = f_2 f_5$ and the design is a GEB design with $\mathbf{s} = \frac{1}{4}[\mathbf{1}_8' \;\; 4\mathbf{1}_2']$, g=3 and $\theta = 4$.

An overview of block designs for comparing test treatments with control treatments was given in [11]. A method of constructing GEB block designs with unequal block sizes for comparing two disjoint sets of treatments, with each set consisting of two or more treatments, has been developed by [22]. Optimal first order circular block designs with fewer blocks considering the correlated observations for an even number of treatments have been constructed in [31]. They developed GEB circular block designs with correlated observations for an even number of treatments. A-optimal/efficient designs for making the comparison between treatments that belongs to two disjoint sets with equal and unequal blocks were obtained by different authors (see for details [25, 21, 19, 20, 23, 15, 12]). Some methods of construction of BBPB designs using incidence matrices of BIB designs and two-associate-class partially balanced incomplete block group divisible designs were discussed in [33]. In another case of block design setup, experiments may be carried out using plots occurring in long, narrow rows wherein spatial fertility trends may occur. In such situations, the response may also depend on the spatial position of the experimental unit within a block. One way to overcome such situations is the suitable arrangement of treatments over plots within a block such that the arranged design is capable of completely eliminating the effects of defined components of a common trend. Such designs have been called Trend Free Block (TFB) designs (see in [6]). These designs are constructed so that treatment effects and trend effects are orthogonal. A necessary and sufficient condition for a block design to be linear trend free was obtained in [35], and the concepts and properties of Nearly TFB designs with linear and quadratic trends over plots within blocks were highlighted in [36]. A lot of literature is also available which deals with different aspects of block designs incorporating trend effects (see, for instance, [4, 5, 3, 13, 14, 16, 17, 26]). An algorithm to construct a series of exact optimum designs resistant to linear and quadratic time trends has been developed by [1]. An integer programming approach for the construction of trend-free split-plot designs was developed by [7].

This article deals with Trend Resistant General Efficiency Balanced Bipartite Block (TR-GEBBPB) designs when there are two disjoint sets of treatments (one set may be tests and other may be controls). Series of TR-GEBBPB designs for comparing a treatment from set 1 to a treatment from set 2, with more precision have been developed. The interest here is to estimate the contrasts pertaining to test treatments vs. control treatments with higher precision in the presence of trend.

2. GEBBPB DESIGNS IN THE PRESENCE OF TREND

Consider the following model in block design set-up for v treatments ($v = v_1 + v_2$; v_1 treatments in first set and v_2 treatments in second set) and b blocks of size k each incorporating trend component (within-block trend effects are represented by orthogonal polynomials of p-th degree, $p \leq k$):

$$\mathbf{Y} = \mu\mathbf{1} + \boldsymbol{\Delta}'\boldsymbol{\tau} + \mathbf{D}'\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\rho} + \mathbf{e}, \tag{2.1}$$

where $\mathbf{Y}$ is a $n \times 1$ vector of observations, μ is general mean, $\mathbf{1}$ is a $n \times 1$ vector of unity, $\boldsymbol{\Delta'}$ is a $n \times (v_1 + v_2)$ matrix of observations versus treatments, τ is a $(v_1 + v_2) \times 1$ vector of treatment effects, $\mathbf{D}'$ is a $n \times b$ incidence matrix of observations versus blocks, $\boldsymbol{\beta'}$ is a $b{\times}1$ vector of block effects, $\mathbf{Z}\boldsymbol{\rho}$ represents the trend effects. The matrix $\mathbf{Z}$, of order $n{\times}p$, is the matrix of coefficients given by $\mathbf{Z} = \mathbf{1}_b \otimes \mathbf{F}$ where $\mathbf{F}$ is a $k \times p$ matrix with columns representing the (normalized) orthogonal polynomials and $\mathbf{e}$ is a $n{\times}1$ vector of errors with $\mathrm{E}(\mathbf{e}) = 0$ and $\mathrm{V}(\mathbf{e}) = \sigma^2\, \mathbf{I}_n$. Further, $\mathbf{1}'\, \mathbf{F} = 0$, $\mathbf{F}'\, \mathbf{F} = \mathbf{I}_p$.

Let $\mathbf{N}$ be a $(v_1 + v_2) \times b$ incidence matrix, which is partitioned as

$$\boldsymbol{\Delta}\mathbf{D}' = \mathbf{N} = \begin{pmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \end{pmatrix},$$

where $\mathbf{N}_1$ is a $v_1 \times b$ incidence matrix pertaining to v_1 treatments and $\mathbf{N}_2$ is a $v_2 \times b$ incidence matrix pertaining to v_2 treatments. The model (2.1) can be written as

$$\mathbf{Y} = \mathbf{X}_1\boldsymbol{\theta}_1 + \mathbf{X}_2\boldsymbol{\theta}_2 + \mathbf{e}, \tag{2.2}$$

where $\mathbf{X}_1 = [\boldsymbol{\Delta}'] = [\boldsymbol{\Delta}_1' \;\; \boldsymbol{\Delta}_2']$, $\mathbf{X}_2 = [1 \;\; \mathbf{D}' \;\; \mathbf{Z}]$, $\boldsymbol{\theta}_1 = \boldsymbol{\tau}$ and $\boldsymbol{\theta}_1 = [\mu \;\; \boldsymbol{\beta}' \;\; \boldsymbol{\rho}']'$.

$\mathbf{X}_1$ is the matrix of effects of interest and $\mathbf{X}_2$ is the matrix of nuisance effects. The joint information matrix for estimating different effects is obtained as:

$$\mathbf{C} = \begin{bmatrix} r_1\mathbf{I}_{v_1} - \frac{1}{k}\mathbf{N}_1\mathbf{N}_1' - \frac{1}{b}\boldsymbol{\Delta}_1\mathbf{Z}\mathbf{Z}'\boldsymbol{\Delta}_1' & -\frac{1}{k}\mathbf{N}_1\mathbf{N}_2' - \frac{1}{b}\boldsymbol{\Delta}_1\mathbf{Z}\mathbf{Z}'\boldsymbol{\Delta}_2' \\ -\frac{1}{k}\mathbf{N}_2\mathbf{N}_1' - \frac{1}{b}\boldsymbol{\Delta}_2\mathbf{Z}\mathbf{Z}'\boldsymbol{\Delta}_1' & r_2\mathbf{I}_{v_2} - \frac{1}{k}\mathbf{N}_2\mathbf{N}_2' - \frac{1}{b}\boldsymbol{\Delta}_2\mathbf{Z}\mathbf{Z}'\boldsymbol{\Delta}_2' \end{bmatrix},$$

where r_1 and r_2 are the replications of the first and second set of treatments, respectively.

Definition 2.1. A bipartite block design is said to be balanced with respect to set 1 vs set 2 if each treatment from a set appears together with every other treatment of the same set a constant number of times (say, λ^*_{ii}, i = 1,2) and each treatment from a set appears together with every other treatment of a different set a constant number of times (say, λ^*_{12}).

Definition 2.2. A bipartite block design is said to be general efficiency balanced i.e. GEBBPB if its information matrix ($\mathbf{C}$) is of the form (1.3).

Definition 2.3. A GEBBPB design is said to be Trend Resistant (TR-GEBBPB) design if the adjusted treatment sum of squares of block model with trend is same as adjusted treatment sum of squares of block model without trend.

3. METHODS OF CONSTRUCTING TR-GEBBPB DESIGNS

3.1. Method 1

Consider a Semi-regular (SR) group divisible design with parameters $v_1 = mn$ $(m < n)$, b_1, r_1, k_1, λ_{11} and λ_{12}. Consider the $(m,\ n)$ group divisible association scheme in m blocks each of size n each with $v_1 = mn$, $b_2 = m$, $r_2 = 1$, $k_2 = n$, $\lambda_{21} = 1$ and $\lambda_{22} = 0$. Augment $(k_2 - k_1) = v_2$ number of treatments to the SR design and juxtapose both the design and the association scheme. Fold-over the whole plan and the resultant design is a TR-GEBBPB design with parameters $v_1 = mn$, v_2, $b = 2(b_1 + b_2)$, $\mathbf{r}' = [2(r_1 + r_2)\mathbf{1}'_{v_1}\ \ 2b_1\mathbf{1}'_{v_2}], k = k_2, \lambda^*_{11} = 2\lambda_{12}, \lambda^*_{12} = 2r_1$ and $\lambda^*_{22} = 2b_1$. The information matrix for this design is given by

$$\mathbf{C} = \frac{2}{k}\begin{bmatrix} r_1^2\mathbf{I}_{v_1} - \mathbf{1}_{v_1}\mathbf{1}'_{v_1} & -r_1\mathbf{1}_{v_1}\mathbf{1}'_{v_2} \\ -r_1\mathbf{1}_{v_2}\mathbf{1}'_{v_1} & r_1^2 k_2\mathbf{I}_{v_2} - r_1^2\mathbf{1}_{v_2}\mathbf{1}'_{v_2} \end{bmatrix}.$$

Example 3.1.1. Consider a SR group divisible design (SR 9 in [8]) with parameters $v_1 = 8$, $b_1 = 16$, $r_1 = 4$, $k_1 = 2$, $m = 2$, $n = 4$, $\lambda_{11} = 0$ and $\lambda_{12} = 1$. The (2, 4) group divisible association scheme with two blocks each of size four is (1, 3, 5, 7) and (2, 4, 6, 8) with $v_1 = 8$, $b_2 = 2$, $r_2 = 1$, $k_2 = 4$, $\lambda_{21} = 1$ and $\lambda_{22} = 0$. Following above procedure a TR-GEBBPB design with parameters $v_1 = 8$, $v_2 = 2$, $b = 36$, $k = 4$, $\mathbf{r}' = [10\mathbf{1}'_8\ \ 32\mathbf{1}'_2], \lambda^*_{11} = 2, \lambda^*_{12} = 8$ and $\lambda^*_{22} = 32$ is obtained with block contents as (1, 2, 9, 10), (3,4, 9, 10), (5, 6, 9, 10), (7, 8, 9, 10), (6, 1, 9, 10), (8, 3, 9, 10), (2, 5, 9, 10), (4, 7, 9, 10), (1, 4, 9, 10), (3, 2, 9, 10), (5, 8, 6, 8), (7, 6, 9, 10), (8,1, 9, 10), (6, 3, 9, 10), (4, 5, 9, 10), (2, 7, 9, 10), (1, 3, 5, 7), (2, 4, 6, 8), (10, 9, 2, 1), (10, 9, 2, 1), (10, 9, 4, 3), (10, 9, 6, 5), (10, 9, 8, 7), (10, 9, 1, 6), (10, 9, 3, 8), (10, 9, 5, 2), (10, 9, 7, 4), (10, 9, 4, 1), (10, 9, 2, 3), (10, 9, 8, 5), (10, 9, 6, 7), (10, 9, 1, 8), (10, 9, 3, 6), (10, 9, 5, 4), (10, 9, 7, 2), (7, 5, 3, 1) and (8, 6, 4, 2). Here, for the given design, the normalized orthogonal polynomial of degree 1 is given as

$$\mathbf{F} = \left[\frac{-3}{\sqrt{20}}\ \frac{-1}{\sqrt{20}}\ \frac{1}{\sqrt{20}}\ \frac{3}{\sqrt{20}}\right]' = \left[-0.67\ \ -0.22\ \ 0.22\ \ 0.67\right]'.$$

The information matrix for this design is given as

$$\mathbf{C} = \frac{1}{2}\begin{bmatrix} 16\mathbf{I}_8 - \mathbf{1}_8\mathbf{1}'_8 & -4\mathbf{1}_8\mathbf{1}'_2 \\ -4\mathbf{1}_2\mathbf{1}'_8 & 64\mathbf{I}_2 - 16\mathbf{1}_2\mathbf{1}'_2 \end{bmatrix}.$$

It can be seen that here $f_2 f_5 = f_3^2$. Variance of any estimated elementary contrast among the treatments belonging to the first set is $V_{11} = 0.2500\ \sigma^2$ and the variance of any estimated elementary contrast between the treatments belonging to the first and second set is $V_{12} = 0.1562\ \sigma^2$.

Example 3.1.2. Consider a SR group divisible design (SR 11 in [8]) with parameters $v_1 = 10$, $b_1 = 25$, $r_1 = 5$, $k_1 = 2$, $m = 2$, $n = 5$, $\lambda_{11} = 0$ and $\lambda_{12} = 1$. The (2, 5) group divisible association scheme with $v_1 = 10$, $b_2 = 2$, $r_2 = 1$, $k_2 = 5$, $\lambda_{21} = 1$ and $\lambda_{22} = 0$ is as follows:

1	3	5	7	9
2	4	6	8	10

The block contents of the TR-GEBBPB design obtained with parameters $v_1 = 10$, $v_2 = 3$, $b = 54$, $k = 5$, $\mathbf{r}' = \left[\, 12\mathbf{1}'_{10} \;\; 50\mathbf{1}'_{2} \,\right]$, $\lambda^*_{11} = 2$, $\lambda^*_{12} = 10$ and $\lambda^*_{22} = 50$, are (1, 2, 11, , 12, 13), (3, 10, 11, 12, 13), (5, 8, 11, 12, 13), (7, 6, 11, 12, 13), (9, 4, 11, 12, 13), (1, 8, 11, 12, 13), (3, 6, 11, 12, 13), (5, 4, 11, 12, 13), (7, 2, 11, 12, 13), (9, 10, 11, 12, 13), (1, 4, 11, 12, 13), (3, 2, 11, 12, 13), (5, 10, 11, 12, 13), (7, 8, 11, 12, 13), (9, 6, 11, 12, 13), (1, 10, 11, 12, 13), (3, 8, 11, 12, 13), (5, 6, 11, 12, 13), (7, 4, 11, 12, 13), (9, 2, 11, 12, 13), (1, 6, 11, 12, 13), (3, 4, 11, 12, 13), (5, 2, 11, 12, 13), (7, 10, 11, 12, 13), (9, 8, 11, 12, 13), (1, 3, 5, 7, 9), (2, 4, 6, 8, 10), (13, 12, 11, 2, 1), (13, 12, 11, 10, 3), (13, 12, 11, 8, 5), (13, 12, 11, 6, 7), (13, 12, 11, 4, 9), (13, 12, 11, 8, 1), (13, 12, 11, 6, 3), (13, 12, 11, 4, 5), (13, 12, 11, 2, 7), (13, 12, 11, 10, 9), (13, 12, 11, 4,1), (13, 12, 11, 2, 3), (13, 12, 11, 10, 5), (13, 12, 11, 8, 7), (13, 12, 11, 6, 9), (13, 12, 11, 10, 1), (13, 12, 11, 8, 3), (13, 12, 11, 6, 5), (13, 12, 11, 4, 7), (13, 12, 11, 2, 9), (13, 12, 11, 6, 1), (13, 12, 11, 4, 3), (13, 12, 11, 2, 5), (13, 12, 11, 10, 7), (13, 12, 11, 8, 9), (9, 7, 5, 3, 1) and (10, 8, 6, 4, 2).

The normalized orthogonal polynomial of degree 1 for the design is

$$\mathbf{F} = \begin{bmatrix} \frac{-2}{\sqrt{10}} & \frac{-1}{\sqrt{10}} & 0 & \frac{1}{\sqrt{10}} & \frac{2}{\sqrt{10}} \end{bmatrix}' = \begin{bmatrix} -0.63 & -0.32 & 0 & 0.32 & 0.63 \end{bmatrix}'.$$

The information matrix obtained for this design

$$\mathbf{C} = \frac{2}{5} \begin{bmatrix} 25\mathbf{I}_{10} - \mathbf{1}_{10}\mathbf{1}'_{10} & -5\mathbf{1}_{10}\mathbf{1}'_{3} \\ -5\mathbf{1}_{3}\mathbf{1}'_{10} & 125\mathbf{I}_{3} - 25\mathbf{1}_{3}\mathbf{1}'_{3} \end{bmatrix}.$$

The variance of any estimated elementary contrast among the treatments belonging to the first set is $V_{11} = 0.200\sigma^2$ and the variance of any estimated elementary contrast between the treatments belonging to the first and second set is $V_{12} = 0.120\sigma^2$.

3.2. Method 2

Consider a BIB design with parameters v^*, b^*, r^*, k^* and λ^*. From each block of this design, develop $(k^* - 1)$ more blocks by rotating the treatments clockwise resulting into b^*k^* blocks. Substitute the last u (u = 2, 3,...,v^*−2) set of treatments of the design with the last treatment of the second set, the second last set of treatments with the second last treatment of the second set, likewise v^*−3 number of treatments can be replaced by p number of treatment of the second set. The resulting design is a TR-GEBBPB with parameters $v_1 = (v^* - pu)$, $v_2 = p$, $b = k^*b^*$, $\mathbf{r}' = [k^*r^*\mathbf{1}'_{v_1} \;\; 2k^*r^*\mathbf{1}'_{v_2}]$, $k = k^*$, $\lambda^*_{11} = k^*\lambda^*$, $\lambda^*_{12} = 2k^*\lambda^*$ and $\lambda^*_{22} = 5k^*\lambda^*$.

The information matrix for this design is

$$\mathbf{C} = \lambda^* \begin{bmatrix} v^*\mathbf{I}_{v_1} - \mathbf{1}_{v_1}\mathbf{1}'_{v_1} & -u\mathbf{1}_{v_1}\mathbf{1}'_{v_2} \\ -u\mathbf{1}_{v_2}\mathbf{1}'_{v_1} & uv^*\mathbf{I}_{v_2} - u^2\mathbf{1}_{v_2}\mathbf{1}'_{v_2} \end{bmatrix},$$

with $V_{11} = \dfrac{2}{v^*\lambda^*}\sigma^2$ and $V_{12} = \dfrac{(u+1)}{uv^*\lambda^*}\sigma^2$.

Example 3.2.1. Let $v^* = 9$,$b^* = 12$, $r^* = 4$, $k^* = 3$ and $\lambda^* = 1$ be the parameters of a BIB design with blocks as (1, 2, 3), (4, 5, 6), (7, 8, 9), (1, 4, 7), (2, 5, 8), (3, 6, 9), (1, 6, 8), (2, 4, 9), (3, 5, 7), (1, 5, 9), (2, 6, 7) and (3, 4, 8). From each block of this design, develop two more blocks by rotating the treatments clockwise resulting into 36 blocks. Let $p = 2$ and $u = 3$, substitute the last two treatments of the design with the last treatment of the second set, second last two treatments with second last treatment of the second set, i.e., substitute treatments (8, 9) by treatment number 5 and treatments (6, 7) by treatment number 4. The resulting design is TR-GEBBPB with parameters $v_1 = 3$, $v_2 = 2$, $b = 36$, $\mathbf{r}' = [12\mathbf{1}_3' \;\; 36\mathbf{1}_2']$, $k = 3$, $\lambda_{11}^* = 3$, $\lambda_{12}^* = 6$ and $\lambda_{22}^* = 15$. The blocks of the design are: (1, 2, 3), (2, 3, 1), (3, 1, 2), (4, 5, 4), (5, 4, 4), (4, 4, 5), (4, 5, 5), (5, 5, 4), (5, 4, 5), (1, 4, 4), (4, 4, 1), (4, 1, 4), (2, 5, 5), (5, 5, 2), (5, 2, 5), (3, 4, 5), (4, 5, 3), (5, 3, 4), (1, 4, 5), (4, 5, 1), (5, 1, 4), (2, 4, 5), (4, 5, 2), (5, 2, 4), (3, 5, 4), (5, 4, 3), (4, 3, 5), (1, 5, 5), (5, 5, 1), (5, 1, 5), (2, 4, 4), (4, 4, 2), (4, 2, 4), (3, 4, 5), (4, 5, 3) and (5, 3, 4).

For the above design, the normalized orthogonal polynomial of degree 1 is given as

$$\mathbf{F} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}' = \begin{bmatrix} -0.71 & 0 & 0.71 \end{bmatrix}'.$$

The information matrix for this design is given as

$$\mathbf{C} = \begin{bmatrix} 9\mathbf{I}_3 - \mathbf{1}_3\mathbf{1}_3' & -3\mathbf{1}_3\mathbf{1}_2' \\ -3\mathbf{1}_2\mathbf{1}_3' & 27\mathbf{I}_2 - 9\mathbf{1}_2\mathbf{1}_2' \end{bmatrix},$$

with $V_{11} = 0.2222\sigma^2$ and $V_{12} = 0.1481\sigma^2$.

3.3. Method 3

Consider a BIB design with parameters $v^* = sm+1$ (prime or prime power), $b^* = sv^*$, $r^* = sm$, $k^* = m$ and $\lambda^* = m-1$ obtained by developing following initial block(s) modulo v:

$$x^w, x^{w+s}, x^{w+2s}, ..., x^{(m-1)s}, \qquad \text{for} \quad w = 0, 1, ..., s-1,$$

where x is the primitive element of GF (v^*). Substitute the last u set of treatments of the design with the last treatment of the second set, second last u set of treatment with second last treatment of the second set, likewise, v^*-3 number of treatments can be replaced by p number of treatment of the second set. The resulting design is a TR-GEBBPB design with parameters $v_1 = (v^* - pu)$, $v_2 = p$, $b = sv^*$, $r_1 = sm$, $r_2 = usm$, $k = m$, $\lambda_{11}^* = \lambda^*$, $\lambda_{12}^* = 2\lambda^*$ and $\lambda_{22}^* = 4\lambda^*$.

The joint information matrix for this design is given as

$$\mathbf{C} = \frac{(k-1)}{k} \begin{bmatrix} v^*\mathbf{I}_{v_1} - \mathbf{1}_{v_1}\mathbf{1}_{v_1}' & -u\mathbf{1}_{v_1}\mathbf{1}_{v_2}' \\ -u\mathbf{1}_{v_2}\mathbf{1}_{v_1}' & u(v^*\mathbf{I}_{v_2} - u\mathbf{1}_{v_2}\mathbf{1}_{v_2}') \end{bmatrix},$$

with $V_{11} = \dfrac{uk}{v^*(k-1)}\sigma^2$ and $V_{12} = \dfrac{k(u+1)}{uv^*(k-1)}\sigma^2$.

Example 3.3.1. The blocks of a TR-GEBBPB design with parameters $v_1 = 3$, $v_2 = 2$, $b = 7$, $r_1 = 6$, $r_2 = 12$, $k = 6$, $\lambda_{11}^* = 5$, $\lambda_{12}^* = 10$ and $\lambda_{22}^* = 20$ obtained from BIB design of parameters $v^* = 7$ ($s = 1$, $m = 6$), $b^* = 7$, $r^* = 6$, $k^* = 6$ and $\lambda^* = 6$ by taking $p = 2$ and $u = 2$ are given as: (1, 3, 2, 5, 4, 4), (2, 4, 3, 5, 4, 5), (3, 4, 4, 1, 5, 5), (4, 5, 4, 2, 5, 1), (4, 5, 5, 3, 1, 2), (5, 1, 5, 4, 2, 3) and (5, 2, 1, 4, 3, 4).

The normalized orthogonal polynomial of degree 1 for the above design is

$$\mathbf{F} = \left[\frac{-5}{\sqrt{70}}\ \frac{-3}{\sqrt{70}}\ \frac{-1}{\sqrt{70}}\ \frac{1}{\sqrt{70}}\ \frac{3}{\sqrt{70}}\ \frac{5}{\sqrt{70}}\right]' = \left[-0.60\ \ -0.36\ \ -0.12\ \ 0.12\ \ 0.36\ \ 0.60\right]'.$$

The information matrix for the above design is

$$\mathbf{C} = \frac{5}{6}\begin{bmatrix} 7\mathbf{I}_3 - \mathbf{1}_3\mathbf{1}_3' & -2\mathbf{1}_3\mathbf{1}_2' \\ -2\mathbf{1}_2\mathbf{1}_3' & 14\mathbf{I}_2 - 4\mathbf{1}_2\mathbf{1}_2' \end{bmatrix},$$

with $V_{11} = 0.3428\sigma^2$ and $V_{12} = 0.2571\sigma^2$.

4. DISCUSSION

This article attempts to study general efficiency balanced block designs for comparing treatments belonging to two disjoint sets in the presence of systematic trend. The advantage of the block designs, named TR-GEBBPB, obtained here is that these are robust against the presence of trend effects. Besides, these designs are general efficiency balanced and are more efficient for estimating the contrasts pertaining to two treatments from two different sets. As the designs are completely trend resistant, the analysis of the data generated from these designs can be carried out in the usual manner as if no trend effect is present in the model. A possible extension of the present study is to develop some methods to obtain smaller designs under the present experimental situation, for which an algorithmic approach can be an alternative. Attempts can also be made to obtain designs for comparing treatments belonging to two disjoint sets in the presence of trend under unequal block structure. The effects of repeated blocks (see for instance [28], [27], [29]) in TR-GEBBPB designs obtained through BIB designs can also be explored in selecting optimal designs for testing block effects.

ACKNOWLEDGMENTS

We are very much thankful to the editor and reviewers for their constructive criticism that has led to substantial improvement of the manuscript. The facilities provided by the Indian Council of Agricultural Research for carrying out this research work are duly acknowledged.

REFERENCES

[1] Atkinson, A.C. and Donev, A.N. (1996). Experimental design optimally balanced for trend, *Technometrics*, **38**(4), 333–341.

[2] Bechoffer, R.E. and Tamhane, A.C. (1981). Incomplete block designs for comparing treatments with a control (I): General theory, *Technometrics*, **23**, 45–47.

[3] Bhowmik, A.; Gupta, R.K.; Jaggi, S.; Varghese, E.; Harun, M.; Varghese, C. and Datta, A. (2021). On the Construction of Trend Resistant PBIB Designs, *Communications in Statistics – Simulation and Computation*, DOI: 10.1080/03610918.2021.1951763.

[4] Bhowmik, A.; Jaggi, S.; Varghese, C. and Varghese, E. (2014). Trend free block designs balanced for interference effects from neighbouring experimental units, *Journal of Combinatorics, Information and System Sciences*, **39**(1-4), 117–133.

[5] Bhowmik, A.; Jaggi, S.; Varghese, C. and Varghese, E. (2015). Trend free second order neighbour balanced block designs, *Journal of the Indian Statistical Association*, **53**(1-2), 63–78.

[6] Bradley, R.A. and Yeh, C.M. (1980). Trend free block designs: Theory, *The Annals of Statistics*, **8**(4), 883–893.

[7] Carrano, A.L.; Thorn, B.K. and Lopez, G. (2006). An integer programming approach to the construction of trend-free experimental plans on split-plot designs, *Journal of Manufacturing Systems*, **25**(1), 39–44.

[8] Clatworthy, W.H. (1973). *Tables of two-associate partially balanced designs*, National Bureau of Standards, Applied Mathematics, Series No. 63, Washington D.C.

[9] Das, M.N. and Ghosh, D.K. (1985). Balancing incomplete block designs, *Sankhya Series B*, **47**, 67–77.

[10] Gupta, V.K. and Prasad, N.S.G. (1991). On construction of general efficiency balanced block designs, *Sankhya Series B*, **53**(1), 89–96.

[11] Gupta, V.K. and Parsad, R. (2001). Block designs for comparing test treatments with control treatments – An overview, *Special issue of Statistics and Applications to felicitate the 80th Birthday of Dr. M.N. Das*, **3**(1-2), 133–146.

[12] Iqbal, I.; Tahir, M.H. and Nasir, J.A. (2011). New optimal designs for comparing test treatments with a control, *The Aligarh Journal of Statistics*, **31**, 33–51.

[13] Jacroux, M. (1990). Some optimal designs for comparing a set of test treatments with a set of controls, *Annals of Institute of Statistical Mathematics*, **42**, 173–185.

[14] Jacroux, M. (1993). On the construction of trend-resistant designs for comparing a set of test treatments with a set of controls, *Journal of the American Statistical Association*, **88**, 1398–1403.

[15] Jacroux, M. (2003). Some MV-optimal group divisible type block designs for comparing a set of test treatments with a set of standard treatments, *Statistical Planning and Inference*, **113**(2), 597–615.

[16] Jacroux, M.; Majumdar, D. and Shah, K.R. (1995). Efficient block designs in the presence of trends, *Statistica Sinica*, **5**, 605–615.

[17] Jacroux, M.; Majumdar, D. and Shah, K.R. (1997). On the determination and construction of optimal block designs in the presence of linear trend, *Journal of the American Statistical Association*, **92**, 375–382.

[18] Jaggi, S. (1992). Study on optimality of one-way heterogeneity designs for comparing two disjoint sets of treatments, *Unpublished Ph.D. Thesis*, IARI, New Delhi.

[19] Jaggi, S. (1996). A-efficient block designs with unequal block sizes for comparing two sets of treatments, *Journal of Indian Society of Agricultural Statistics*, **48**(2), 125–139.

[20] Jaggi, S. and Gupta, V.K. (1997). A-optimal block designs with unequal block sizes for comparing two disjoint sets of treatments, *Sankhya: Series B*, **59**(2), 164–180.

[21] Jaggi, S.; Gupta, V.K. and Parsad, R. (1996). A-efficient block designs for comparing two disjoint sets of treatments, *Communications in Statistics – Theory and Methods*, **25**(5), 967–983.

[22] Jaggi, S.; Parsad, R. and Gupta, V.K. (1997). General efficiency balanced block designs with unequal block sizes for comparing two sets of treatments, *Journal of Indian Society of Agricultural Statistics*, **50**(1), 37–46.

[23] Jaggi, S.; Parsad, R. and Gupta, V.K. (1999). Construction of non-proper balanced bipartite block designs, *Calcutta Statistical Association Bulletin*, **49**, 55–63.

[24] Kageyama, S. and Mukerjee, R. (1986). General balanced designs through reinforcement, *Sankhya: Series B*, **48**, 380–387.

[25] Majumdar, D. (1986). Optimal designs for comparisons between two sets of treatments, *Journal of Statistical Planning and Inference*, **14**, 359–372.

[26] Majumdar, D. and Martin, R.J. (2002). Finding optimal designs in the presence of trends, *Journal of Statistical Planning and Inference*, **106**, 177–190.

[27] Oliveira, T.A. (2002). BIB Designs with Repeated Blocks: Review and perspectives, *Proceedings of ICCS-X Conference*, Tenth Islamic Countries Conference on Statistical Sciences – Statistics for Development and Good Governance. Editors Zeinab Amin and Ali, S., Hadi, The American University in Cairo, **1**, 82–96.

[28] Oliveira, T.A.; Ceranka, B. and Graczyk, M. (2006). The variance of the difference of block effects in the balanced incomplete block designs with repeated blocks, *Colloquium Biometryczne*, **36**, 115–124.

[29] Oliveira, T.A. and Oliveira, A. (2012). Ineffectiveness of the FIM in selecting Optimal BIB Designs for testing block effects, *Proceedings of COMPSTAT*.

[30] Patel, B.L. (2018). Efficiency balanced designs, *Journal of Harmonized Research in Applied Sciences*, **6**(3), 217–221.

[31] Patil, S.G.; Majumder, A. and Manjunatha, G.R. (2016). General efficiency balanced designs in circular blocks with correlated observations, *Journal of Indian Society of Agricultural Statistics*, **70**(2), 159–166.

[32] Prasad, N.S.G. (1989). Some investigations on general efficiency balanced designs, *Unpublished Ph.D. Thesis*, IARI, New Delhi.

[33] Pravender and Patel, B.L. (2016). Construction of Balanced Bipartite Block Designs, *Research Journal of Recent Sciences*, **5**(5), 41–46.

[34] Srivastav, S.K. and Srivastav, A. (2021). Efficiency balanced designs for bootstrap simulations, *Communications in Statistics – Theory and Methods*, **50**(19), 4510–4527.

[35] Yeh, C.M. and Bradley, R.A. (1983). Trend free block designs, existence and construction results, *Communications in Statistics – Theory and Methods*, **12**, 1–24.

[36] Yeh, C.M.; Bradley, R.A. and Netz, W.I. (1985). Nearly trend free block designs, *Journal of American Statistical Association*, **80**, 985–992.

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 321–342
https://doi.org/10.57805/revstat.v22i3.507

Assessing Influence on Partially Varying-Coefficient Generalized Linear Model

Authors: Germán Ibacache-Pulgar ✉
– Institute of Statistics, Universidad de Valparaíso, Chile
– Center for Atmospheric Studies and Climate Change (CEACC), Universidad de Valparaíso, Chile
german.ibacache@uv.cl

Valeria Lira
– Institute of Statistics, Universidad de Valparaíso, Chile
valeria.s.lira@gmail.com

Cristian Villegas
– Exact Science Department, University of São Paulo, São Paulo, Brazil
clobos@usp.br

Received: August 2021 Revised: July 2022 Accepted: July 2022

Abstract:

- In this paper we discuss estimation and diagnostic procedures in partially varying-coefficient generalized linear models based in the penalized likelihood function. Specifically, we derive a weighted back-fitting algorithm to estimate the model parameters using smoothing spline. Moreover, we developed the local influence method to assess the sensitivity of maximum penalized likelihood estimators when small perturbations are introduced into the model or data. Finally, an example with real data of ozone concentration is given for illustration.

Keywords:

- *exponential family; maximum penalized likelihood estimators; likelihood displacement; semiparametric models; weighted back-fitting algorithm.*

AMS Subject Classification:

- 49A05, 78B26.

✉ Corresponding author

1. INTRODUCTION

Partially varying-coefficient generalized linear model (PVCGLM) is an extension of generalized linear model (GLM), and have received special attention in recent years. These models have the same characteristics as GLM (see, for instance, McCullagh and Nelder, 1989 [27]), in the sense of encompassing different families of distributions for the response variable, allowing for non-linear dependence between the mean of the response variable and the explanatory variables (linear predictor) through a link function, and allowing for non-constant variance in the data. In addition, PVCGLM have the flexibility to model explanatory variables effects that can contribute parametrically and explanatory variables effects in which the coefficients are allowed to vary as smooth functions of other variables (for example, time variable). The model is a very useful tool for exploring dynamic patterns in some scientific areas, such as environmental, epidemiology, medical science, ecology and so on; see Fan and Zhang (2008) [9], Finley (2011) [14], Ma *et al.* (2011) [26], Li *et al.* (2018) [24], and He *et al.* (2022) [18].

As was noted by some authors (see, for example, Ouwens *et al.*, 2001 [29]), GLM parameter estimators can be higly impacted by outlying observations. For this reason, diagnostic analysis is of fundamental importance in the statistical modelling of any data set. The main idea of the local influence technique, introduced by Cook (1986) [5], is to evaluate the sensitivity of parameter estimators when small perturbations are introduced in the assumptions of the model or in the data. Some of the works related to the technique of local influence applied to different regression models are the following. Thomas and Cook (1989) [33] extended the method of local influence proposed by Cook to generalized linear models, with the purpose to asses the effect of small perturbations in the data. Ouwens *et al.* (2001) [29] developed local influence to detect influential data structures under a generalized linear mixed model; specifically, they proposed a two-stage diagnostic procedure, the first to measure the influence of the subjects and the second to measure the influence of the observations. Zhu and Lee (2001) [35] extended the method of local influence for incomplete data based on the conditional expectation of the complete-data log-likelihood function, and applied the results to the generalized linear mixed model; see also Zhu and Lee (2003) [36]. Espinheira *et al.* (2008) [8] developed the local influence method for beta regressions model under different perturbation schemes. Rocha and Simas (2011) [31] extended the local influence method to a general formulation of the class of the beta regression models, whereas Ferrari *et al.* (2011) [12] derived the normal curvatures of local influence for beta regression models with varying dispersion. Ferreira and Paula (2016) [13] extended the local influence technique for different perturbation schemes considering a skew-normal partially linear model and Emami (2016) [6] applied local influence analysis to the Liu penalized least squares estimator.

In semiparametric context, Thomas (1991) [33] constructed local influence diagnostics to evaluate the sensitivity of the smoothing parameter estimate obtained by cross-validation criterion. Zhu *et al.* (2003) [36] and Ibacache-Pulgar and Paula (2011) [21] provide local influence measures to evaluate the sensitivity of the maximum penalized likelihood estimator in normal and Student-t partially linear models, respectively. Ibacache-Pulgar *et al.* (2012, 2013) [19, 20] derived the local influence curvature for elliptical semiparametric mixed and symmetric semiparametric additive models, respectively. Zhang *et al.* (2015) [34] and Ibacache-Pulgar and Reyes (2018) [22] developed local influence measures for normal and elliptical partially varying-coefficient models, respectively. Recently, Ibacache-Pulgar *et al.* (2021) [23] developed the local influence method to semiparametric additive beta regression models and Sanchez *et al.* (2021) [32] derived the normal curvature for a new quantile regression model.

The aim of this paper is to apply local influence to the PVCGLM. The paper is organized as follows. In Section 2, the PVCGLM is presented. A discussion on the process used to obtain the maximum likelihood (ML) estimator based on the penalized likelihood, the derivation of a back-fitting algorithm and some inferential result are given in Section 3. In Section 4 the main concepts of local influence are considered and normal curvatures for different perturbations schemes are derived. An illustration of the methodology is presented in Section 5. Finally, in Section 6, some concluding remarks are given.

2. STATISTICAL MODEL

In this section we present the PVCGLM and the penalized log-likelihood function used to carry out parameter estimation.

2.1. Formulation

Consider a data set that is composed of a response variable y_i, for $i \in \{1, ..., n\}$, that follows a distribution in the exponential family with density function

$$f_y(y_i; \theta_i, \phi) = \exp\left[\frac{y_i\theta_i - \psi(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right], \tag{2.1}$$

where θ_i is the canonical form of the location parameter and is a function of the mean μ_i, $a_i(\phi)$ is a known function of the unknown dispersion parameter ϕ (or a vector of unknown dispersion parameters), c is a function of the dispersion parameter and the responses, and ψ is a known function, such that the mean and variance of y_i are equals to $\mu_i = \mathrm{E}(y_i) = \partial\psi(\theta_i)/\partial\theta_i$ and $\mathrm{Var}(y_i) = a_i(\phi)\,\mathrm{V}_i$, with $\mathrm{V}_i = \mathrm{V}(\mu_i) = \partial^2\psi(\theta_i)/\partial\theta_i^2$, respectively. The PVCGLM is defined by Equation (2.2) and the following systematic component:

$$g(\mu_i) = \eta_i = \boldsymbol{w}_i^\top \boldsymbol{\alpha} + \sum_{k=1}^{s} \mathrm{x}_i^{(k)}\, \beta_k(\mathrm{t}_{k_i}), \tag{2.2}$$

where $\boldsymbol{w}_i$ is a $(p \times 1)$ vector of predictors variables, $\boldsymbol{\alpha} = (\alpha_1, ..., \alpha_p)^\top$ is a vector of regression coefficients, $\beta_k(\cdot)$ for $k \in \{1, ..., s\}$ are unknown smooth arbitrary functions of t_k, associated with the predictor variable $\mathrm{x}_i^{(k)}$. Here, the superscript k refers to the relationship of the predictor variable x_i with the k-th nonparametric component. Note that Model (2.2) can be written in a matrix form as

$$\boldsymbol{\eta} = \boldsymbol{W}\boldsymbol{\alpha} + \sum_{k=1}^{s} \widetilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k\ , \tag{2.3}$$

where $\boldsymbol{W} = (\boldsymbol{w}_1^\top, ..., \boldsymbol{w}_n^\top)$, $\widetilde{\boldsymbol{N}}_k = \boldsymbol{X}^{(k)}\boldsymbol{N}_k$, $\boldsymbol{X}^{(k)} = \mathrm{diag}_{1\le i\le n}(\mathrm{x}_i^{(k)})$, $\boldsymbol{N}_k$ is an $(n \times r_k)$ incidence matrix with the (i, l)-th element equal to the indicator $I(\mathrm{t}_{k_i} = \mathrm{t}^0_{k_l})$ with $\mathrm{t}^0_{k_l}$ denoting the distinct and ordered values of the explanatory variable t_k, and $\boldsymbol{\beta}_k = (\psi_{k_1}, ..., \psi_{k_r})^\top$ is a $(r_k \times 1)$ vector of parameters with $\psi_{k_l} = \beta_k(\mathrm{t}^0_{k_l})$ for $l \in \{1, ..., r_k\}$.

2.2. Penalized log-likelihood function

Let $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\beta}_1^\top, ..., \boldsymbol{\beta}_s^\top, \phi)^\top \in \boldsymbol{\Theta} \subseteq \mathcal{R}^{p^*}$, with $p^* = p + r + 1$ and $r = \sum_{k=1}^{s} r_k$, be the vector of unknown parameters associated to Model (2.1). Then, the log-likelihood function is given by

$$L(\boldsymbol{\theta}) = \sum_{i=1}^{n} L_i(\boldsymbol{\theta}), \tag{2.4}$$

where

$$L_i(\boldsymbol{\theta}) = \left[\frac{y_i\theta_i - \psi(\theta_i)}{a_i(\phi)} + c(y_i, \phi)\right]. \tag{2.5}$$

Since the β_k's belong to a space of infinite dimension and are considered parameters with respect to the expected value of y_i, it is necessary to define a restricted subspace for these functions so that the identifiability of the parameters holds. This choice typically depends on the domain of the function, on a priori knowledge of form of the function, on constraints to ensure identifiability, or simply on some specific application. In this paper, we will assume that the function β_k belongs to the Sobolev function space

$$\mathcal{W}_2^{(l)} = \left\{\beta_k : \beta_k, \beta_k^{(1)}, ..., \beta_k^{(l-1)} \text{abs. cont.}, \beta_k^{(l)} \in \mathcal{L}^2[a_k, b_k]\right\},$$

where $\beta_k^{(l)}(\mathrm{t}_k) = \mathrm{d}^l\beta(\mathrm{t}_k)/\mathrm{dt}_k^l$, with $\mathrm{t}^0_k \in [a_k, b_k]$. To ensure the identifiability of the parameters and an adequate fit of the model, we incorporate a penalty term in the original log-likelihood function over each function β_k. In this way, we obtain a penalized version of the log-likelihood function of the form (see details in Green and Silverman, 1994 [15])

$$L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) = L(\boldsymbol{\theta}) - \sum_{k=1}^{s} \frac{\lambda_k}{2}\, \boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k\,, \tag{2.6}$$

where $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_s)^\top$ denotes a $(s \times 1)$ vector of smoothing parameters that controls the tradeoff between goodness of fit and the smoothness estimated functions, and $\boldsymbol{K}_k$ is a $(r_k \times r_k)$ nonnegative definite smoothing matrix associated with the k-th explanatory variable that only depends on the knots. In this case, the estimation of β_k leads to a smooth cubic spline with knots at the points $\mathrm{t}^0_{k_l}$ for $l \in \{1, ..., r_k\}$.

3. ESTIMATION AND INFERENCE

In this section we outlying the estimation of the parameters of the PVCGLM. Specifically, we propose an iterative process based on the Fisher score and back-fitting algorithms to estimate the regression coefficients and the nonparametric functions, and respective standard errors being obtained from the penalized Fisher information matrix. More details about estimation procedure can be found, for example, in Hastie and Tibshirani (1993) [16], Cai *et al.* (2000) [4], Fang and Huang (2005) [10] and Rigby and Stasinopoulos (2005) [30].

3.1. Weighted maximum penalized likelihood estimator

Assuming that the function (2.6) is regular with respect to $\boldsymbol{\alpha}$, $\boldsymbol{\beta}_k$'s and ϕ, the penalized score function vector of $\boldsymbol{\theta}$ is given by

$$\mathbf{U}_{\mathrm{P}}(\boldsymbol{\theta}) = \frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\theta}}\,.$$

After some algebraic manipulations (see, for instance, Liu *et al.*, 2021 [25], for details of the calculation of derivatives of matrix or vectors), we obtain the following:

$$\begin{aligned}
\frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\alpha}} &= \boldsymbol{W}^{\top}\boldsymbol{T}(\boldsymbol{y}-\boldsymbol{\mu})\,,\\
\frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\beta}_k} &= \widetilde{\boldsymbol{N}}_k^{\top}\boldsymbol{T}(\boldsymbol{y}-\boldsymbol{\mu}) - \lambda_k\boldsymbol{K}_k\boldsymbol{\beta}_k \qquad k\in\{1,...,s\}\,,\\
\frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \phi} &= \sum_{i=1}^{n} -(\mathrm{a}_i(\phi))^{-2}\{y_i\theta_i-\psi(\theta_i)\} + \sum_{i=1}^{n} c'(y_i,\phi),
\end{aligned}$$

where $\boldsymbol{W}$ is a $(n\times p)$ matrix whose i-th row is $\boldsymbol{w}_i^{\top}$, $\boldsymbol{T}=\mathrm{diag}_{1\le i\le n}\big((\mathrm{a}_i(\phi))^{-1}(\partial\mu_i/\partial\eta_i)V_i^{-1}\big)$ with $V_i=V(\mu_i)=\partial^2\psi(\theta_i)/\partial\theta_i^2$ the variance function, $\mathrm{a}_i(\phi)$ is a function of ϕ, $\boldsymbol{y}=(y_1,...,y_n)^{\top}$, $\boldsymbol{\mu}=(\mu_1,...,\mu_n)^{\top}$ and $c'(y_i,\phi)=\partial c(y_i,\phi)/\partial\phi$. To estimate $\boldsymbol{\theta}$, we have to solve $\mathbf{U}_{\mathrm{P}}(\boldsymbol{\theta})=\mathbf{0}$. However, the estimating equations are nonlinear and require an iterative method. For example, maximum penalized likelihood (MPL) estimator for $\boldsymbol{\theta}$ can be performed by using the Fisher scoring algorithm. Let $\boldsymbol{\beta}_0=\boldsymbol{\alpha}$, $\widetilde{\boldsymbol{N}}_0=\boldsymbol{W}$, and $\boldsymbol{\lambda}$ fixed. Then, the Fisher scoring algorithm is given by

$$\begin{pmatrix} \boldsymbol{I} & \boldsymbol{S}_0^{(u)}\widetilde{\boldsymbol{N}}_1 & \dots & \boldsymbol{S}_0^{(u)}\widetilde{\boldsymbol{N}}_s\\ \boldsymbol{S}_1^{(u)}\widetilde{\boldsymbol{N}}_0 & \boldsymbol{I} & \dots & \boldsymbol{S}_1^{(u)}\widetilde{\boldsymbol{N}}_s\\ \vdots & \vdots & \ddots & \vdots\\ \boldsymbol{S}_s^{(u)}\widetilde{\boldsymbol{N}}_0 & \boldsymbol{S}_s^{(u)}\widetilde{\boldsymbol{N}}_1 & \dots & \boldsymbol{I}\end{pmatrix}\begin{pmatrix}\boldsymbol{\beta}_0^{(u+1)}\\ \boldsymbol{\beta}_1^{(u+1)}\\ \vdots\\ \boldsymbol{\beta}_s^{(u+1)}\end{pmatrix} = \begin{pmatrix}\boldsymbol{S}_0^{(u)}\boldsymbol{z}^{(u)}\\ \boldsymbol{S}_1^{(u)}\boldsymbol{z}^{(u)}\\ \vdots\\ \boldsymbol{S}_s^{(u)}\boldsymbol{z}^{(u)}\end{pmatrix}, \tag{3.1}$$

where $\boldsymbol{z}^{(u)}=(\boldsymbol{y}-\boldsymbol{\mu}^{(u)})+\left(\sum_{k=0}^{s}\widetilde{\boldsymbol{N}}_k\boldsymbol{\beta}_{\boldsymbol{k}}{}^{(u)}\right)$ and

$$\boldsymbol{S}_k^{(u)} = \begin{cases} (\widetilde{\boldsymbol{N}}_0^{\top}\boldsymbol{M}^{(u)}\widetilde{\boldsymbol{N}}_0)^{-1}\widetilde{\boldsymbol{N}}_0^{\top}\boldsymbol{M}^{(u)} & k=0\\ (\widetilde{\boldsymbol{N}}_k^{\top}\boldsymbol{M}^{(u)}\widetilde{\boldsymbol{N}}_k+\lambda_k\boldsymbol{K}_k)^{-1}\widetilde{\boldsymbol{N}}_k^{\top}\boldsymbol{M}^{(u)} & k\in\{1,...,s\}\,,\end{cases}$$

where $\boldsymbol{M} = \mathrm{diag}_{1\le i\le n}\big((\mathrm{a}_i(\phi))^{-1}(\partial\mu_i/\partial\eta_i)^2 V_i^{-1}\big)$. Consequently, the weighted back-fitting (Gauss-Seidel) iterations that are used to solve the equations system (3.1) take the form

$$\boldsymbol{\beta}_k^{(u+1)} = \boldsymbol{S}_k^{(u)}\Bigg(\boldsymbol{z}^{(u)} - \sum_{l=0,l\neq k}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}\Bigg), \tag{3.2}$$

for $u \in \{0,1,...\}$. On the other hand, the MPL estimator of the dispersion parameter, $\widehat{\phi}$, can be obtained by solving the following iterative process:

$$\phi^{(u+1)} = \phi^{(u)} - \mathrm{E}\left\{\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\phi^2}\right\}^{-1} \frac{\partial L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\phi}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(u)}},$$

for $u \in \{0,1,...\}$.

Algorithm 1 – Joint iterative process for estimating the parameters of the PVCGLM.

(i) Initialize:

- **(a)** Provide values for $\boldsymbol{\beta}_0^{(0)}, \boldsymbol{\beta}_1^{(0)}, ..., \boldsymbol{\beta}_s^{(0)}$.
- **(b)** Get starting value for ϕ by using the fitted values from (a).
- **(c)** From the current value $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\beta}_0^{(0)\top}, \boldsymbol{\beta}_1^{(0)\top}, ..., \boldsymbol{\beta}_s^{(0)\top}, \phi^{(0)})^\top$ obtaining the weight matrix $\boldsymbol{M}^{(0)}$. Then, obtain

$$\begin{aligned}
\boldsymbol{z}^{(0)} &= (\boldsymbol{y} - \boldsymbol{\mu}^{(0)}) + \Bigg(\sum_{k=0}^{s} \widetilde{\boldsymbol{N}}_k \boldsymbol{\beta}_k^{(0)}\Bigg),\\
\boldsymbol{S}_0^{(0)} &= (\widetilde{\boldsymbol{N}}_0^\top \boldsymbol{M}^{(0)} \widetilde{\boldsymbol{N}}_0)^{-1} \boldsymbol{N}_0^\top \boldsymbol{M}^{(0)},\\
\boldsymbol{S}_k^{(0)} &= (\widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M}^{(0)} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M}^{(0)}, \qquad k \in \{1,...,s\}.
\end{aligned}$$

(ii) Step 1: Iterate repeatedly by cycling between the following equations:

$$\begin{aligned}
\boldsymbol{\beta}_0^{(u+1)} &= \boldsymbol{S}_0^{(u)}\Bigg(\boldsymbol{z}^{(u)} - \sum_{l=1}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}\Bigg),\\
\boldsymbol{\beta}_1^{(u+1)} &= \boldsymbol{S}_1^{(u)}\Bigg(\boldsymbol{z}^{(u)} - \widetilde{\boldsymbol{N}}_0 \boldsymbol{\beta}_0^{(u+1)} - \sum_{l=2}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}\Bigg),\\
&\vdots\\
\boldsymbol{\beta}_s^{(u+1)} &= \boldsymbol{S}_s^{(u)}\Bigg(\boldsymbol{z}^{(u)} - \sum_{l=0}^{s-1} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u+1)}\Bigg),
\end{aligned}$$

for $u \in \{0,1,...\}$. Repeat (ii) replacing $\boldsymbol{\beta}_\jmath^{(u)}$ by $\boldsymbol{\beta}_\jmath^{(u+1)}$ until convergence criterion $\Delta_u(\boldsymbol{\beta}_\jmath^{(u+1)}, \boldsymbol{\beta}_\jmath^{(u)}) = \sum_{\jmath=0}^{s} \| \boldsymbol{\beta}_\jmath^{(u+1)} - \boldsymbol{\beta}_\jmath^{(u)} \| / \sum_{\jmath=0}^{s} \| \boldsymbol{\beta}_\jmath^{(u)} \|$ is below some small threshold (Hastie and Tibshirani, 1990 [17]).

(iii) Step 2: For current values $\boldsymbol{\beta}_\jmath^{(u+1)}$ for $\jmath \in \{0,1...,s\}$, obtaining $\phi^{(u+1)}$ by using

$$\phi^{(u+1)} = \phi^{(u)} - \mathrm{E}\left\{\frac{\partial^2 L_{\mathrm{p}}(\phi,\boldsymbol{\lambda})}{\partial\phi\partial\phi}\right\}^{-1} \frac{\partial L_{\mathrm{p}}(\phi,\boldsymbol{\lambda})}{\partial\phi}\Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}^{(u)}}.$$

(iv) Iterating between steps (ii) and (iii) by replacing $\boldsymbol{\beta}_\jmath^{(0)}$ for $\jmath \in \{0,1...,s\}$ and $\phi^{(0)}$ by $\boldsymbol{\beta}_\jmath^{(u+1)}$ and $\phi^{(u+1)}$, respectively, until convergence.

Note that the system of equations (3.1) is consistent and the back-fitting algorithm (3.2) converges to a solution for any starting values if the weights matrix involved is symmetric and positive definite. Additionally, the solution is unique when there is not concurvity in the data, that is, nonlinear dependencies among the predictor variables. However, in the presence of concurvity, the starting functions will determine the final solution, while in presence exact concurvity is highly unlikely, except in the case of symmetric smoothers with eigenvalues in [0,1]; see, for instance, Berhane and Tibshirani (1998) [1]. The summary, the solution of the estimating equation system (3.1) to obtain the MPL estimates of $\boldsymbol{\theta}$ may be attained by iterating between a weighted back-fitting algorithm with weight matrix $\boldsymbol{M}$ and a Fisher score algorithm to obtain ML estimation of ϕ, which is equivalent to the iterative process in Algorithm 1.

3.2. Standard error of MPL estimator

Similarly to the classical theory of generalized linear models, the variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$ can be approximated through the inverse of Fisher information matrix obtained from penalized log-likelihood function, $L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})$. Assuming that the penalized log-likelihood function (2.6) is twice differentiable with respect to $\boldsymbol{\theta}$, we have that the penalized Fisher information matrix is given by

$$\boldsymbol{\mathcal{I}}_{\mathrm{p}} = -\mathrm{E}\left(\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\right).$$

This matrix assumes the following diagonal structure in blocks:

$$\boldsymbol{\mathcal{I}}_{\mathrm{p}}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{\mathcal{I}}_{\mathrm{p}}^{\alpha\beta_k}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\mathcal{I}}_{\mathrm{p}}^{\phi\phi}(\boldsymbol{\theta}) \end{pmatrix},$$

where

$$\boldsymbol{\mathcal{I}}_{\mathrm{p}}^{\alpha\beta_k}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{W}^\top\boldsymbol{M}\boldsymbol{W} & \boldsymbol{W}^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_1 & \ldots & \boldsymbol{W}^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_s \\ \widetilde{\boldsymbol{N}}_1^\top\boldsymbol{M}\boldsymbol{W} & \widetilde{\boldsymbol{N}}_1^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_1 + \lambda_1\boldsymbol{K}_1 & \ldots & \widetilde{\boldsymbol{N}}_1^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_s \\ \vdots & \vdots & \ddots & \vdots \\ \widetilde{\boldsymbol{N}}_s^\top\boldsymbol{M}\boldsymbol{W} & \widetilde{\boldsymbol{N}}_s^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_1 & \ldots & \widetilde{\boldsymbol{N}}_s^\top\boldsymbol{M}\widetilde{\boldsymbol{N}}_s + \lambda_s\boldsymbol{K}_s \end{pmatrix}$$

and

$$\boldsymbol{\mathcal{I}}_{\mathrm{p}}^{\phi\phi}(\boldsymbol{\theta}) = \sum_{i=1}^{n} -2(\mathrm{a}_i(\phi))^{-3}(\mu_i\theta_i - \psi(\theta_i)) - \sum_{i=1}^{n}\mathrm{E}(c''(y_i,\phi)),$$

with $c''(y_i,\phi) = \partial^2 c(y_i,\phi)/\partial\phi^2$ for $i \in \{1,\ldots,n\}$. Therefore, the approximate variance--covariance matrix of $\widehat{\boldsymbol{\theta}}$ and an approximate pointwise standard error band (SEB) for $\beta_k(\cdot)$, that allows us to assess the accuracy of $\widehat{\beta_k}(\cdot)$ at different locations within the range of interest, are given by

$$\widehat{\mathrm{Cov}}(\widehat{\boldsymbol{\theta}}) \approx \boldsymbol{\mathcal{I}}_{\mathrm{p}}^{-1}\big|_{\widehat{\boldsymbol{\theta}}},$$

$$\mathrm{SEB}_{\mathrm{approx}}(\beta_k(\mathrm{t}_l^0)) = \widehat{\beta}_k(\mathrm{t}_l^0) \pm 2\sqrt{\widehat{\mathrm{Var}}(\widehat{\beta}_k(\mathrm{t}_l^0))} \qquad l \in \{1,\ldots,r_k\},$$

where $\mathrm{Var}(\widehat{\beta}_k(\mathrm{t}_l))$, for $k \in \{1,\ldots,s\}$, is the l-th principal diagonal element of the corresponding block-diagonal matrix of $\boldsymbol{\mathcal{I}}_{\mathrm{p}}^{-1}$.

3.3. Effective degrees of freedom and smoothing parameters

In the iterative process defined in the Equation (3.2), considering ϕ as known, we can write the expression of the estimator of $\boldsymbol{\beta}_k$ at step u as

$$\boldsymbol{\beta}_k^{(u+1)} = (\widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \boldsymbol{M} \boldsymbol{z}^{*(u)} \qquad k \in \{1, ..., s\}, \tag{3.3}$$

where $\boldsymbol{z}^{*(u)} = \boldsymbol{z}^{(u)} - \sum_{l=0, l\neq k}^{s} \widetilde{\boldsymbol{N}}_l \boldsymbol{\beta}_l^{(u)}$. From the convergence of the iterative process given in the Equation (3.3), we obtain

$$\widehat{\boldsymbol{\beta}}_k = (\widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}} \widehat{\boldsymbol{z}^*} \qquad k \in \{1, ..., s\},$$

where $\widehat{\boldsymbol{z}^*} = \Big[(\boldsymbol{y} - \widehat{\boldsymbol{\mu}}) + \big(\sum_{k=0}^{s} \widetilde{\boldsymbol{N}}_k \widehat{\boldsymbol{\beta}}_k\big)\Big] - \sum_{l=0, l\neq k}^{s} \widetilde{\boldsymbol{N}}_l \widehat{\boldsymbol{\beta}}_l$. In this paper we define the effective degrees of freedom (df) associated with the smooth functions as (see, for instance, Hastie and Tibshirani, 1990 [17])

$$\mathrm{edf}(\lambda_k) = \mathrm{tr}\{\widetilde{\boldsymbol{N}}_k (\widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_k + \lambda_k \boldsymbol{K}_k)^{-1} \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{M}}\}.$$

Following Ibacache-Pulgar and Reyes (2018) [22], we choose the optimal smoothing parameter for each smooth functions by specifying an appropriate $\mathrm{edf}(\lambda_k)$ value.

4. LOCAL INFLUENCE

In this section we obtain the normal curvature for PVCGLM. Specifically, the Hessian and perturbations matrices for different perturbations schemes.

4.1. The method

To assess the influence of minor perturbations on the MPL estimator of $\boldsymbol{\theta}$, $\widehat{\boldsymbol{\theta}}$, we can consider the likelihood displacement $\mathrm{LD}(\boldsymbol{\omega}) = 2\Big[L_{\mathrm{P}}(\widehat{\boldsymbol{\theta}}, \boldsymbol{\lambda}) - L_{\mathrm{P}}(\widehat{\boldsymbol{\theta}}_\omega, \boldsymbol{\lambda})\Big] \geq 0$, where $\widehat{\boldsymbol{\theta}}_\omega$ is the MPL estimador under the perturbed penalized log-likelihood function, denoted by $L_{\mathrm{P}}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega})$, and $\boldsymbol{\omega} = (\omega_1, ..., \omega_n)^\top$ be an n-dimensional vector of perturbations restricted to some open subset $\boldsymbol{\Omega} \in \mathcal{R}^n$. It is assumed that there exists $\boldsymbol{\omega}_0 \in \boldsymbol{\Omega}$, a vector of no perturbation, such that $L_{\mathrm{P}}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega}_0) = L_{\mathrm{P}}(\boldsymbol{\theta}, \boldsymbol{\lambda})$. Cook (1986) [5] suggests to study the local behavior of $\mathrm{LD}(\boldsymbol{\omega})$ around $\boldsymbol{\omega}_0$ selecting a unit direction $\boldsymbol{\ell} \in \Omega$ ($\|\boldsymbol{\ell}\| = 1$), and then to consider the plot of $\mathrm{LD}(\boldsymbol{\omega}_0 + a\,\boldsymbol{\ell})$ (called lifted line) against a, where $a \in \mathcal{R}$. Each lifted line can be characterized by considering the normal curvature $C_\ell(\boldsymbol{\theta})$ around $a = 0$. The suggestion is to consider the direction $\boldsymbol{\ell} = \boldsymbol{\ell}_{\max}$ corresponding to the largest curvature $C_{\ell_{\max}}(\boldsymbol{\theta})$. The index plot of $\boldsymbol{\ell}_{\max}$ may reveal those observations that under small perturbations exercise notable influence on $\mathrm{LD}(\boldsymbol{\omega})$. According to Cook (1986) [5], the normal curvature at the unit direction $\boldsymbol{\ell}$ is given by $C_\ell(\boldsymbol{\theta}) = -2\big[\boldsymbol{\ell}^\top \boldsymbol{\Delta}_{\mathrm{p}}^\top \boldsymbol{L}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}} \boldsymbol{\ell}\big]$, which represents the local influence on $\widehat{\boldsymbol{\theta}}$ after perturbing the model or data, where $\boldsymbol{L}_{\mathrm{p}}$ is the Hessian matrix evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\Delta}_{\mathrm{p}} = \partial^2 L_{\mathrm{P}}(\boldsymbol{\theta}, \boldsymbol{\lambda} \,|\, \boldsymbol{\omega}) / \partial\boldsymbol{\theta}\partial\boldsymbol{\omega}^\top$ is the perturbation matrix evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$.

Escobar and Meeker (1992) [7] proposed to study the normal curvature at the direction $\boldsymbol{\ell} = \boldsymbol{\epsilon}_i$, where $\boldsymbol{\epsilon}_i$ is an n-dimensional vector with 1 at the i-th position and zeros at the remaining positions. In this case, the normal curvature, called total local influence of the i-th individual, takes the form $C_{\boldsymbol{\epsilon}_i}(\boldsymbol{\theta}) = 2|\,\mathrm{c}_{ii}\,|$ for $i \in \{1, ..., n\}$, where c_{ii} is the i-th principal diagonal element of the matrix $\boldsymbol{C} = \boldsymbol{\Delta}_{\mathrm{p}}^{\top} \boldsymbol{L}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}}$. In order to have a invariant curvature under uniform change of scale, Poon and Poon (1999) [28] proposed the conformal normal curvature defined as

$$B_{\ell}(\boldsymbol{\theta}) = \frac{C_{\ell}(\boldsymbol{\theta})}{2\sqrt{\mathrm{tr}(\boldsymbol{\Delta}_{\mathrm{p}}^{\top} \boldsymbol{L}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}})^2}} = -\frac{\boldsymbol{\ell}^{\top} \boldsymbol{\Delta}_{\mathrm{p}}^{\top} \boldsymbol{L}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}} \boldsymbol{\ell}}{\sqrt{\mathrm{tr}(\boldsymbol{\Delta}_{\mathrm{p}}^{\top} \boldsymbol{L}_{\mathrm{p}}^{-1} \boldsymbol{\Delta}_{\mathrm{p}})^2}} \,.$$

This curvature is characterized to allow for any unit direction $\boldsymbol{\ell}$ that $0 \leq B_{\ell}(\theta) \leq 1$. A suggestion is to consider the direction $\boldsymbol{\ell} = \boldsymbol{\ell}_{\max}$ corresponding to the largest curvature $B_{\ell_{\max}}(\boldsymbol{\theta})$ or, alternatively, to evaluate the normal curvature at the direction $\boldsymbol{\ell} = \boldsymbol{\epsilon}_i$ and analyse the index plot of $B_{\boldsymbol{\epsilon}_i}(\boldsymbol{\theta})$.

4.2. Derivation of normal curvature

The perturbation schemes that are considered in the analysis of local influence depend on the structure of the proposed model (see, for instance, Billor and Loynes, 1993 [2]), and can be classified into two broad groups: perturbation to the model (in order to study modifications in the assumptions) or in the data. For example, we might be interested in perturbing the response or the explanatory variables. The reasons for considering such perturbation schemes are, for example, the existence of outliers or measures with measurement errors, respectively. However, the perturbation scheme to be considered should be formulated in a way that responds to questions previously established by the researcher. We will present in what follows expressions of the $\boldsymbol{L}_{\mathrm{p}}$ and $\boldsymbol{\Delta}_{\mathrm{p}}$ matrices for some perturbations schemes.

Hessian matrix

Let $\boldsymbol{L}_{\mathrm{p}}$ $(p^* \times p^*)$ be the Hessian matrix with (j^*, ℓ^*)-th element given by $\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda}) / \partial\theta_{j^*}\theta_{\ell^*}$ for $j^*, \ell^* \in \{1, ..., p^*\}$, where $p^* = p + r + 1$, with $r = \sum_{k=1}^{s} r_k$. After some algebraic manipulations we find

$$\begin{aligned}
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{\alpha}^{\top}} &= -\boldsymbol{W}^{\top} \widetilde{\boldsymbol{M}} \boldsymbol{W}\,, \\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \boldsymbol{\beta}_k \partial \boldsymbol{\beta}_k^{\top}} &= \begin{cases} -\widetilde{\boldsymbol{N}}_k^{\top} \widetilde{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_k - \lambda_k \boldsymbol{K}_k & k = k' \\ -\widetilde{\boldsymbol{N}}_k^{\top} \widetilde{\boldsymbol{M}} \widetilde{\boldsymbol{N}}_{k'} & k \neq k'\,, \end{cases} \\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta}, \boldsymbol{\lambda})}{\partial \phi^2} &= \sum_{i=1}^{n} 2(\mathrm{a}_i(\phi))^{-3}(y_i\theta_i - \psi(\theta_i)) + \sum_{i=1}^{n} c''(y_i, \phi))\,,
\end{aligned}$$

$$\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \boldsymbol{\alpha}\partial \boldsymbol{\beta}_k^\top} = -\boldsymbol{W}^\top \widetilde{\boldsymbol{M}}\widetilde{\boldsymbol{N}}_k\,,$$

$$\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \alpha_j\partial \phi} = -\sum_{i=1}^{n}({\rm a}_i(\phi))^{-2}\left\{(y_i-\mu_i)V_i^{-1}\frac{\partial\mu_i}{\partial\eta_i}\,{\rm w}_i\right\},$$

$$\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda})}{\partial \psi_{k_l}\partial \phi} = -\sum_{i=1}^{n}({\rm a}_i(\phi))^{-2}\left\{(y_i-\mu_i)V_i^{-1}\frac{\partial\mu_i}{\partial\eta_i}\,{\rm n}_{k_{il}}\right\},$$

where $c''(y_i,\phi)=\partial^2 c(y_i,\phi)/\partial\phi^2$, $\widetilde{\boldsymbol{M}}=\mathrm{diag}_{1\le i\le n}\big(({\rm a}_i(\phi))^{-1}(\partial\mu_i/\partial\eta_i)^2V_i^{-1}\rho_i\big)$, $\rho_i=\kappa(\mu_i)/\{g'(\mu_i)^2V_i\}$, with $\kappa(\mu_i)=1+(y_i-\mu_i)\{V_i'/V_i+g''(\mu_i)/g'(\mu_i)\}$ and $g'(\mu_i)=\mathrm{d}\eta_i/\mathrm{d}\mu_i$, and ${\rm n}_{k_{li}}$ denotes the (i,l)-th element of the matrix $\boldsymbol{N}_k$.

Cases-weight perturbation

Let us consider the attributed weights for the observations in the penalized log-likelihood function as

$$L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})=L(\boldsymbol{\theta}|\boldsymbol{\omega})-\sum_{k=1}^{s}\frac{\lambda_k}{2}\,\boldsymbol{\beta}_k^\top\boldsymbol{K}_k\boldsymbol{\beta}_k\,,$$

where $L(\boldsymbol{\theta}|\boldsymbol{\omega})=\sum_{i=1}^n\omega_i\,L_i(\boldsymbol{\theta})$, $\boldsymbol{\omega}=(\omega_1,...,\omega_n)^\top$ is the vector of weights, with $0\le\omega_i\le 1$, and $\boldsymbol{\omega}_0=(1,...,1)^\top$ the vector of no perturbation. Differentiating $L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and $\boldsymbol{\omega}$, we obtain after some algebraic manipulation

$$\left.\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \boldsymbol{W}^\top\widehat{\boldsymbol{D}}_\tau\,,$$

$$\left.\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\beta}_k\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \widetilde{\boldsymbol{N}}_k^\top\widehat{\boldsymbol{D}}_\tau \qquad k\in\{1,...,s\}\,,$$

$$\left.\frac{\partial^2 L_{\rm p}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\phi\partial\boldsymbol{\omega}^\top}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} = \widehat{\boldsymbol{u}}^\top\,,$$

where $\boldsymbol{D}_\tau=\mathrm{diag}_{1\le i\le n}(\tau_i)$ and $\boldsymbol{u}=(u_1,...,u_n)^\top$, with $\tau_i=({\rm a}_i(\phi))^{-1}(y_i-\partial\psi(h(\eta_i))/\partial h(\eta_i))\cdot$ $\cdot\,\partial h(\eta_i)/\partial\eta_i$, $h(\eta_i)=\psi'^{-1}(\eta_i)$, $\psi'^{-1}(\cdot)$ denoting the inverse function of $\psi'(\cdot)$, $u_i=-({\rm a}_i(\phi))^{-2}\cdot$ $\cdot\,(y_ih(\eta_i)-\psi(h(\eta_i))+{\rm c}'(y_i,\phi)\boldsymbol{e}_{in}^\top$, and $\boldsymbol{e}_{in}$ a vector with 1 at the i-th position and zero elsewhere.

Response variable perturbation

In general, the response variable can be perturbed in two ways:

$$y_{i\omega}=\begin{cases} y_i+\ \omega_i & \text{additive perturbation} \qquad i\in\{1,...,n\}\\ y_i\times\ \omega_i & \text{multiplicative perturbation}\,.\end{cases}$$

In this paper we consider $y_{i\omega}=y_i+\omega_i$, where $\boldsymbol{\omega}=(\omega_1,...,\omega_n)^\top$ is the vector of perturbations and $\boldsymbol{\omega}_0=(0,...,0)^\top$ the vector of no perturbation. The perturbed penalized log-likelihood function is constructed from expression (2.6) with y_i replaced by $y_{i\omega}$, that is,

$$L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega}) = L(\boldsymbol{\theta}|\boldsymbol{\omega}) - \sum_{k=1}^{s}\frac{\lambda_k}{2}\,\boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k\ ,$$

where $L(\cdot)$ is given by Equation (2.4) with $y_{i\omega}$ in the place of y_i. Differentiating $L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and ω_i we obtain, after some algebraic manipulation, that

$$\begin{aligned}
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \boldsymbol{W}^\top \widehat{\boldsymbol{D}}_c\ ,\\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\beta}_k\partial\boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{D}}_c \qquad k\in\{1,\dots,s\}\ ,\\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\phi\partial\boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \widehat{\boldsymbol{d}}^\top\ ,
\end{aligned}$$

where $\boldsymbol{D}_c = \mathrm{diag}_{1\le i\le n}(c_i)$ and $\boldsymbol{d} = (d_1,\dots,d_n)^\top$, with $c_i = \partial h(\eta_i)/\partial\eta_i$ and $d_i = -(\mathrm{a}_i(\phi))^{-2}\cdot$ $\cdot\,(h(\eta_i)\boldsymbol{e}_{in}^\top + c'(y_{i\omega},\phi)/\partial\omega_i)$, with $\boldsymbol{e}_{in}$ denoting a vector with 1 at the i-th position and zero elsewhere..

Explanatory variable perturbation

The explanatory variable can be perturbed in two ways:

$$\mathrm{w}_{i\omega} = \begin{cases}\mathrm{w}_{i\omega}+\omega_i & \text{additive perturbation} \qquad\qquad i\in\{1,\dots,n\}\\ \mathrm{w}_{i\omega}\times\omega_i & \text{multiplicative perturbation}\ .\end{cases}$$

Here the d-th explanatory variable, assumed continuous, is perturbed by considering the additive perturbation scheme, namely $\mathrm{w}_{id\omega} = \mathrm{w}_{id} + \omega_i$, where $\boldsymbol{\omega} = (\omega_1,\dots,\omega_n)^\top$ is the vector of perturbations such as $\omega_i\in\mathcal{R}$. The vector of no perturbation is given by $\boldsymbol{\omega_0} = (0,\dots,0)^\top$. The perturbed penalized log-likelihood function is given by

$$L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega}) = L(\boldsymbol{\theta}|\boldsymbol{\omega}) - \sum_{k=1}^{s}\frac{\lambda_k}{2}\,\boldsymbol{\beta}_k^\top \boldsymbol{K}_k \boldsymbol{\beta}_k\ ,$$

where $L(\cdot)$ is given by Equation (2.4) with $\mu_{i\omega} = g^{-1}(\eta_{i\omega})$ in the place of μ_i, for $\eta_{i\omega} = \boldsymbol{w}_{i\omega}^\top\boldsymbol{\alpha} + \sum_{k=1}^{s}\mathrm{x}_i^{(k)}\,\beta_k(\mathrm{t}_{k_i})$, with w_{id} replaced by $\mathrm{w}_{id\omega}$. Differentiating $L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})$ with respect to the elements of $\boldsymbol{\theta}$ and ω_i we obtain

$$\begin{aligned}
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \boldsymbol{e}_p\widehat{\boldsymbol{\tau}}^\top - \alpha_d\boldsymbol{W}^\top \widehat{\boldsymbol{D}}_b\ ,\\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\boldsymbol{\beta}_k\partial\boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= \boldsymbol{e}_p\widehat{\boldsymbol{\tau}}^\top - \alpha_d\widetilde{\boldsymbol{N}}_k^\top \widehat{\boldsymbol{D}}_b \qquad k\in\{1,\dots,s\}\ ,\\
\frac{\partial^2 L_{\mathrm{p}}(\boldsymbol{\theta},\boldsymbol{\lambda}|\boldsymbol{\omega})}{\partial\phi\partial\boldsymbol{\omega}^\top}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\,\boldsymbol{\omega}=\boldsymbol{\omega}_0} &= -\sum_{i=1}^{n}(\mathrm{a}_i(\phi))^{-2}\left\{y_i\frac{\partial(\eta_{i\omega})}{\partial\omega_i} - \frac{\partial\psi(h(\eta_{i\omega}))}{\partial\omega_i}\right\}\boldsymbol{e}_{in}^\top,
\end{aligned}$$

where $\boldsymbol{\tau} = (\tau_1,\dots,\tau_n)^\top$, $\boldsymbol{D}_b = \mathrm{diag}_{1\le i\le n}(b_i)$ and $\boldsymbol{e}_p$ is a vector with 1 at the p-th position and zero elsewhere, $\tau_i = (\mathrm{a}_i(\phi))^{-1}(y_i - \partial\psi(h(\eta_{i\omega}))/\partial h(\eta_{i\omega}))\partial h(\eta_{i\omega})/\partial\eta_{i\omega}$ and $b_i = (\mathrm{a}_i(\phi))^{-1}\cdot$ $\cdot\,(y_i - \partial\psi(h(\eta_{i\omega}))/\partial h(\eta_{i\omega}))\partial^2 h(\eta_{i\omega})/\partial\eta_{i\omega}^2 - (\partial^2\psi(h(\eta_{i\omega}))/\partial^2 h(\eta_{i\omega}))(\partial h(\eta_{i\omega})/\partial\eta_{i\omega})^2$.

5. APPLICATION

In this section, we illustrate the applicability of the PVCGLM and the local influence method through an application based on a set of real data. For our analysis, we consider the Poisson distribution.

5.1. Data set and problem statement

To motivate the use of the PVCGLM and the local influence method developed in this work, we consider a set of real data from a study conducted in the city of Los Angeles during 1976 (see, for instance, Breiman and Friedman, 1985 [3] and Faraway, 2006 [11]) with the purpose of describing the relationship between the outcome variable O3 (concentration of ozone per hour in Upland, CA, measured in parts per million (ppm) and a set of nine explanatory variables, for a sample of 330 days. The description of such variables is as follows. VH (pressure height 500 millibar, measured at the base of the air force of Vandenberg, in meters), WIND (wind speed, in miles per hour), HUM (humidity in percentage), TEMP (sandburg Air Base temperature, in Celsius), IBH (inversion base height, in foot), DPG (dagget pressure gradient, in mmHg), IBT (inversion base temperature, in Fahrenheit), VIS (visibility, in miles), DAY (calendar day).

5.2. Model fit

In our application we will consider only four explanatory variables, specifically, the variables VIS, TEMP, IBT and DAY. Figure 1 contains the dispersion graphs between the outcome variable and each one of the explanatory variables.

We see in Figure 1a that the relationship between O3 and the explanatory variable VIS is approximately linear, whereas the relationship between O3 and DAY appear to be nonlinear (Figure 1b). Note that there is a significant increase in the level of O3 from January to July with a decrease until December. This suggests that the incorporation of a quadratic or nonparametric term in the model can account for the behavior of O3 over time. On other hand, Figures 1c and 1d suggest that the explanatory variables TEMP and IBT might be interacting with the variable DAY in nonlinear fashion. Figures 2a and 2b shows the graph of autocorrelation and partial autocorrelation, respectively. Following the same analysis of Faraway (2006) [11], in this work we will no consider the possible temporal correlation for O3.

Initially, we will fit a GLM assuming that the response variable O3 follows a Poisson distribution with mean μ_i and logarithmic link function considering different structures of the linear predictor for the explanatory variables VIS, TEMP, IBT and DAY (see Table 1).

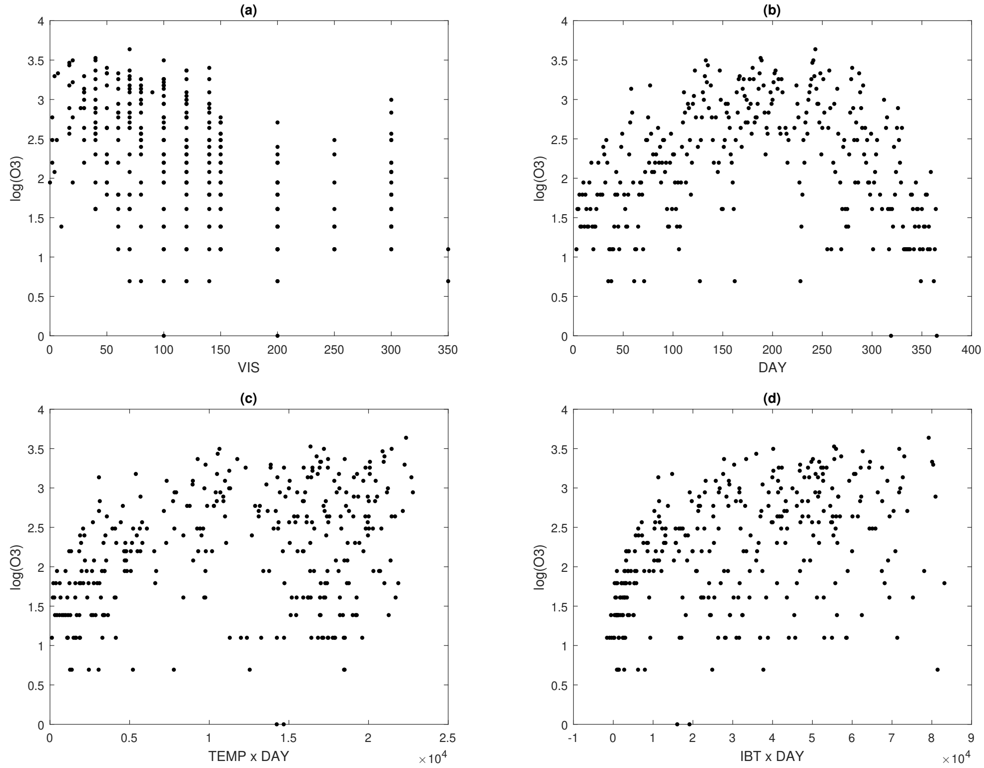


Figure 1: Scatter plots: log(O3) versus VIS (a), log(O3) versus DAY (b), log(O3) versus TEMP × DAY (c) and log(O3) versus IBT × DAY (d).

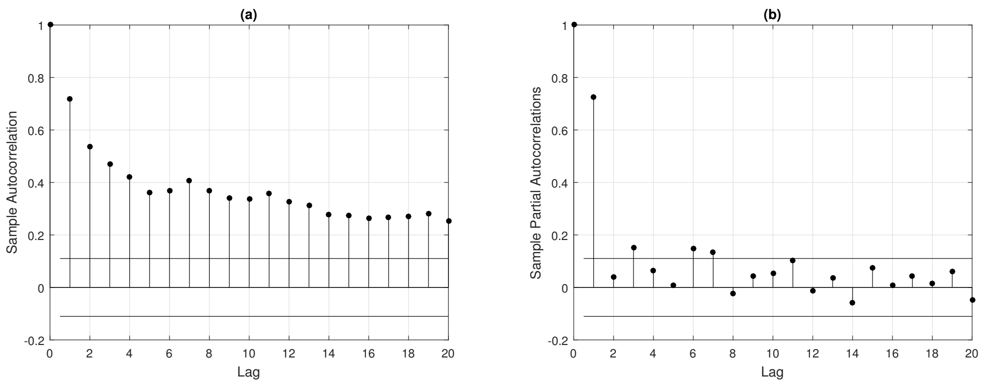


Figure 2: Autocorrelation (a) and partial autocorrelation (b) for Ozone data.

Table 1: Different structures of the linear predictor for the explanatory variables VIS, TEMP, IBT and DAY assuming that the response variable O3 $\sim$ Poisson(μ_i).

Model	Systematic component $g(\mu_i) = \log(\mu_i)$
I	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i$
II	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i + \alpha_4 \mathrm{DAY}_i$
III	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i + f(\mathrm{DAY}_i)$
IV	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \alpha_2 \mathrm{TEMP}_i + \alpha_3 \mathrm{IBT}_i + \alpha_4 \mathrm{DAY}_i + \alpha_5 \mathrm{TEMP}_i \times \mathrm{DAY}_i + \alpha_6 \mathrm{IBT}_i \times \mathrm{DAY}_i$
V	$\alpha_0 + \alpha_1 \mathrm{VIS}_i + \mathrm{TEMP}_i \beta_1(\mathrm{DAY}_i) + \mathrm{IBT}_i \beta_2(\mathrm{DAY}_i)$

For Model I, only the individual effect of the VIS, TEMP and IBT explanatory variables were considered. In Model II, the individual effects of these three covariates plus the effect of the DAY variable were incorporated in a linear manner, whereas in the Model III the individual effect of the DAY explanatory variable is included nonlinearly by using a smooth function. Model IV considers the individual contributions of VIS, TEMP, IBT and DAY explanatory variables, plus the interaction effects of the TEMP and IBT explanatory variables with the DAY variable. Finally, Model V corresponds to a PVCGLM where the explanatory variables TEMP and IBT interact with the variable DAY in nonlinear fashion. Table 2 contains the ML and MPL estimates associated with the parametric component for the five fitted models; the respective standard errors appear in parentheses.

Table 2: AIC, R^2, ML and MPL estimates (standard error) for all five fitted models to the Ozone data.

Parameters	Model				
	I	II	III	IV	V
α_0	0.65 (0.11)	0.79 (0.11)	1.09 (0.16)	0.88 (0.21)	1.18 (0.17)
α_1	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)	−0.00 (0.00)
α_2	0.02 (0.00)	0.03 (0.00)	0.01 (0.00)	0.02 (0.00)	—
α_3	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	—
α_4	—	−0.001 (0.00)	—	−0.001 (0.00)	—
α_5	—	—	—	−0.00 (0.00)	—
α_6	—	—	—	0.00 (0.00)	—
AIC	1890.71	1861.27	1752.56	1863.66	1735.76
R^2	0.682	0.691	0.752	0.690	0.754

It should be noted that the p-values (omitted here) associated with the parameters of each fitted model are smaller than 0.05, thus indicating the contributions of the individual and interaction effects are statistically significant. Note also that the parameter estimates (associated with the parametric component) obtained from the different fitted models are quite similar and accurate. The last two rows of the Table 2 shows the Akaike Information Criterion (AIC) and R^2 values, respectively. It is clear that the PVCGLM, for which the $\text{AIC}(\lambda_1, \lambda_2) = 1735.76$, presents the best fit to the Ozone data, followed by Model III with an $\text{AIC} = 1752.56$, which is confirmed by the QQ-plots presented in Figure 3; see, specifically, Figures 3(c) and 3(e). Note also that the R^2 associated with our model is higher than Models I, II and IV, and slightly higher that Model III.

For the PVCGLM the estimates of the smoothing parameters λ_1 and λ_2 as well as the corresponding df's were obtained by the procedure proposed by Ibacache-Pulgar *et al.* (2013) [20], and are presented in Table 3. Figures 4(a) and 4(b) show the estimated smooth functions under PVCGLM and the corresponding approximate SEB (dashed curves).

Table 3: Fit summary for smoothing components under PVCGLM fitted to Ozone data.

	Smooth function	
	β_1(DAY)	β_2(DAY)
df(λ_k)	6.894	7.228
λ_k	89050.050	5886.339

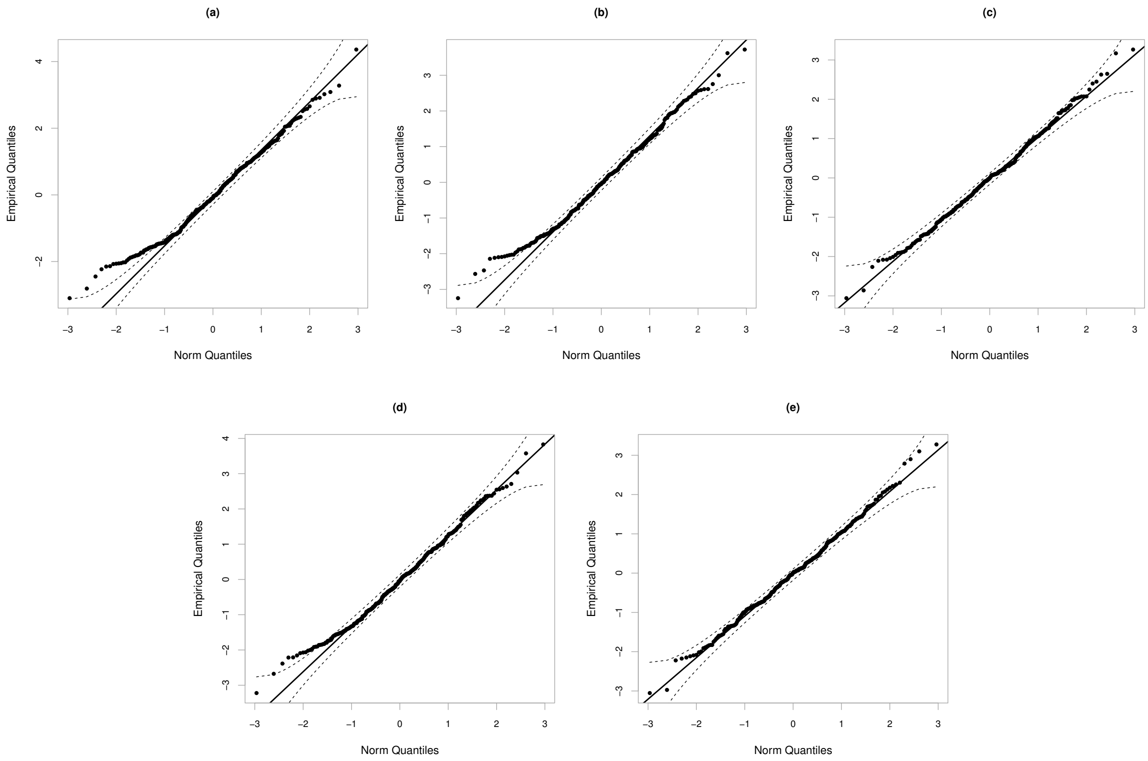


Figure 3: Normal probability plots to Ozone data: Model I (a), Model II (b), Model III (c), Model IV (d) and Model V (e).

Note that the plots confirm the nonlinear trends of the interaction effects between (TEMP, DAY) and (IBT, DAY).

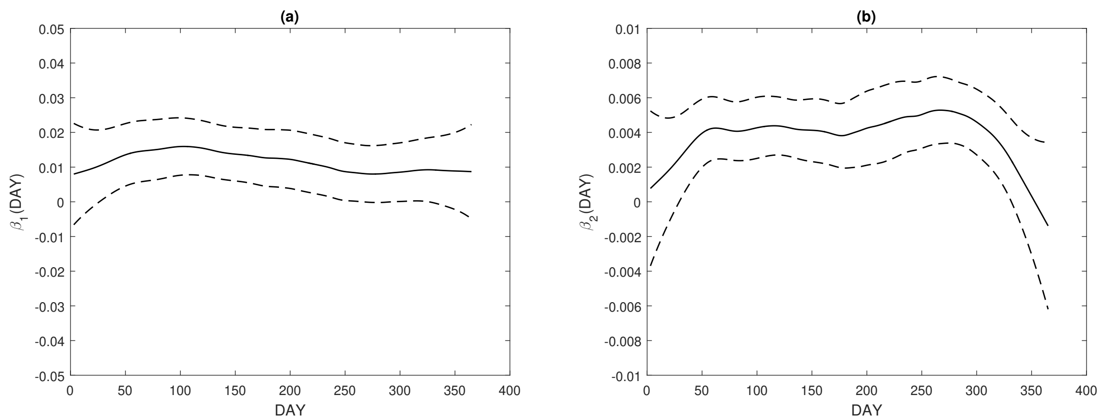


Figure 4: Plots of the estimated smooth functions β_1 (a) and β_2 (b) and their approximate pointwise SEB denoted by the dashed lines, Ozone data.

5.3. Local influence analysis

As mentioned earlier, the measure LD($\boldsymbol{\omega}$) is useful for assessing the distance between $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_\omega$. In order to identify influential potentially observations on MPL estimators under the fitted PVCGLM model to Ozone data, we present some index plots of $B_i = B_{\boldsymbol{e}_i}(\boldsymbol{\gamma})$, for $\boldsymbol{\gamma} = \boldsymbol{\alpha}, \boldsymbol{\beta}_k$ and $k \in \{1, 2\}$.

Case-weight perturbation

Figure 5 shows the index plot B_i for the case-weight perturbation scheme under the fitted model. Note at Figure 5, that observations #125, #219, #167 and #258 are more influential for the MPL estimator $\widehat{\boldsymbol{\alpha}}$, whereas observations #219, #221 and #222 are influential for the MPL estimator $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$, respectively. When we introduce an additive perturbation to the response variable, the results are analogous to those observed under the case-weight perturbation scheme, and therefore the graphs are omitted.

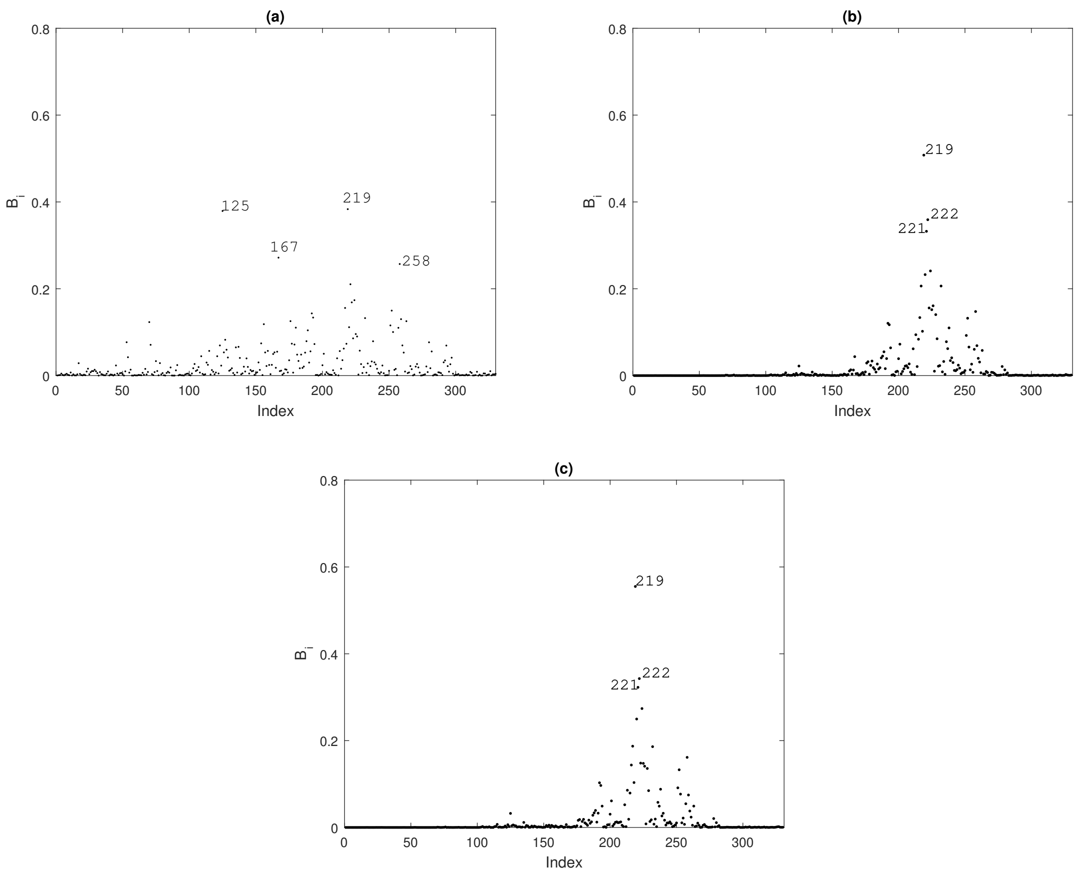


Figure 5: Index plots of B_i for assessing local influence on $\widehat{\boldsymbol{\alpha}}$ (a), $\widehat{\boldsymbol{\beta}}_1$ (b) and $\widehat{\boldsymbol{\beta}}_2$ (c) considering case-weight perturbation, Ozone data.

Explanatory variable additive perturbation

By perturbing the explanatory variable in an additive way, it becomes clear that s observations #125, #219 and #167 are more influential for the MPL estimator $\widehat{\boldsymbol{\alpha}}$, whereas observations #219, #221 and #222 are influential for the MPL estimator $\widehat{\boldsymbol{\beta}}_1$ and $\widehat{\boldsymbol{\beta}}_2$, respectively; see Figure 6.

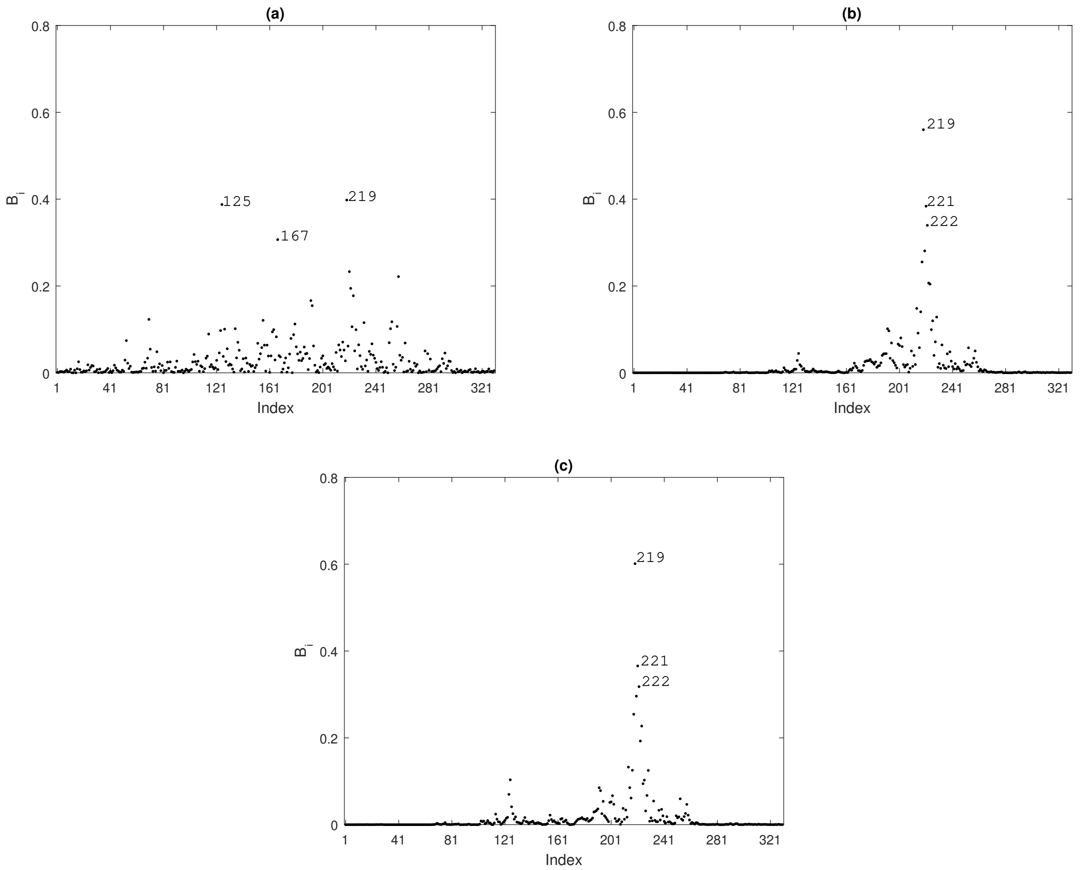


Figure 6: Index plots of B_i for assessing local influence on $\widehat{\boldsymbol{\alpha}}$ (a), $\widehat{\boldsymbol{\beta}}_1$ (b) and $\widehat{\boldsymbol{\beta}}_2$ (c) considering explanatory variable perturbation, Ozone data.

Based on the local influence analysis, we conclude that the MPL estimators of the regression coefficient and of the smooth functions are sensitive to perturbations introduced into the data or to the model. In addition, this analysis revealed that the observations that were detected as influential for the parametric component are not necessarily influential for the nonparametric component, and vice versa. For instance, under the case-weight perturbation scheme, observations #125, #219, #167 and #258 were detected as influential for the parametric component. However, of these three observations, only observation #219 is indicated as influential for the nonparametric component, in addition to observations #221 and #222. In general, similar results were obtained when the explanatory variable is additively perturbed.

5.4. Confirmatory analysis

In order to investigate the impact on the model inference when the observations detected as potentially influential in the diagnostic analysis are removed, we present the relative changes (RCs) in the MPL estimate of α_j for $j \in \{1,2\}$ after removing from the data set the influential potentially observations (%). The RC is defined as $\mathrm{RC}_\xi = |(\widehat{\xi} - \widehat{\xi}_{(\mathrm{I})})/\widehat{\xi}| \times 100\%$, where $\widehat{\xi}_{(\mathrm{I})}$ denotes the MPL estimate of ξ, with $\xi = \alpha_j$, after the corresponding observation(s) are removed according to the set I. Table 4 presents the RCs in the regression coefficient estimates after removing the observations indicated as potentially influential for the parametric component of the model.

Table 4: Relative changes (in %) in the MPL estimates of α_j under the PVCGLM.

Dropped observation	Parameters		Relative changes	
	α_0	α_1	RC_{α_0}	RC_{α_1}
125	1.17365	−0.001616	0.977	1.635
167	1.16798	−0.001592	1.455	0.125
219	1.18324	−0.001626	0.167	2.264
258	1.21007	−0.001622	2.096	2.013
125–167	1.17727	−0.001623	0.672	2.075
125–219	1.18273	−0.001628	0.211	2.389
125–258	1.52686	−0.001638	28.823	3.019
167–219	1.17701	−0.001603	0.694	0.817
167–258	1.53185	−0.001614	29.245	1.509
219–258	1.17689	−0.001609	0.703	1.195
125–167–219	1.15265	−0.001637	2.748	2.955
167–219–258	1.17625	−0.001593	0.758	0.189
125–167–219–258	1.51397	−0.001654	27.737	4.025

On the other hand, Table 5 shows the RCs observed in the estimation of the regression coefficient once the observations detected as potentially influential for the nonparametric component of the model are excluded.

Table 5: Relative changes (in %) in the MPL estimates of α_j under the PVCGLM considering the observations detected as influential on the nonparametric component.

Dropped observation	Parameters		Relative changes	
	α_0	α_1	RC_{α_0}	RC_{α_1}
none	1.18	−0.001		
219	1.183	−0.002	0.167	2.264
221	1.161	−0.002	2.041	1.132
222	1.552	−0.002	30.949	2.075
219–221	1.142	−0.002	3.641	1.635
219–222	1.521	−0.002	28.330	4.779
221–222	1.151	−0.002	2.865	1.258
219–221–222	1.186	−0.002	0.092	2.955

Considering these results, we conclude that, although some RCs are large, inferential changes are not detected. It is interesting to notice from Tables 4 and 5 the coherence with the local influence diagnostic shown previously. For instance, removal of the observations sets $\mathrm{I} = \{167, 258\}$ and $\mathrm{I} = \{125, 258\}$, which contain observations detected as influential potentially for the parametric component, leads to significant changes in the MPL estimates, mainly in $\hat{\alpha}_0$, of the order of 29.245% and 28.823%, respectively; see Table 4.

Note also that the individual removal of observation #258 produces a RC of order of 2.096%. On the other hand, the removal of the observations set $\mathrm{I} = \{219, 222\}$, whose observations were detected as influential potentially for nonparametric component, leads to significant changes in the MPL estimate of α_0, 28.330%. It is also observed that the removal of observation #222 produces a RC of 30.949%. This indicates the need of a diagnostic examination. The changes produced in the estimates of the smooth functions are presented in Figure 7.

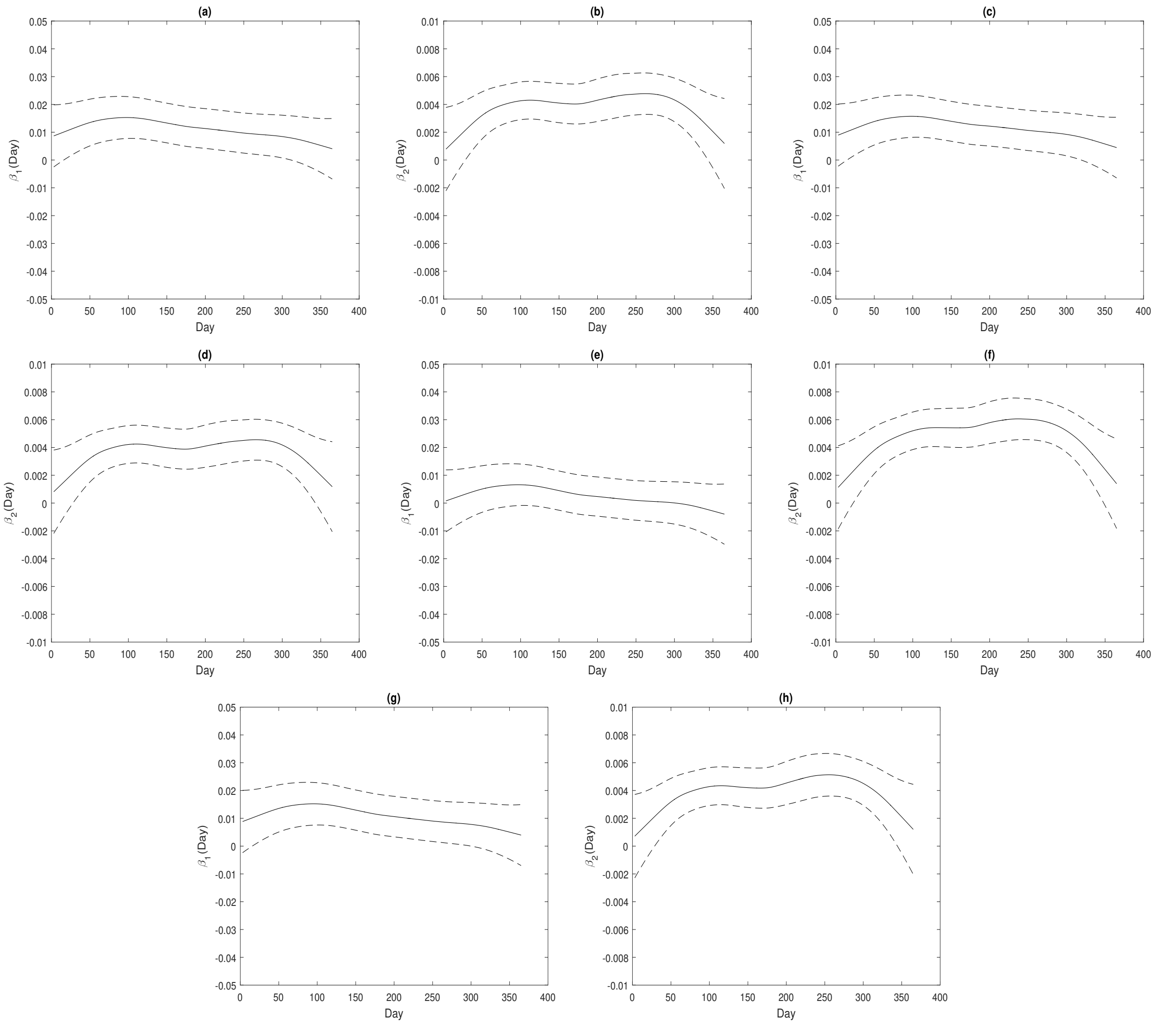


Figure 7: Plots of estimated smooth functions, $\widehat{\beta}_1$ and $\widehat{\beta}_2$, for the Ozone data and their approximate pointwise SEB denoted by the dashed lines: excluding observations #219 and #221 (a)–(b), excluding observations #219 and #222 (c)–(d), excluding observations #221 and #222 (e)–(f), excluding observations #219, #221 and #222 (g)–(h).

5.5. Computational aspects and summary of our methodology

The fitted models, quantile-quantile plots (qqplot) with simulated envelopes, and local influence were done in Matlab version R2015a and are available via email for people interested in replicating our analyses. Additionally, it is important to note that there are at least two libraries in the free R software that can be used to fit our models, for example, the mgcv library (`https://cran.r-project.org/web/packages/mgcv/index.html`) and gamlss (`https://cran.r-project.org/web/packages/gamlss/index.html`). However, there is no R library that performs local influence for the models studied in our work. Next, we summarize all the stages of our methodology through an algorithm (Algorithm 2).

Algorithm 2 – Some guidelines for applying the analysis of local influence on the PVCGLM.

1. Make a scatterplot and analyze the trend of the variables. Depending on the trend of your data you should use linear, quadratic, or polynomial function (parametric function). Alternatively, you could use a non-linear parametric form or nonparametric function (cubic spline for instance).

2. Decide if your response variable is a discrete random variable (Bernoulli, Binomial, Poisson, etc) or continuous (Normal, Gamma, etc) belonging to the exponential family. After that, decide which is the best option for your link function (log, square root, inverse, logit, probit, etc) and try different parametric, nonparametric, or semiparametric for the systematic component of a generalized linear model.

3. Choose the best model based on some criterion such as R-square or AIC.

4. Apply the local influence method and if you have some outlying observation study the relative changes deleting some observations. If you do not have outlying observations, make some conclusions about your data set.

6. CONCLUSIONS, LIMITATIONS, AND FUTURE RESEARCH

In this paper we study some aspects of the partially varying-coefficient generalized linear models. Specifically, we derive a weighted back-fitting iterative process to estimate the regression coefficients, the smooth functions and the dispersion parameter associated with our model. The variance-covariance matrix of the maximum penalized likelihood estimators was approximated by the inverse of penalized Fisher information matrix, and the effective degrees of the freedom of the nonparametric components were calculated from the estimates obtained in convergence of the iterative process. Furthermore, we extended the local influence method and obtained closed expressions for the Hessian matrix and the perturbation matrix under different perturbation schemes. We performed a statistical data analysis with a real data set on ozone concentration and some meteorological variables. The study showed the advantage of incorporating a semiparametric additive term when there are predictors whose interactions contribute nonlinearly to the model, and the utility of the local influence method to detect influential observations on the maximum penalized likelihood estimators. One of the main limitations of our model is the absence of a structure that allows modeling the correlation in those data sets that have a time component, this being one of the main lines of research to be developed. In addition, we believe that the exploration of new perturbation schemes is necessary, mainly in the interaction components and the smoothing parameter.

Finally, we recommend the use of partially varying-coefficient generalized linear models and the local influence method when the response variable belongs to the exponential family and the interactions between the explanatory variables can be modelled through smooth functions, and our interest is to evaluate the sensitivity of the maximum penalized likelihood estimator.

ACKNOWLEDGMENTS

The authors thanks Editors and Referees for their suggestions which allowed us to improve the presentation of this work. G. Ibacache-Pulgar acknowledges funding support by grant FONDECYT 11130704, Chile.

REFERENCES

[1] Berhane, K. and Tibshirani, R. (1998). Generalized additive models for longitudinal data, *Canadian Journal of Statistics*, **26**, 517–535.

[2] Billor, N. and Loynes, R. (1993). Local influence: a new approach, *Communications in Statistics – Theory and Methods*, **22**, 1595–1611.

[3] Breiman, L. and Friedman, J. (1985). Estimating optimal transformations for multiple regression and correlation, *Journal of the American Statistical Association*, **80**, 580–598.

[4] Cai, Z.; Fan, J. and Li, R. (2000). Efficient estimation and inferences for varying-coefficient models, *Journal of the American Statistical Association*, **95**, 888–902.

[5] Cook, R. (1986). Assessment of local influence (with discussion), *Journal of the Royal Statistical Society B*, **48**, 133–169.

[6] Emami, H. (2017). Local influence for Liu estimators in semiparametric linear models, *Statistical Papers*, **19**, 529–544.

[7] Escobar, L. and Meeker, W. (1992). Assessing local influence in regression analysis with censored data, *Biometrics*, **48**, 507–528.

[8] Espinheria, P.; Ferrari, S. and Cribari-Neto, F. (2008). On beta regression residuals, *Journal of Applied Statistics*, **35**, 407–419.

[9] Fan, J. and Zhang, W. (2008). Statistical methods with varying coefficient models, *Statistics and Interface*, **1**, 179–195.

[10] Fan, J. and Huang, T. (2005). Profile likelihood inferences on semiparametric varying-coefficient partially linear models, *Bernoulli*, **11**, 1031–1057.

[11] Faraway, J. (2006). *Extending the Linear Model with R Generalized Linear, Mixed Effects and Nonparametric Regression Models*, Chapman and Hall.

[12] Ferrari, S.; Espinheira, P. and Cribari-Neto, F. (2011). Diagnostics tools in beta regression with varying dispersion, *Statistica Neerlandica*, **65**, 337–351.

[13] Ferreira, C. and Paula, G. (2016). Estimation and diagnostic for skew-normal partially linear models, *Journal of Applied Statistics*, **44**, 3033–3053.

[14] Finley, A. (2011). Comparing spatially-varying coefficients models for analysis of ecological data with non-stationary and anisotropic residual dependence, *Methods in Ecology and Evolution*, **2**, 142–154.

[15] Green, P. and Silverman, B. (1994). *Nonparametric Regression and Generalized Linear Models*, Chapman and Hall, Boca Raton.

[16] Hastie, T. and Tibshirani, R. (1993). Varying-coefficient models, *Journal of the Royal Statistical Society B*, **55**, 757–796.

[17] Hastie, T. and Tibshirani, R. (1990). Generalized additive models, *Statistical Science*, **3**, 297–318.

[18] He, K.; Wang, Y.; Su, W. and Yang, H. (2022). A varying-coefficient regression approach to modelling the effects of wind speed on the dispersion of pollutants, *Environmental and Ecological Statistics*. http://doi.org/10.1007/s10651-022-00535-6

[19] Ibacache-Pulgar, G.; Paula, G. and Galea, M. (2012). Influence diagnostics for elliptical semiparametric mixed models, *Statistical Modelling*, **12**, 165–193.

[20] Ibacache, G.; Paula, G. and Cysneiros, F. (2013). Semiparametric additive models under symmetric distributions, *TEST*, **22**, 103–121.

[21] Ibacache, G. and Paula, G. (2011). Local Influence for student-t partially linear models, *Computational Statistics and Data Analysis*, **55**, 1462–1478.

[22] Ibacache-Pulgar, G. and Reyes, S. (2018). Local influence for elliptical partially varying-coefficient model, *Statistical Modelling*, **18**, 149–174.

[23] Ibacache-Pulgar, G.; Figueroa-Zuñiga, J. and Marchant, C. (2021). Semiparametric additive beta regression models: inference and local influence diagnostics, *REVSTAT – Statistical Journal*, **19**, 255–274.

[24] Li, J.; Xia, X.; Wong, W.E. and Nott, D. (2018). Varying-coefficient semiparametric model averaging predicction, *Biometrics*, **14**, 1417–1426.

[25] Liu, S.; Leiva, V.; Zhuang, D.; Ma, T. and Figueroa-Zuñiga, J.I. (2021). Matrix differential calculus with applications in the multivariate linear model and its diagnostics, *Journal of Multivariate Analysis*, pages in press available at https://doi.org/10.1016/j.jmva.2021.104849

[26] Ma, S.; Yang, L.; Romero, R. and Cui, Y. (2011). Varying coefficient model for gene-environment interacction: a non-linear look, *Bioinformatics*, **15**, 2119–2126.

[27] McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, 2nd ed., London: Chapman and Hall.

[28] Poon, W. and Poon, Y. (1999). Conformal normal curvature and assessment of local influence, *Journal of the Royal Statistical Society B*, **61**, 51–61.

[29] Ouwens, M.; Tan, F. and Berger, M. (2001). Local influence to detect influential data structures for generalized linear mixed models, *Biometrics*, **57**, 1166–1172.

[30] Rigby, R. and Stasinopoulos, D. (2005). Generalized additive models for location, scale and shape, *Appl. Statist.*, **54**, 507–554.

[31] Rocha, A. and Simas, A. (2011). Influence diagnostics in a general class of beta regression models, *TEST*, **20**, 95–119.

[32] Sanchez, L.; Leiva, V.; Helto, S.; Marchant, C. and Sarabia, J.M. (2021). A new quantile regression model and its diagnostics analytics for a weibull distributed response with applications, *Mathematics*, **9**, 2768. http://doi.org/10.3390/math9212768

[33] Thomas, W. (1991). Influence diagnostics for the cross-validated smoothing parameter in spline smoothing, *Journal of the American Statistical Association*, **9**, 693–698.

[34] Zhang, J.; Zhang, X.; Ma, H. and Zhiya, C. (2015). Local influence analysis of varying coefficient linear model, *Journal of Interdisciplinary Mathematics*, **3**, 293–306.

[35] Zhu, H. and Lee, S. (2001). Local influence for incomplete-data models, *Journal of the Royal Statistical Society B*, **63**, 111–126.

[36] Zhu, H. and Lee, S. (2003). Local influence for generalized linear mixed models, *Canadian Journal of Statistics*, **31**, 293–309.

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 343–367
https://doi.org/10.57805/revstat.v22i3.539

Estimation for Inverse Burr Distribution under Generalized Progressive Hybrid Censored Data with an Application to Wastewater Engineering Data

Authors: Hanieh Panahi ✉
– Department of Mathematics and Statistics, Lahijan Branch, Islamic Azad University, Lahijan, Iran
panahi@liau.ac.ir

Saeid Asadi
– Department of Mechanical Engineering, Payame Noor University (PNU), Tehran, Iran
s_asadi@pnu.ac.ir

Parya Parviz
– Department of Statistics, Science and Research Branch, Islamic Azad University, Tehran, Iran
pariyaparviz@gmail.com

Received: February 2021 Revised: July 2022 Accepted: July 2022

Abstract:

- The inverse Burr distribution is a significant and commonly used lifetime distribution, which plays an important role in reliability engineering. In this article, the estimation of parameters of the inverse Burr distribution based on generalized Type II progressive hybrid censored sample is studied. The expectation-maximization (EM) algorithm is employed for computing the maximum likelihood estimates of the unknown parameters. It is shown that the maximum likelihood estimates exist uniquely. The asymptotic confidence intervals for the parameters are constructed using the missing value principle. Under Bayesian framework, the Bayes estimators are developed based on Lindley's technique and Metropolis–Hastings algorithm. Furthermore, the highest posterior density (HPD) credible intervals are successively constructed. Finally, simulation experiments are implemented to access performance of several proposed methods in this article, and sewer invert trap real data is presented to exemplify the theoretical outcomes.

Keywords:

- *Bayes estimators; EM algorithm; generalized Type II progressive hybrid censoring; HPD credible interval; inverse Burr distribution; separation of sewer solids.*

AMS Subject Classification:

- 62N02, 62P12.

✉ Corresponding author

1. INTRODUCTION

Discharge of industrial and domestic effluent wastes, leakage from water tanks, marine dumping and atmospheric deposition are major causes of pollution. The removal of suspended solids from any sewer plays an important part in its overall waste treatment program. There are several methods for separating suspended particles from sewers. One of these methods is the use of invert traps. Several researcher have obtained experiment and simulation results on the invert trap.

For instance, Buxton *et al.* [7] presented the results from a laboratory investigation comparing the trapping performance of three slot size configurations of a laboratory-scale invert trap. Thinglas [25] studied flow field prediction and optimization of invert trap configuration using three-dimensional computational fluid dynamics (CFD) modeling. Mohsin and Kaushal [21] considered the experimental and discrete phase modeling for sediment retention ratio for invert traps. Moreover, in invert trap data analysis, the complete information is generally difficult to acquire on account of experimental cost and time-consuming of simulation. Therefore, censored data is more common whose censoring schemes are mainly divided into Type I and Type II censoring.

Furthermore, if Type I and Type II censored schemes are mixed together, it is hybrid censoring scheme (Epstein [13]). In Type I hybrid censored sample, the experiment stops at time $T^* = \min\{x_{m:m:n}, T\}$, where $X_{m:m:n}$ means the m-th failure time from n units, and T is the predetermined experiment time. Based on this censoring, it is a possibility that very few failures may occur before time T^*. So, Childs *et al.* [10] introduced the Type II hybrid censoring scheme that would terminate the experiment at $T^* = \max\{x_{m:m:n}, T\}$. Based on these censoring schemes, many statistical inferences have been carried out by several authors, see for example, Balakrishnan *et al.* [5], Banerjee and Kundu [4], Kundu and Howlader [17], Gupta and Singh [14].

A progressive censoring scheme (PCS) was then proposed to permit more flexibility in the conduct of the experiment, where individuals can be removed at several stages of the experiment rather than at the end. It can be classified into progressive Type I (PICS) and progressive Type II censoring schemes (PIICS). In PICS, let the number of items used in a life testing experiment be n. In this scheme, $R_1, R_2, ..., R_m$ items are randomly withdrawn at pre-specified time points $T_1, T_2, ..., T_m$, respectively. The test will be terminated at prefixed time point T_m in this scheme. Now, we describe PIICS. Consider n number of total units at initial time on an experiment. We remove randomly R_1 number of survival units when first failure time $X_{1:m:n}$ is observed. This process continues till m-th failure occurs. We assume that the m-th failure takes place at time $X_{m:m:n}$ and the remaining number of surviving units is $R_m = n - (m + \sum_{i=1}^{m-1} R_i)$.

Today, due to the high lifespan of many products, the total experimental time can be very long if PCS is used. Consequently, with the aim of enhancing the experimental efficiency and accuracy, it was further proposed as a progressive hybrid censoring scheme (PHCS). For various applications of the progressive hybrid sampling schemes in life testing experiments, we refer to Panahi [22] and El-Sherpieny *et al.* [12]. The main limitation of PHCS is that the number of observed failures is random and it can turn out to be a very small number, thus, any inference procedure will be invalid or its accuracy will be extremely low.

To overcome this drawback, a new hybrid censoring scheme has been proposed by Cho *et al.* [16] and is referred to generalized progressive hybrid censoring scheme (GPHCS), which maintains the experimental time in an acceptable range for researchers and guarantees a sufficient number of failed individuals. This scheme provides not only time and cost savings but also promotes more efficient statistical inference based on more observable data. The procedure of generalized Type II progressive hybrid (GIIPH) censoring scheme can be described as follows:

Suppose that n units are put on a test and the number of failures m, two time points T_1 and T_2 $(0 < T_1 < T_2 < \infty)$ and also the progressive censoring scheme $R_1, R_2, ..., R_m$ $(\sum_{j=1}^{m} R_j + m = n)$ are fixed beforehand. At the first failure time, (say $X_{1:m:n}$), R_1 number of live items are selected and randomly removed from the experiment. At the second failure time $(X_{2:m:n})$, R_2 units are removed from the remaining test items and so on, until the termination time $T^* = \max\{T_1,\ \min(x_{m:m:n}, T_2)\}$ failure observed and then all the remaining units are removed from the experiment. Let Q_1 and Q_2 denoted the number of observed failures up to time T_1 and T_2, respectively. Therefore,

- If $X_{m:m:n} < T_1$, then the experiment continue to observe failures until times T_1. In this case the failure times are denoted by $x_{1:m:n}, ..., x_{m:m:n}, x_{m+1:m:n}, ..., x_{q_1:n}$ (say Case I).
- If $T_1 < X_{m:m:n} < T_2$, then the experiment terminate at the m-th failure. In this case the failure times are represented by $x_{1:m:n}, ..., x_{q_1:m:n}, ..., x_{m:m:n}$ (say Case II).
- If $X_{m:m:n} > T_2$, then the experiment terminate at time T_2. In this case the failure times are denoted by $x_{1:m:n}, ..., x_{q_2:m:n}, ..., x_{m:m:n}$ (say Case III).

Where, q_1 and q_2 are the observed values of Q_1 and Q_2 respectively.

There are some authors studying this scheme under different lifetime distributions, see for example, Chan *et al.* [9], Gorny *et al.* [15], Koley and Cramer [18]. Based on the observed GIIPH censored sample, the likelihood function can be written as:

$$L(\alpha, \beta) = \begin{cases} \Im_i \prod_{j=1}^{Q_1} f(x_{j:m:n})[1 - F(x_{j:m:n})]^{R_j} & \text{Case I,} \\ \Im_i \prod_{j=1}^{m} f(x_{j:m:n})[1 - F(x_{j:m:n})]^{R_j} & \text{Case II,} \\ \Im_i \prod_{j=1}^{Q_2} f(x_{j:m:n})[1 - F(x_{j:m:n})]^{R_j} & \text{Case III,} \end{cases} \tag{1.1}$$

$$\Im_i = \begin{cases} [1 - F(T_1)]^{\tilde{R}_{Q_1+1}} \prod_{j=1}^{Q_1} \sum_{k=j}^{m} (1 + R_k), & \text{Case I,} \\ \prod_{j=1}^{m} \sum_{k=j}^{m} (1 + R_k) & \text{Case II,} \\ [1 - F(T_2)]^{\tilde{R}_{Q_2+1}} \prod_{j=1}^{Q_2} \sum_{k=j}^{m} (1 + R_k) & \text{Case III.} \end{cases}$$

In our work, estimation problems of unknown parameters of the inverse Burr (Burr III) distribution under GIIPH censoring scheme gets discussed. The Burr III distribution is one from twelve distributions was explored by using the method of differential equation (Burr [6]). This distribution has the following probability density function and the cumulative distribution function as:

$$f(x; \alpha, \beta) = \alpha\beta x^{-\beta-1}(1 + x^{-\beta})^{-(\alpha+1)};\ \ x > 0, \alpha > 0, \beta > 0, \tag{1.2}$$

and

$$F(x; \alpha, \beta) = (1 + x^{-\beta})^{-\alpha};\ \ x > 0, \alpha > 0, \beta > 0. \tag{1.3}$$

From Figure 1, it can be noticed that the inverse Burr distribution has two important shapes of its hazard rate function: decreasing and upside-down bathtub (or unimodal). It is worth mentioning that in reliability engineering, biology and several statistical modelling, different shaped hazard rate functions are used with different interpretations. We would like to mention that because of various shapes of the hazard rate function of inverse Burr distribution, it can be applied in many areas of research. Further, for fitting various lifetime data, inverse Burr distribution can be treated as an alternative model to other distributions such as gamma, Weibull and log-normal. Moreover, there are various real engineering data sets, for which inverse Burr (Burr III) distribution fits better than Weibull distribution.

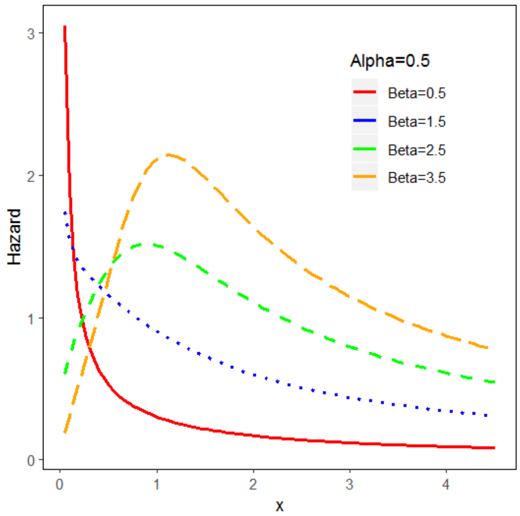


Figure 1: Graphs of the hazard rate function of the B-III distribution for different sets of parameters.

For example, the inverse Burr distribution fits the nano droplet dispersion data set (see Panahi and Asadi [23]).

The inverse Burr distribution has been studied by many researchers based on different censoring schemes. Abd-Elfattah and Alharbey [1] discussed the parameter estimations of this distribution based on a trimmed samples. Singh *et al.* [24] considered statistical inferences for the unknown parameters based on Type II progressive censoring scheme. Altindag *et al.* [2] studied the estimation and prediction problems for the inverse Burr distribution under Type II censored data. Panahi and Asadi [23] studied the application of this distribution on the Nano droplet censored data.

To the best of our knowledge, nobody has considered the inverse Burr distribution for the purpose of statistical inference based on GIIPH censoring scheme. Thus, our objectives in this study to close this gap are: First, estimating the parameters of the inverse Burr distribution using the EM algorithm. Using the Fisher information matrix, the approximate confidence intervals (ACIs) for unknown parameters are obtained.

Second objective is to obtain the Bayes estimates of the unknown inverse Burr parameters using independent gamma priors. Since the Bayes estimates cannot be obtained in closed expressions, Lindley's approximation and Markov chain Monte Carlo technique are considered to compute the complex posterior functions and in turn calculating Bayes estimates and the associated highest posterior density (HPD) credible intervals. Using various choices of the censoring schemes, the performance of the proposed methods is compared through an extensive simulation study in terms of their simulated mean squared-error (MSE) and average confidence lengths.

Also, third objective is to show the practical application of this distribution in separation of sewer solids data which are obtained by the authors using the computational fluid dynamics (CFD) method. The rest of the paper is organized as follows.

In Section 2, it is represented how the EM algorithm is utilized to obtain the maximum likelihood estimators (MLEs) of the unknown parameters as well as Fisher information matrix of the inverse Burr distribution under GIIPH censored sample. The existence and uniqueness properties of the MLEs have also been studied graphically. In Section 3, we derive the approximate explicit expressions for the Bayesian estimates using Lindley's approximation and Markov chain Monte Carlo technique. The Markov chain Monte Carlo samples are also used to construct the HPD credible intervals of the unknown parameters. Section 4 is devoted a simulation study to compare the proposed point and interval estimators. One real data set is analyzed for illustration in Section 5. Conclusions are given in Section 6.

2. MAXIMUM LIKELIHOOD ESTIMATORS

In this Section, the maximum likelihood method is carried out on the model based on the GIIPH censoring scheme. By (1.2), the likelihood function without additive constant is presented as follows:

$$L(\alpha,\beta)=\begin{cases}\prod_{j=1}^{q_1}\sum_{k=j}^{m}(1+R_k)(\alpha\beta)^{q_1}\prod_{j=1}^{q_1}x_{j:m:n}^{-\beta-1}A_j^{-(\alpha+1)}B_j^{R_j}D^{\tilde{R}_{Q_1+1}} & \text{case I}\\ \prod_{j=1}^{m}\sum_{k=j}^{m}(1+R_k)(\alpha\beta)^{m}\prod_{j=1}^{m}x_{j:m:n}^{-\beta-1}A_j^{-(\alpha+1)}B_j^{R_j} & \text{case II}\\ \prod_{j=1}^{q_2}\sum_{k=j}^{m}(1+R_k)(\alpha\beta)^{q_2}\prod_{j=1}^{q_2}x_{j:m:n}^{-\beta-1}A_j^{-(\alpha+1)}B_j^{R_j}D^{\tilde{R}_{Q_2+1}} & \text{case III,}\end{cases}\tag{2.1}$$

where $A_j=(1+x_{j:m:n}^{-\beta})$, $B_j=(1-(1+x_{j:m:n}^{-\beta})^{-\alpha})$, $\tilde{R}_{Q_1+1}=n-q_1-\sum_{j=1}^{m-1}R_j$, $\tilde{R}_{Q_2+1}=n-q_2-\sum_{j=1}^{q_2}R_j$ and

$$D=\begin{cases}1-(1+T_1^{-\beta})^{-\alpha} & \text{for case I}\\ 1-(1+T_2^{-\beta})^{-\alpha} & \text{for case III.}\end{cases}$$

The corresponding log-likelihood function is given by:

$$\begin{aligned}l(\alpha,\beta)&=\ln L(\alpha,\beta)\\ &=\iota(\ln\alpha+\ln\beta)-(\beta+1)\sum_{j=1}^{\iota}\ln x_{j:m:n}-(\alpha+1)\sum_{j=1}^{\iota}\ln A_j+\sum_{j=1}^{\iota}R_j\ln B_j+(\tilde{R}_{Q_j+1})\ln D,\end{aligned}$$

where

$$(\iota,\tilde{R}_{Q_j+1})=\begin{cases}(q_1,\tilde{R}_{Q_1+1}), & \text{Case I,}\\ (m,0) & \text{Case II,}\\ (q_2,\tilde{R}_{Q_2+1}), & \text{Case III.}\end{cases}$$

After differentiating the function $l(\alpha,\beta)$ with respect to α and β, we have

$$\hat{\alpha} = \iota\left(\sum_{j=1}^{\iota} \ln A_j - \sum_{j=1}^{\iota} R_j \frac{A_j^{-\alpha}\ln A_j}{B_j} - \varpi_1\right)^{-1}, \tag{2.2}$$

(2.3)
$$\hat{\beta} = \iota\left(\sum_{j=1}^{\iota} \ln x_{j:m:n} + \sum_{j=1}^{\iota} R_j\alpha \frac{x_{j:m:n}^{-\beta}\ln x_{j:m:n} A_j^{-\alpha-1}}{B_j} - \sum_{j=1}^{\iota}(\alpha+1)\frac{x_{j:m:n}^{-\beta}\ln x_{j:m:n}}{A_j} + \varpi_2\right)^{-1},$$

where

$$(\iota, \varpi_1) = \begin{cases} (q_1, \tilde{R}_{Q_1+1}\dfrac{(1+T_1^{-\beta})^{-\alpha}\ln(1+T_1^{-\beta})}{D}), & \text{Case I,} \\ (m, 0) & \text{Case II,} \\ (q_2, \tilde{R}_{Q_2+1}\dfrac{(1+T_2^{-\beta})^{-\alpha}\ln(1+T_2^{-\beta})}{D}), & \text{Case III,} \end{cases}$$

$$\varpi_2 = \begin{cases} \tilde{R}_{Q_1+1}\dfrac{\alpha T_1^{-\beta}\ln T_1(1+T_1^{-\beta})^{-\alpha-1}}{D}, & \text{Case I,} \\ 0 & \text{Case II,} \\ \tilde{R}_{Q_2+1}\dfrac{\alpha T_2^{-\beta}\ln T_2(1+T_2^{-\beta})^{-\alpha-1}}{D}, & \text{Case III,} \end{cases}$$

respectively. Now, we show the existence and uniqueness of the maximum likelihood estimates of the parameters of the inverse Burr distribution under GIIPH censored data using the graphical method (Ateya [3]) as:

- A sample of size 50 from the inverse Burr distribution are generated.
- Based on certain case of censored data ($T_1 = .8$, $T_2 = 2$, $m = 30$, $R_{15} = 20$, $R_j = 0$, $j \neq 15$), the curves of the equations $\partial(l(\alpha,\beta))/\partial(\alpha)$ and $\partial(l(\alpha,\beta))/\partial(\beta)$ are presented in Figure 2.
- The curve of $l(\hat{\alpha},\beta)$ and $l(\alpha,\hat{\beta})$ are also drawn in Figures 2 and 3, respectively.
- It is easy to see from Figure 1 that there exist one intersection point (1.1922,1.6455) which indicates that the solution of $\frac{\partial l(\alpha,\beta)}{\partial\alpha} = 0$ and $\frac{\partial l(\alpha,\beta)}{\partial\beta} = 0$, exists and is unique. This concludes that the maximum likelihood estimates of the parameters α and β exist and are unique.
- The Figure 3 shows that the previous intersection point maximizes the $l(\alpha,\hat{\beta})$.
- Similarly from Figure 4, it is observed the intersection point is the maximization point of the $l(\hat{\alpha},\beta)$.
- An important implication is that the maximum likelihood estimates of the parameters α and β exist and are unique for other generalized Type II progressive hybrid censored cases.

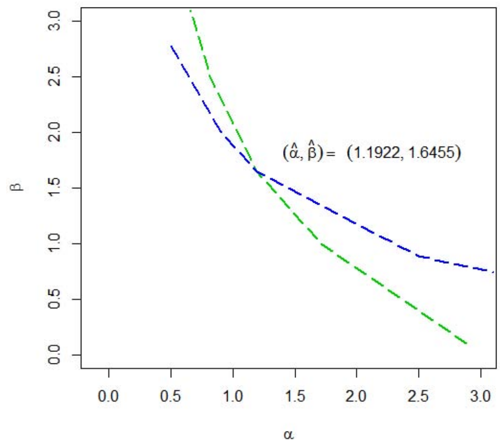


Figure 2: The plot of the ML estimates of α and β graphically.

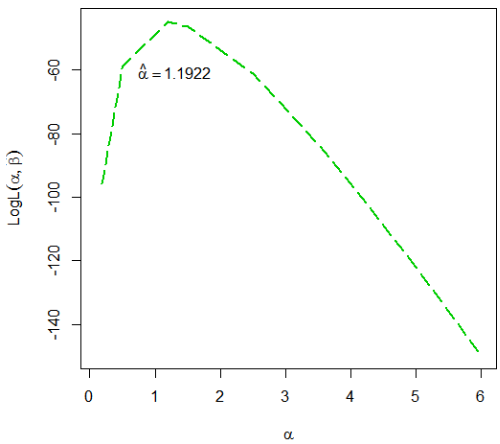


Figure 3: The curve of the log-likelihood function $l(\alpha, \hat{\beta})$.

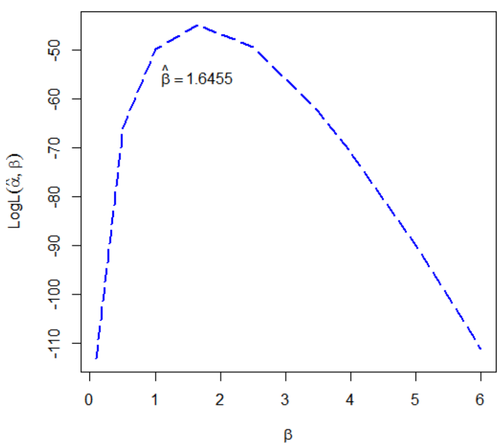


Figure 4: The curve of the the log-likelihood function $l(\hat{\alpha}, \beta)$.

2.1. EM algorithm

It is found that there is no explicit solution of (2.2) and (2.3), making them incredibly difficult to get the exact form of their solutions, thus utilizing the EM algorithm to work out these equations. Suppose that $\mathbf{X} = (X_1, X_2, ..., X_\iota)$ denotes the observed and $(\mathbf{Z}_j, \mathbf{Z}')$ represent the censored data. Where, $\mathbf{Z}_j = (Z_{j_1}, Z_{j_2}, ..., Z_{jR_j})$ and

$$\mathbf{Z}' = \begin{cases} (Z_1, Z_2, ..., Z_{\tilde{R}_{Q_1+1}}) & \text{Case I} \\ (Z_1, Z_2, ..., Z_{\tilde{R}_{Q_2+1}}) & \text{Case III,} \end{cases}$$

The log-likelihood function of (α, β) under the complete data is:

$$l_{\text{Complete}}(\alpha, \beta) = \begin{cases} \Delta + \mathtt{C}_1 & \text{Case I} \\ \Delta & \text{Case II} \\ \Delta + \mathtt{C}_2 & \text{Case III,} \end{cases} \tag{2.4}$$

$$\begin{aligned} \Delta = n\ln(\alpha) + n(\ln\beta) - (\beta+1)\sum_{j=1}^{\iota} \ln x_{j:m:n} - (\alpha+1)\sum_{j=1}^{\iota} \ln(1 + x_{j:m:n}^{-\beta}) \\ -(\beta+1)\sum_{j=1}^{\iota}\sum_{k=1}^{R_j} E[\ln Z_{jk}|Z_{jk} > x_{j:m:n}] - (\alpha+1)\sum_{j=1}^{\iota}\sum_{k=1}^{R_j} E[\ln(1 + Z_{jk}^{-\beta})|z_{jk} > x_{j:m:n}], \end{aligned}$$

$$\mathtt{C}_1 = -(\beta+1)\sum_{p=1}^{\tilde{R}_{Q_1+1}} E[\ln(Z'_p)|Z'_p > T_1] - (\alpha+1)\sum_{p=1}^{\tilde{R}_{Q_1+1}} E[\ln(1 + (Z'_p)^{-\beta})|Z'_p > T_1],$$

and

$$\mathtt{C}_2 = -(\beta+1)\sum_{p=1}^{\tilde{R}_{Q_2+1}} E[\ln Z'_p|Z'_p > T_2] - (\alpha+1)\sum_{p=1}^{\tilde{R}_{Q_2+1}} E[\ln(1 + (Z'_p)^{-\beta})|Z'_p > T_2].$$

The E-step of the EM-iteration needs the following conditional expectations:

$$E[\ln Z_{jk}|Z_{jk} > c] = \frac{\alpha\beta}{1 - F_X(c;\alpha,\beta)} \int_c^\infty x^{-\beta-1}(1 + x^{-\beta})^{-(\alpha+1)} \ln x dx = H_1(x_j, \alpha, \beta),$$

$$\begin{aligned} E[\ln(1 + Z_{jk}^{-\beta})|Z_{jk} > c] &= \frac{\alpha\beta}{1 - F_X(c;\alpha,\beta)} \int_c^\infty x^{-\beta-1}(1 + x^{-\beta})^{-(\alpha+1)} \ln(1 + x^{-\beta}) dx \\ &= H_2(x_j, \alpha, \beta), \end{aligned}$$

$$E[\ln Z'_p|Z'_p > T_1] = \frac{\alpha\beta}{1 - F_X(T_1;\alpha,\beta)} \int_{T_1}^\infty x^{-\beta-1}(1 + x^{-\beta})^{-(\alpha+1)} \ln x dx = H_3(x_j, \alpha, \beta),$$

$$E[\ln Z'_p|Z'_p > T_2] = \frac{\alpha\beta}{1 - F_X(T_2;\alpha,\beta)} \int_{T_2}^\infty x^{-\beta-1}(1 + x^{-\beta})^{-(\alpha+1)} \ln x dx = H_4(x_j, \alpha, \beta),$$

$$\begin{aligned} E[\ln(1 + (Z'_p)^{-\beta})|Z'_p > T_1] &= \frac{\alpha\beta}{1 - F_X(T_1;\alpha,\beta)} \int_{T_1}^\infty x^{-\beta-1}(1 + x^{-\beta})^{-(\alpha+1)} \ln(1 + x^{-\beta}) dx \\ &= H_5(T_1, \alpha, \beta), \end{aligned}$$

and

$$E[\ln(1+(Z'_p)^{-\beta})|Z'_p > T_2] = \frac{\alpha\beta}{1-F_X(T_2;\alpha,\beta)} \int_{T_2}^{\infty} x^{-\beta-1}(1+x^{-\beta})^{-(\alpha+1)} \ln(1+x^{-\beta})dx$$
$$= H_6(T_2,\alpha,\beta),$$

The M-step in a EM-iteration is maximizing the likelihood under complete sample over (α, β), with the missing values replaced by their conditional expectations.

2.2. Approximate Confidence Interval

For each unknown parameter, the approximate confidence intervals (ACIs) are presented by utilizing the observed Fisher information matrix. We have

$$I_{\mathbf{X}}(\theta) = I_{\mathbf{W}}(\theta) - I_{\mathbf{Z}|\mathbf{X}}(\theta). \tag{2.5}$$

Where,

$$I_{\mathbf{W}}(\theta) = -E_\theta\Big[\frac{\partial^2 l_{\text{Complete}}(\theta)}{\partial\theta^2}\Big]; \quad \theta = (\alpha,\beta), \tag{2.6}$$

and

$$\hat{l}_{\alpha\alpha} = \frac{\partial^2 l}{\partial\alpha^2}|_{\alpha=\hat{\alpha},\beta=\hat{\beta}} = -\frac{\iota}{\hat{\alpha}^2} - \sum_{j=1}^{\iota} R_j \frac{A_j^\alpha \ln^2(A_j)}{(A_j^\alpha - 1)^2} - \varpi_3,$$

Also, ϖ_3 is equal to $\tilde{R}_{Q_1+1}\{\frac{S_1^\alpha \ln^2(S_1)}{(S_1^\alpha-1)^2}\}$, 0 and $\tilde{R}_{Q_2+1}\{\frac{S_2^\alpha \ln^2 S_2}{(S_2^\alpha-1)^2}\}$ for cases I, II and III respectively.

$$\hat{l}_{\beta\beta} = \frac{\partial^2 l}{\partial\beta^2}|_{\alpha=\alpha,\beta=\hat{\beta}} = -\frac{\iota}{\beta^2} - (\alpha+1)\sum_{i=1}^{\iota}\frac{x_j^\beta \ln^2 x_j}{(1+x_j^\beta)^2} + \sum_{j=1}^{\iota}\alpha R_j \frac{x_j^\beta \ln^2 x_j (A_j^{\alpha+1}-1)}{(x_j^\beta(A_j^{\alpha+1}-1)-1)^2}$$
$$- \sum_{j=1}^{\iota}\alpha(\alpha+1)R_j\frac{\ln^2 x_j (A_j)^\alpha}{(x_j^\beta(A_j^{\alpha+1}-1)-1)^2} + \varpi_4,$$

$$\varpi_4 = \tilde{R}_{Q_1+1}\Big\{\frac{\alpha T_1^\beta \ln^2 T_1 (S_1^{\alpha+1}-1)}{(T_1^\beta(S_1^{\alpha+1}-1)-1)^2} - \frac{\alpha(\alpha+1)\ln^2 T_1 S_1^\alpha}{(T_1^\beta(S_1^{\alpha+1}-1)-1)^2}\Big\} \quad \text{for case I}$$
$$\varpi_4 = 0 \quad \text{for case II}$$
$$\varpi_4 = \tilde{R}_{Q_2+1}\Big\{\frac{\alpha T_2^\beta \ln^2 T_2 (S_2^{\alpha+1}-1)}{(T_2^\beta(S_2^{\alpha+1}-1)-1)^2} - \frac{\alpha(\alpha+1)\ln^2 T_2 (S_2)^\alpha}{(T_2^\beta(S_2^{\alpha+1}-1)-1)^2}\Big\} \quad \text{for case III}$$

and

$$\hat{l}_{\beta\alpha} = \hat{l}_{\alpha\beta} = \frac{\partial^2 l}{\partial\beta\partial\alpha}|_{\alpha=\hat{\alpha},\beta=\hat{\beta}} = \sum_{j=1}^{\iota}\frac{x_j}{1+x_j^\beta} - \sum_{j=1}^{\iota}R_j\frac{\ln x_j}{x_j^\beta(A_j^{\alpha+1}-1)-1}$$
$$+ \sum_{j=1}^{\iota}\alpha R_j\frac{A_j^{\alpha+1}x_j^\beta \ln x_j ln(A_j)}{(x_j^\beta(A_j^{\alpha+1}-1)-1)^2} - \varpi_5$$

$$\varpi_5 = \tilde{R}_{Q_1+1}\Big\{\frac{\ln T_1}{T_1^{\beta}(S_1^{\alpha+1}-1)-1} - \frac{\alpha T_1^{\beta}\ln T_1 S_1^{\alpha+1}\ln(S_1)}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2}\Big\} \quad \text{for case I}$$
$$\varpi_5 = 0 \quad \text{for case II}$$
$$\varpi_5 = \tilde{R}_{Q_2+1}\Big\{\frac{\ln T_2}{T_2^{\beta}(S_2^{\alpha+1}-1)-1} - \frac{\alpha T_2^{\beta}\ln T_1 S_2^{\alpha+1}\ln(S_2)}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2}\Big\} \quad \text{for case III.}$$

Note that, we consider $x_{j:m:n} = x_j$, $S_1 = (1 + T_1^{-\beta})$, $S_2 = (1 + T_2^{-\beta})$, and $A_j = (1 + x_j^{-\beta})$. Based on the conditional distribution, the Fisher information in the j-th observation can be evaluated as

$$I^{(j)}_{\mathbf{Z}|\mathbf{X}}(\theta) = -E\big[\frac{\partial^2}{\partial\theta^2}\ln(f(z_{jk}|x_{j:m:n},\theta))\big]. \tag{2.7}$$

Therefore, we have

$$I_{\mathbf{Z}|\mathbf{X}}(\theta) = \begin{cases} \sum_{j=1}^{q_1} R_i I^{(j)}_{\mathbf{Z}|\mathbf{X}}(\theta) + \tilde{R}_{Q_1+1} I^{*}_{\mathbf{Z}|\ \mathbf{X}}(\theta), & \text{Case I} \\ \sum_{j=1}^{m} R_i I^{(j)}_{\mathbf{Z}|\mathbf{X}}(\theta), & \text{Case II} \\ \sum_{j=1}^{q_2} R_i I^{(j)}_{\mathbf{Z}|\mathbf{X}}(\theta) + \tilde{R}_{Q_2+1} I^{*}_{\mathbf{Z}|\mathbf{X}}(\theta) & \text{Case III.} \end{cases}$$

Where, $I^{(j)}_{\mathbf{Z}|\mathbf{X}}(\theta)$ and $I^{*}_{\mathbf{Z}|\mathbf{X}}(\theta)$ are the information matrix of a single observation for the truncated inverse Burr distribution. Therefore, the $100(1-\gamma)\%$ ACIs for the parameters are given by:

$$\Big(\hat{\alpha} - Z_{\gamma/2}\sqrt{\mathrm{Var}(\hat{\alpha})}\ ,\ \hat{\alpha} + Z_{\gamma/2}\sqrt{\mathrm{Var}(\hat{\alpha})}\Big)\ \ \textit{and}\ \ \Big(\hat{\beta} - Z_{\gamma/2}\sqrt{\mathrm{Var}(\hat{\beta})}\ ,\ \hat{\beta} + Z_{\gamma/2}\sqrt{\mathrm{Var}(\hat{\beta})}\Big).$$

3. BAYESIAN ETIMATION

In contrast to traditional frequentist methods, the Bayesian approaches take advantage of available data information and incorporate prior information of parameters, thereby attracting much attention in statistical inference. For obtaining the Bayesian estimates, we consider independent gamma prior distributions for α and β with hyper-parameters (a_1, b_1) and (a_2, b_2) respectively, that reflect prior beliefs. Hence the PDF of the joint prior distribution takes the following expression:

$$\pi(\alpha,\beta) \propto \alpha^{b_1-1}e^{-a_1\alpha}\beta^{b_2-1}e^{-a_2\beta};\ \alpha > 0,\ \beta > 0,\ a_1 > 0,\ a_2 > 0,\ b_1 > 0,\ b_2 > 0, \tag{3.1}$$

In prior distributions, hyper parameters a_i and $b_i, i = 1, 2$ are assumed as non-negative and known. In the case of noninformative priors, very small non-negative values of the hyper-parameters , *i.e.* $a_1 = a_2 = b_1 = b_2 = 0.0001$, are used as suggested by Congdon [8] which are almost like Jeffrey's priors, but they are proper, inversely. As more informative priors, different cases of the hyperparameters can be evaluated. Therefore, Bayes estimation of a general function of parameters $(\Upsilon(\alpha, \beta))$ with the square error loss function can be derived as

$$\tilde{\Upsilon}(\alpha,\beta) = E(\Upsilon(\alpha,\beta)|Data) = \top^{-1}\iint \Upsilon(\alpha,\beta)\pi(\alpha,\beta|Data)d\alpha d\beta, \tag{3.2}$$

where $\top = \iint \pi(\alpha,\beta|Data)d\alpha d\beta$ and
(3.3)

$$\pi(\alpha,\beta|Data)=\begin{cases}\psi\alpha^{q_1+b_1-1}\beta^{q_1+b_2-1}\prod_{j=1}^{q_1}x_{j:m:n}^{-\beta-1}A_j^{-(\alpha+1)}e^{-\alpha a_1-\beta a_2}B_j^{R_j}D^{\tilde{R}_{Q_1+1}} & \text{case I},\\ \psi\alpha^{m+b_1-1}\beta^{m+b_2-1}\prod_{j=1}^{m}x_{j:m:n}^{-\beta-1}A_j^{-(\alpha+1)}e^{-\alpha a_1-\beta a_2}B^{R_j} & \text{case II},\\ \psi\alpha^{q_2+b_1-1}\beta^{q_2+b_2-1}\prod_{j=1}^{q_2}x_{j:m:n}^{-\beta-1}A_j^{-(\alpha+1)}e^{-\alpha a_1-\beta a_2}B_j^{R_j}D^{\tilde{R}_{Q_2+1}} & \text{case III}.\end{cases}$$

Here, A_j, B_j, D are introduced previously and ψ and Υ can be written as:

$$\psi=\begin{cases}\prod_{j=1}^{q_1}\sum_{k=j}^{m}(1+R_k) & \text{for case I},\\ \prod_{j=1}^{m}\sum_{k=j}^{m}(1+R_k) & \text{for case II},\\ \prod_{j=1}^{q_2}\sum_{k=j}^{m}(1+R_k) & \text{for case III},\end{cases}$$

and

$$\Upsilon(\alpha,\beta)=\alpha^{\upsilon_1}\beta^{\upsilon_2}\rightsquigarrow\begin{cases}\upsilon_1=1,\upsilon_2=0 & \text{for estimating }\alpha\\ \upsilon_1=0,\upsilon_2=1 & \text{for estimating }\beta.\end{cases}$$

It is clear that the Bayes estimator in (3.2) cannot be obtained analytically. Therefore, some approximation methods are required in order to compute the approximate Bayes estimates. We adopt the Lindley's method and Metropolis–Hastings algorithm to solve the problem.

3.1. Lindley's Approximation

In the above Section, we see that the proposed Bayes estimates are in the form of the ratio of two integrals. These integrals can not be evaluated in terms of some closed-form expressions. So, we developed the Bayesian estimates using the Lindley's approximation (Lindley [19]). Based on the Lindley's method, the Bayes estimations of parameters have the following expression:

$$\tilde{\alpha}=\hat{\alpha}+\frac{1}{2}\big[2\hat{\rho}_\alpha\hat{\upsilon}_{\alpha\alpha}+2\hat{\rho}_\beta\hat{\upsilon}_{\alpha\beta}+\hat{\upsilon}^2_{\alpha\alpha}\hat{l}_{\alpha\alpha\alpha}+\hat{\upsilon}_{\alpha\alpha}\hat{\upsilon}_{\beta\beta}\hat{l}_{\beta\beta\alpha}+2\hat{\upsilon}_{\alpha\beta}\hat{\upsilon}_{\beta\alpha}\hat{l}_{\alpha\beta\beta}+\hat{\upsilon}_{\alpha\beta}\hat{\upsilon}_{\beta\beta}\hat{l}_{\beta\beta\beta}\big],$$

and

$$\tilde{\beta}=\hat{\alpha}+\frac{1}{2}\big[2\hat{\rho}_\alpha\hat{\upsilon}_{\beta\beta}+2\hat{\rho}_\beta\hat{\upsilon}_{\beta\alpha}+\hat{\upsilon}^2_{\beta\beta}\hat{l}_{\beta\beta\beta}+3\hat{\upsilon}_{\beta\beta}\hat{\upsilon}_{\alpha\beta}\hat{l}_{\alpha\beta\beta}+\hat{\upsilon}_{\alpha\alpha}\hat{\upsilon}_{\beta\alpha}\hat{l}_{\alpha\alpha\alpha}\big].$$

Here, $\hat{\rho}_\alpha=\frac{b_1-1}{\hat{\alpha}}-a_1$, $\hat{\rho}_\beta=\frac{b_2-1}{\hat{\beta}}-a_2$, $\hat{l}_{\alpha^n\beta^m}=\partial^{n+m}l(\alpha,\beta)/\partial\alpha^n\partial\beta^m; n,m=0,1,...$ and $\hat{\upsilon}_{ij}$ are the (ij)-th elements of matrix $\big[-\partial^2 l(\alpha,\beta)/\partial\alpha\partial\beta\big]^{-1}; i,j=1,2$. Also, we have

$$\hat{l}_{\alpha\alpha\alpha}=\frac{\partial^3 l}{\partial\alpha^3}=\frac{2\iota}{\alpha^3}-\sum_{j=1}^{\iota}R_j\frac{A_j^\alpha\ln^3(A_j)}{(A_j^\alpha-1)^2}+2\sum_{j=1}^{\iota}R_j\frac{A_j^{2\alpha}\ln^3(A_j)}{(A_j^\alpha-1)^3}-\varpi_6,$$

$$\begin{aligned}\varpi_6 &= \tilde{R}_{Q_1+1}\Big\{\frac{S_1^\alpha\ln^3(S_1)}{(S_1^\alpha-1)^2}-2\frac{S_1^{2\alpha}\ln^3(S_1)}{(S_1^\alpha-1)^3}\Big\} && \text{for case I},\\ \varpi_6 &= 0 && \text{for case II},\\ \varpi_6 &= \tilde{R}_{Q_2+1}\Big\{\frac{S_2^\alpha\ln^3(S_2)}{(S_2^\alpha-1)^2}-2\frac{S_2^{2\alpha}\ln^3(S_2)}{(S_2^\alpha-1)^3}\Big\} && \text{for case III},\end{aligned}$$

$$\hat{l}_{\alpha\alpha\beta} = \hat{l}_{\alpha\beta\alpha} = \hat{l}_{\beta\alpha\alpha} = \frac{\partial^3 l}{\partial\alpha^2\partial\beta}|_{\alpha=\alpha,\beta=\beta} + \sum_{j=1}^{\iota} R_j \frac{\alpha A_j^{\alpha-1}\ln x_j x_j^{-\beta}\ln^2(A_j)}{(A_j^{\alpha}-1)^2}$$
$$+ 2\sum_{j=1}^{\iota} R_j \frac{A_j^{\alpha-1}\ln x_j x_j^{-\beta}\ln(A_j)}{(A_j^{\alpha}-1)^2} - 2\sum_{j=1}^{\iota} R_j \frac{\alpha A_j^{2\alpha-1}\ln x_j x_j^{-\beta}\ln^2(A_j)}{(A_j^{\alpha}-1)^3} + \varpi_7,$$

$$\varpi_7 = \tilde{R}_{Q_1+1}\Big\{\frac{\alpha S_1^{\alpha-1}T_1^{-\beta}\ln T_1\ln^2(S_1)}{(S_1^{\alpha}-1)^2} + 2\frac{\ln(S_1)T_1^{-\beta}\ln T_1 S_1^{\alpha-1}}{(S_1^{\alpha}-1)^2}$$
$$- 2\frac{\alpha S_1^{2\alpha-1}T_1^{-\beta}\ln T_1\ln^2(S_1)}{(S_1^{\alpha}-1)^3}\Big\} \qquad \text{for case I},$$
$$\varpi_7 = 0 \qquad \text{for case II},$$
$$\varpi_7 = \tilde{R}_{Q_2+1}\Big\{\frac{\alpha S_2^{\alpha-1}T_2^{-\beta}\ln T_2\ln^2(S_2)}{(S_2^{\alpha}-1)^2} + 2\frac{\ln(S_2)T_2^{-\beta}\ln T_2 S_2^{\alpha-1}}{(S_2^{\alpha}-1)^2}$$
$$- 2\frac{\alpha S_2^{2\alpha-1}T_2^{-\beta}\ln T_2\ln^2(S_2)}{(S_2^{\alpha}-1)^3}\Big\} \qquad \text{for case III},$$

$$\hat{l}_{\beta\beta\beta} = \frac{\partial^3 l}{\partial\beta^3}|_{\alpha=\hat{\alpha},\beta=\hat{\beta}} = \frac{2\iota}{\beta^3} - \sum_{j=1}^{\iota}(\alpha+1)\frac{x_j^{\beta}\ln^3 x_j}{(1+x_j^{\beta})^2} + \sum_{j=1}^{\iota}2(\alpha+1)\frac{x_j^{2\beta}\ln^3 x_j}{(1+x_j^{\beta})^3}$$
$$+ \sum_{j=1}^{\iota}\alpha R_j\frac{x_j^{\beta}\ln^3 x_j(A_j^{\alpha+1}-1)}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2} - \sum_{j=1}^{\iota}\alpha(\alpha+1)R_j\frac{\ln^3 x_j A_j^{\alpha}}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2}$$
$$- \sum_{j=1}^{\iota}2\alpha R_j\frac{x_j^{2\beta}\ln^3 x_j(A_j^{\alpha+1}-1)^2}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^3} + \sum_{j=1}^{\iota}4\alpha(\alpha+1)R_j\frac{x_j^{\beta}A_j^{\alpha}\ln^3 x_j(A_j^{\alpha+1}-1)}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^3}$$
$$+ \sum_{j=1}^{\iota}\alpha^2(\alpha+1)R_j\frac{x_j^{-\beta}\ln^3 x_j A_j^{\alpha-1}}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2} - \sum_{j=1}^{\iota}2\alpha(\alpha+1)^2R_j\frac{\ln^3 x_j A_j^{2\alpha}}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^3}$$
$$+ \varpi_8,$$

$$\varpi_8 = \tilde{R}_{Q_1+1}\Big\{\frac{\alpha T_1^{\beta}\ln^3 T_1(S_1^{\alpha+1}-1)}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2} - \frac{\alpha(\alpha+1)\ln^3 T_1 S_1^{\alpha}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2}$$
$$- \frac{2\alpha T_1^{2\beta}\ln^3 T_1(S_1^{\alpha+1}-1)^2}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^3} + \frac{4\alpha(\alpha+1)T_1^{\beta}\ln^3 T_1 S_1^{\alpha}(S_1^{\alpha+1}-1)}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^3}$$
$$+ \frac{\alpha^2(\alpha+1)T_1^{-\beta}\ln^3 T_1 S_1^{\alpha-1}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2} - \frac{2\alpha(\alpha+1)^2\ln^3 T_1 S_1^{2\alpha}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^3}\Big\} \quad \text{for case I},$$
$$\varpi_8 = 0 \qquad \text{for case II},$$
$$\varpi_8 = \tilde{R}_{Q_2+1}\Big\{\frac{\alpha T_2^{\beta}\ln^3 T_2(S_2^{\alpha+1}-1)}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2} - \frac{\alpha(\alpha+1)\ln^3 T_2 S_2^{\alpha}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2}$$
$$- \frac{2\alpha T_2^{2\beta}\ln^3 T_2(S_2^{\alpha+1}-1)^2}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^3} + \frac{4\alpha(\alpha+1)T_2^{\beta}\ln^3 T_2 S_2^{\alpha}(S_2^{\alpha+1}-1)}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^3}$$
$$+ \frac{\alpha^2(\alpha+1)T_2^{-\beta}\ln^3 T_2 S_2^{\alpha-1}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2} - \frac{2\alpha(\alpha+1)^2\ln^3 T_2 S_2^{2\alpha}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^3}\Big\} \quad \text{for case III},$$

$$
\begin{aligned}
\hat{l}_{\beta\beta\alpha} &= \hat{l}_{\beta\alpha\beta} = \hat{l}_{\alpha\beta\beta} = \frac{\partial^3 l}{\partial\beta^2\partial\alpha}|_{\alpha=\hat{\alpha},\beta=\hat{\beta}} = -\sum_{i=1}^{\iota} \frac{x_j^{\beta}\ln^2 x_j}{(1+x_j^{\beta})^2} \\
&+ \sum_{j=1}^{\iota} R_j \frac{\ln^2 x_j x_j^{\beta}(A_j^{\alpha+1}-1)}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2} + \sum_{j=1}^{\iota} \alpha R_j \frac{A_j^{\alpha+1}x_j^{\beta}\ln(A_j)\ln^2 x_j}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2} \\
&- \sum_{j=1}^{\iota} 2\alpha R_j \frac{A_j^{\alpha+1}(A_j^{\alpha+1}-1)x_j^{2\beta}\ln^2 x_j \ln(A_j)}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^3} \\
&- \sum_{j=1}^{\iota}(2\alpha+1)R_j \frac{A_j^{\alpha}\ln^2 x_j}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2} - \sum_{j=1}^{\iota}(\alpha^2+\alpha)R_j \frac{A_j^{\alpha}\ln^2 x_j \ln(A_j)}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^2} \\
&+ \sum_{j=1}^{\iota} 2\alpha(\alpha+1)R_j \frac{A_j^{2\alpha+1}x_j^{\beta}\ln^2 x_j \ln(A_j)}{(x_j^{\beta}(A_j^{\alpha+1}-1)-1)^3} + \varpi_9,
\end{aligned}
$$

$$
\begin{aligned}
\varpi_9 &= \tilde{R}_{Q_1+1}\Big\{ \frac{T_1^{\beta}\ln^2 T_1(S_1^{\alpha+1}-1)}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2} + \alpha \frac{\ln^2 T_1 S_1^{\alpha+1}\ln(S_1)T_1^{\beta}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2} \\
&- 2\alpha \frac{T_1^{2\beta}\ln^2 T_1(S_1^{\alpha+1}-1)S_1^{\alpha+1}\ln S_1}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^3} - (2\alpha+1)\frac{\ln^2 T_1(S_1)^{\alpha}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2}
\end{aligned}
$$

$$
\begin{aligned}
&- (\alpha+\alpha^2)\frac{\ln^2 T_1 \ln(S_1)S_1^{\alpha}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^2} + 2\alpha(\alpha+1)\frac{\ln^2 T_1 \ln(S_1)T_1^{\beta}S_1^{2\alpha+1}}{(T_1^{\beta}(S_1^{\alpha+1}-1)-1)^3}\Big\} \quad \text{for case I,} \\
\varpi_9 &= 0 \qquad \text{for case II,}
\end{aligned}
$$

$$
\begin{aligned}
\varpi_9 &= \tilde{R}_{Q_2+1}\Big\{ \frac{T_2^{\beta}\ln^2 T_2(S_2^{\alpha+1}-1)}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2} + \alpha \frac{\ln^2 T_2 S_1^{\alpha+1}\ln(S_2)T_2^{\beta}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2} \\
&- 2\alpha \frac{T_2^{2\beta}\ln^2 T_2(S_2^{\alpha+1}-1)S_2^{\alpha+1}\ln S_2}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^3} - (2\alpha+1)\frac{\ln^2 T_2 S_2^{\alpha}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2} \\
&- (\alpha+\alpha^2)\frac{\ln^2 T_2 \ln(S_2)S_2^{\alpha}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^2} + 2\alpha(\alpha+1)\frac{\ln^2 T_2 \ln(S_2)T_2^{\beta}S_2^{2\alpha+1}}{(T_2^{\beta}(S_2^{\alpha+1}-1)-1)^3}\Big\} \quad \text{for case III.}
\end{aligned}
$$

3.2. Metropolis–Hastings Algorithm

In the previous Subsection, we obtain the Bayes estimates using Lindley's approximation method. One disadvantage of this method is that it requires higher order partial derivatives of the log-likelihood function. Further, the Lindley's approximation can not be used to construct HPD credible intervals. Moreover, it is observed that the conditional posterior distribution of unknown parameters cannot be reduced to any well-known distribution. To overcome this problem, we propose to apply the Metropolis–Hastings (Metropolis *et al.* [20]) algorithm for generating samples from the respective posterior distributions. This algorithm is the most popular example of the Markov chain Monte Carlo (MCMC) method and it is free from the higher order partial derivatives. The basic scheme of the Metropolis–Hastings (M-H) is given as follows:

Step 1: Use the MLEs of (α, β) as the initial point of the iteration, denoted by (α_0, β_0).

Step 2: Generate α_j and β_j from the normal proposal distributions $N(\alpha_{j-1}, \sigma^2)$ and $N(\beta_{j-1}, \sigma^2)$, respectively, for $j = 1, ..., N$.

Step 3: Compute $h = \dfrac{\pi(\alpha_j, \beta_j|Data)}{\pi(\alpha_{j-1}, \beta_{j-1}|Data)}$.

Step 4: Accept the new sample with probability $\min(1, h)$.

Step 5: Set $j = j + 1$.

Step 6: Repeat Step 2–5, up to N times.

So, the Bayes estimates of α and β are respectively obtained as below:

$$\tilde{\alpha} = \frac{1}{N - N_0} \sum_{i=N_0+1}^{N} \alpha_i, \qquad \text{and} \qquad \tilde{\beta} = \frac{1}{N - N_0} \sum_{i=N_0+1}^{N} \beta_i.$$

In order to guarantee the convergence and to remove the affection of the selection of initial values, the first N_0 simulated varieties are discarded (burn-in-period of Markov chain). Also, for computing the confidence interval based on MCMC samples, we first order the samples $\alpha_{1:N}, \alpha_{2:N}, ..., \alpha_{N:N}$ and $\beta_{1:N}, \beta_{2:N}, ..., \beta_{N:N}$, then a $(1 - \gamma) \times 100\%$ HPD credible interval for α and β are obtained as:

$$[\alpha_{N\gamma}, \alpha_{N(1-\gamma)}] \qquad \text{and} \qquad [\beta_{N\gamma}, \beta_{N(1-\gamma)}].$$

Finally, choose the interval which has the smallest width as a HPD credible interval.

4. SIMULATION STUDY

To evaluate the behavior of the theoretical results obtained in the previous Sections, including the classical and Bayesian estimators and the associated confidence/credible intervals, an extensive Monte Carlo simulation study is performed. We simulate GIIPH censored samples for different combinations of (n, m, T_1, T_2) from the inverse Burr (α, β) distribution. We adopted the true values of unknown parameters as $\alpha = 1.2$ and $\beta = 1.6$. Note that all the computations have been performed using R software. Through the sample data, we evaluate the MLEs by employing an EM algorithm. Approximate expressions for the Bayesian estimators have been obtained using the Lindley's approximation and Metropolis–Hastings algorithm. Using the M-H sampler algorithm described in Subsection 3.2, 10000 MCMC samples and discard the first 2000 values as 'burn-in' are generated. In Bayesian paradigm, the choice of the hyper-parameter values is the main issue. For this propose, both non-informative prior (NIP) and informative prior (IP) are taken into account in the Bayesian approach, where all hyper-parameters in the NIP are chosen to be 0.0001 instead of 0, which is more appropriate since the hyper-parameters are greater than 0, and the hyper-parameters in the IP are selected according to this manner: the means of prior (PR) distributions are equal to original parameters($a_1 = 1.2, a_2 = 1.6, b_1 = 1, b_2 = 1$). The %95 approximate confidence (AC) and Bayesian (HPD) intervals for the parameters are also constructed. The HPD credible intervals are computed based on 10000 MCMC samples. We take three different censoring schemes as follows:

Scheme 1: $R_m = n - m$ and $R_j = 0$ for $j \neq m$;

Scheme 2: $R_1 = R_m = (n - m)/2$ and $R_j = 0$ for $j \neq 1, m$;

Scheme 3: $R_{m/2} = n - m$ and $R_j = 0$ for $j \neq m/2$.

Based on these set up assumptions, we show the numerical results in the Table 1, Table 2, Table 3 and Table 4.

Tables 1 and 2 (also, Figure 5 and Figure 6) present the average ML and Bayes estimates and the corresponding MSEs based on 10000 replications. Moreover, the average lower and upper bounds of the AC and HPD intervals are displayed in Tables 3 and 4.

The following conclusions are found from Tables 1–4 and Figures 5–6:

- For fixed n, T_1 and T_2 as m increases, the average estimates and the MSEs of the parameters decreases. Also, with increasing m, the average lengths of all intervals mostly decrease.
- For fixed m, T_1 and T_2 as sample size n increases the MSEs of all the estimators decreases (Figure 6). Similar trend is observed (Figure 6) for fixed n, T_1 and T_2 as m increases.
- The MSEs have a downward trend for fixed n, m, T_1 and increasing T_2 (Figure 5).
- For fixed n, m and T_2 as T_1 increases, the MSEs decreases (Figure 5).
- To evaluate the effect of the proposed estimation methods with respect to the smallest MSE, it is observed that the Bayes estimates work efficiently and provide better performance as compared to those obtained based on MLEs. For the parameters α and β, the MSEs of the maximum likelihood estimates are larger than the Bayes estimates.
- The Bayesian MCMC estimation using M-H algorithm sampler for the unknown parameters under GIIPH censoring is recommended for all values of n, m, T_1 and T_2.
- As expected, the Bayesian estimation with IP tends to be preferable to that with NIP.
- The average lengths of the ACI for α and β are relatively large compared to those of Bayesian credible intervals.
- As for the Bayes method, similar to the findings for the point estimates, the Bayesian intervals under non-informative prior are slightly worse than those under informative prior.

Table 1: The MSEs of the MLEs ($\hat{\alpha}$), Lindleys ($\tilde{\alpha}_{LIN}$) and M-Hs ($\tilde{\alpha}_{MH}$) for $T_1 = 1$, $T_2 = 2.5$.

n	m	*Scheme*	$\hat{\alpha}(MSE)$	*PR*	$\tilde{\alpha}_{LIN}(MSE)$	$\tilde{\alpha}_{MH}(MSE)$
30	15	R_1	1.4754(0.2256)	*IP*	1.4897(0.1945)	1.4270(0.0979)
				NIP	1.5023(0.2094)	1.4328(0.1006)
		R_2	1.4427(0.1789)	*IP*	1.4475(0.1581)	1.3665(0.0733)
				NIP	1.4596(0.1637)	1.3709(0.0769)
		R_3	1.4246(0.1005)	*IP*	1.3937(0.0897)	1.3508(0.0621)
				NIP	1.4214(0.0942)	1.3700(0.0699)
50	24	R_1	1.2982(0.1785)	*IP*	1.3687(0.1586)	1.2843(0.0464)
				NIP	1.3278(0.1669)	1.3199(0.0497)
		R_2	1.3064(0.0939)	*IP*	1.3000(0.0900)	1.1674(0.0175)
				NIP	1.3087(0.0911)	1.1721(0.0244)
		R_3	0.9172(0.0439)	*IP*	0.9450(0.0382)	0.9913(0.0230)
				NIP	0.9608(0.0410)	0.9725(0.0277)
	30	R_1	1.2982(0.1316)	*IP*	1.2786(0.1212)	1.2028(0.0388)
				NIP	1.2865(0.1283)	1.2536(0.0449)
		R_2	1.1464(0.0476)	*IP*	1.1397(0.0455)	0.9741(0.0091)
				NIP	1.1471(0.0462)	01.9932(.0105)
		R_3	1.006(0.0295)	*IP*	0.9995(0.0261)	1.001(0.0141)
				NIP	1.002(0.0269)	1.009(0.0169)
100	48	R_1	1.3875(0.1632)	*IP*	1.3768(0.1544)	1.2434(0.0336)
				NIP	1.3811(0.1598)	1.2709(0.0390)
		R_2	1.1687(0.0283)	*IP*	1.1666(0.0277)	1.1019(0.0100)
				NIP	1.1679(0.0281)	1.1052(0.0133)
		R_3	0.9414(0.0221)	*IP*	0.9551(0.0205)	1.1174(0.0188)
				NIP	0.9532(0.0215)	1.1168(0.0209)
	60	R_1	1.3212(0.1226)	*IP*	1.3105(0.1152)	1.2190(0.0327)
				NIP	1.3176(0.1174)	1.2187(0.0328)
		R_2	0.9028(0.0090)	*IP*	0.9162(0.0070)	0.9786(0.0004)
				NIP	0.9200(0.0084)	0.9921(0.0013)
		R_3	1.0090(0.0145)	*IP*	1.0060(0.0136)	1.0400(0.0016)
				NIP	1.008(0.0140)	1.0488(0.0021)

Table 2: The MSEs of the MLEs ($\hat{\beta}$), Lindleys ($\tilde{\beta}_{LIN}$) and M-Hs ($\tilde{\beta}_{MH}$) for $T_1 = 1$, $T_2 = 2.5$.

n	m	*Scheme*	$\hat{\beta}(MSE)$	PR	$\tilde{\beta}_{LIN}(MSE)$	$\tilde{\beta}_{MH}(MSE)$
30	15	R_1	1.3675(0.2354)	*IP*	1.3954(0.1738)	1.3700(0.1614)
				NIP	1.4186(0.1845)	1.3961(0.1823)
		R_2	1.6325(.1535)	*IP*	1.5885(0.1366)	1.5143(0.1122)
				NIP	1.5899(0.1389)	1.5357(0.1308)
		R_3	1.7859(0.1673)	*IP*	1.7654(0.1611)	1.4998(0.1066)
				NIP	1.7865(0.1738)	1.5403(0.1231)
50	24	R_1	1.2763(0.1807)	*IP*	1.3029(0.1322)	1.3122(0.1316)
				NIP	1.3043(0.1493)	1.3145(0.1387)
		R_2	1.5917(.1068)	*IP*	1.5643(0.1031)	1.5457(0.0900)
				NIP	1.5749(0.1054)	1.5499(0.0938)
		R_3	1.7431(0.1673)	*IP*	1.7183(0.1527)	1.5125(0.0760)
				NIP	1.7327(0.1602)	1.5226(0.0811)
	30	R_1	1.3846(0.1155)	*IP*	1.4220(0.1136)	1.7034(0.0921)
				NIP	1.4165(0.1140)	1.6993(0.0976)
		R_2	1.5075(0.0860)	*IP*	1.5064(0.0850)	1.6737(0.0540)
				NIP	1.5069(0.0857)	1.3199(0.0497)
		R_3	1.6514(0.0897)	*IP*	1.6094(0.0741)	1.5509(0.0240)
				NIP	1.6328(0.0809)	1.5499(0.0296)
100	48	R_1	1.2681(0.1699)	*IP*	1.2796(0.1242)	1.6234(0.1045)
				NIP	1.2765(0.1374)	1.6309(0.1089)
		R_2	1.7893(0.0558)	*IP*	1.7674(0.0480)	1.4064(0.0540)
				NIP	1.7763(0.0518)	1.4078(0.0616)
		R_3	1.6676(0.0573)	*IP*	1.6559(0.0540)	1.4797(0.0244)
				NIP	1.6621(0.0564)	1.4859(0.0289)
	60	R_1	1.3435(0.1047)	*IP*	1.3586(0.0905)	1.6112(0.0446)
				NIP	1.3488(0.0977)	1.6269(0.0535)
		R_2	1.5587(0.0460)	*IP*	1.54974(0.0430)	1.6850(0.0070)
				NIP	1.5546(0.0451)	1.6589(0.0087)
		R_3	1.5973(0.0397)	*IP*	1.5784(0.0370)	1.6405(0.0046)
				NIP	1.5884(0.0383)	1.6712(0.0065)

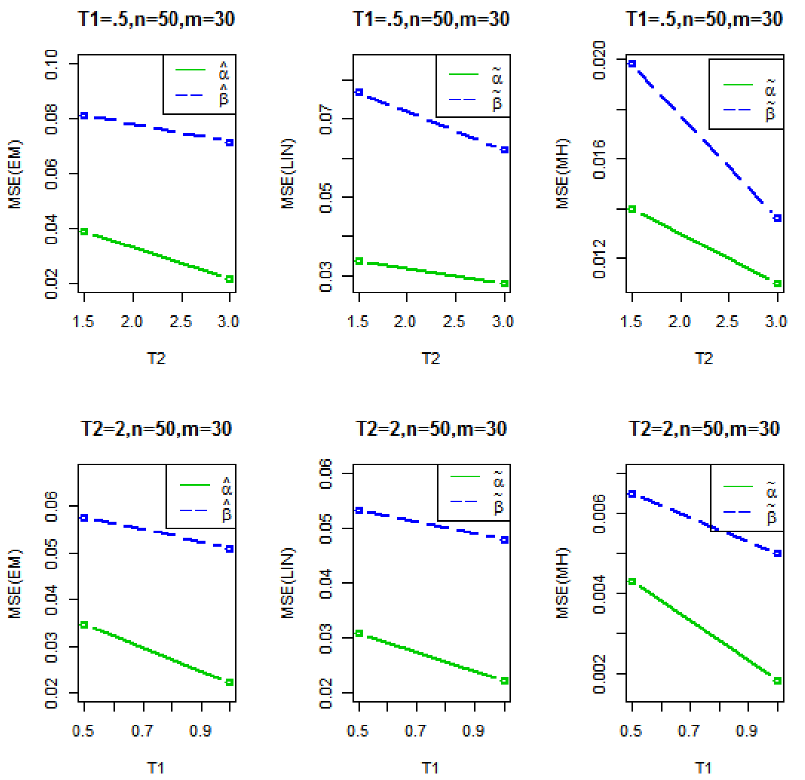


Figure 5: The MSEs of the estimators for different choices of T_1 and T_2.

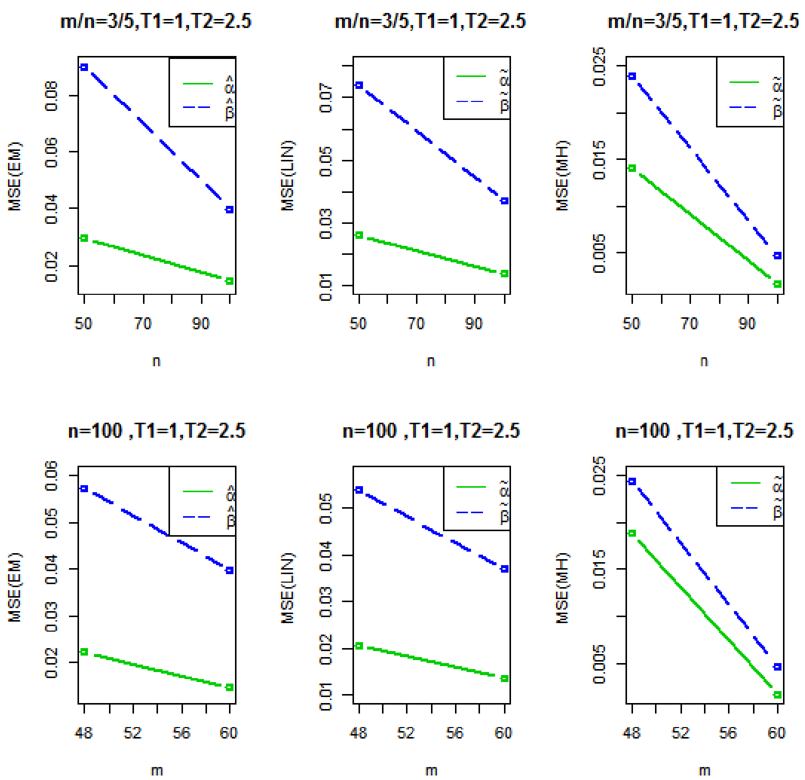


Figure 6: The MSEs of the estimators for different choices of n and m.

Table 3: The average upper and lower bounds for α when $T_1 = 1$, $T_2 = 2.5$.

n	m	$Scheme$	$LACI$	$UACI$	PR	$LHPD$	$UHPD$
30	15	R_1	1.1432	1.8865	IP	1.1675	1.8536
					NIP	1.1604	1.8653
		R_2	0.6884	1.4893	IP	0.7476	1.4280
					NIP	0.7421	1.4452
		R_3	0.7189	1.3966	IP	0.7728	1.3609
					NIP	0.7496	1.3648
50	24	R_1	1.1820	1.7020	IP	1.1224	1.5726
					NIP	1.1148	1.5921
		R_2	0.8440	1.3540	IP	0.9980	1.3932
					NIP	0.9972	1.4266
		R_3	0.7850	1.2893	IP	0.9782	1.3780
					NIP	0.9760	1.3882
	30	R_1	0.9334	1.4357	IP	0.9855	1.3921
					NIP	0.9599	1.3989
		R_2	0.9233	1.3585	IP	0.9315	1.2667
					NIP	0.9035	1.2833
		R_3	0.9540	1.3585	IP	0.8993	1.2264
					NIP	0.8556	1.2345
100	48	R_1	0.9420	1.4441	IP	1.1434	1.5134
					NIP	1.1139	1.5173
		R_2	0.8021	1.2721	IP	1.0019	1.3431
					NIP	1.0011	1.3564
		R_3	0.8527	1.2427	IP	1.0674	1.3174
					NIP	1.0221	1.3308
	60	R_1	0.9378	1.3876	IP	1.0319	1.3108
					NIP	0.9881	1.3288
		R_2	0.8937	1.2437	IP	0.9735	1.2899
					NIP	0.9711	1.3004
		R_3	0.7860	1.2334	IP	0.9989	1.2206
					NIP	0.9366	1.2371

Table 4: The average upper and lower bounds for β when $T_1 = 1$, $T_2 = 2.5$.

n	m	$Scheme$	$LACI$	$UACI$	IP	$LHPD$	$UHPD$
30	15	R_1	0.5265	1.7832	IP	0.6643	1.7548
					NIP	0.6532	1.7802
		R_2	0.9568	2.1005	IP	1.0944	1.9136
					NIP	0.9867	1.9197
		R_3	1.2147	2.2715	IP	0.6957	1.6257
					NIP	0.6809	1.6441
50	24	R_1	0.5669	1.6790	IP	0.5997	1.6212
					NIP	0.5911	1.63211
		R_2	1.0828	2.0995	IP	0.8532	1.6834
					NIP	0.8498	1.6980
		R_3	1.2147	2.2715	IP	0.6957	1.6257
					NIP	0.6709	1.6299
	30	R_1	0.7412	1.8433	IP	1.5999	2.3618
					NIP	1.5799	1.3654
		R_2	1.0339	1.9803	IP	1.4956	2.1097
					NIP	1.4832	2.1217
		R_3	1.1319	2.1708	IP	1.0911	1.6639
					NIP	1.0783	1.6823
100	48	R_1	0.9202	1.6397	IP	1.1000	2.0211
					NIP	1.0906	2.0466
		R_2	1.3920	2.1866	IP	1.2016	1.6581
					NIP	1.1923	1.6734
		R_3	0.9736	1.6248	IP	1.2079	1.6297
					NIP	1.1996	1.6500
	60	R_1	0.9922	1.6392	IP	1.4573	1.9612
					NIP	1.4524	1.9903
		R_2	1.1871	1.9002	IP	1.5417	1.8247
					NIP	1.5289	1.8357
		R_3	1.3432	1.9515	IP	1.4902	1.8750
					NIP	1.4599	1.8786

5. APPLICATIONS OF BIII DISTRIBUTION TO SEPARATION OF SEWER SOLIDS

A real set of experimental data contains the invert trap efficiency. The invert traps are used to separate suspended solids in the sewers and storm water drainage channels. The solid particles are deposited in the bottom of the sewer drainage channel and decreases the channel cross section and thus reduces the hydraulic efficiency. Therefore, increasing invert trap efficiency directly affects the hydraulic efficiency. For computational convenience we divided each data point by 70. Figure 7(a) shows the velocity stream lines of water in channel. The color of the velocity stream lines shows that the velocity decreases in the trap, so the particles entering the low-velocity zone of the invert trap settle in the bottom of the trap. Figure 7(b) shows 3D view of an open rectangular channel fitted with an invert trap at the bottom of the channel. Before we carry out numerical calculations and give way to an advanced point in the analysis of this data, we compute the Kolmogorov-Smirnov (K-S) distances between the empirical distribution and the fitted distribution functions based on MLEs, it is 0.1189, and the associated p-value is 0.8312. We also presented the P-P and CDF (the empirical function and the fitted function) plots for the fitted inverse Burr distribution in Figures 8 and 9 respectively. The result indicates that considered distribution can be used to to obtain inferential results from the considered data set. We have obtained the MLEs by using EM algorithm by taking initial values with the help of contour and 3D profile plot given in Figure 10.

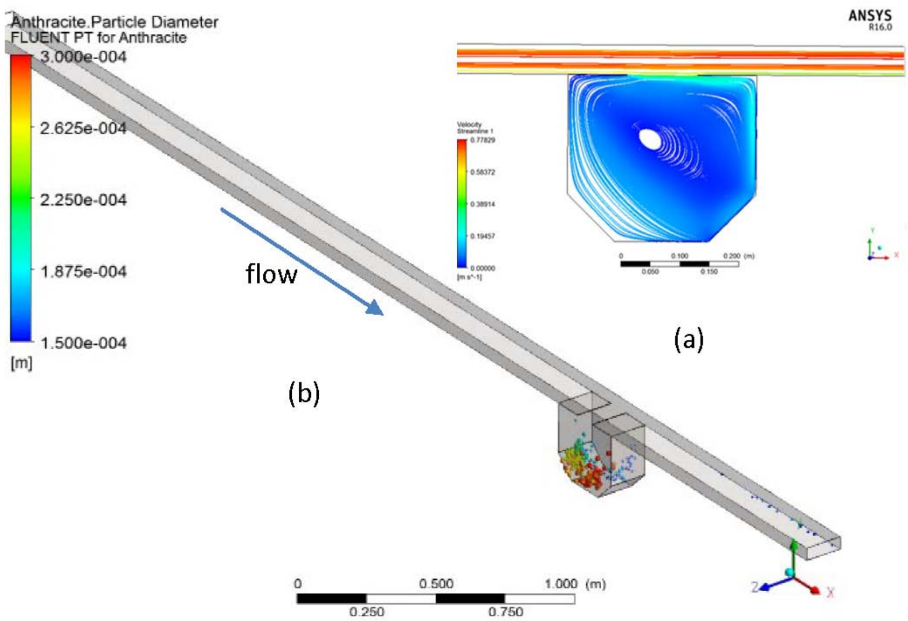


Figure 7: (a) stream lines of water in invert trap. (b) trapping of sewer solids, flowing into a sewer drainage system. Particle traces coloured according to the particle size of 150–300 micron.

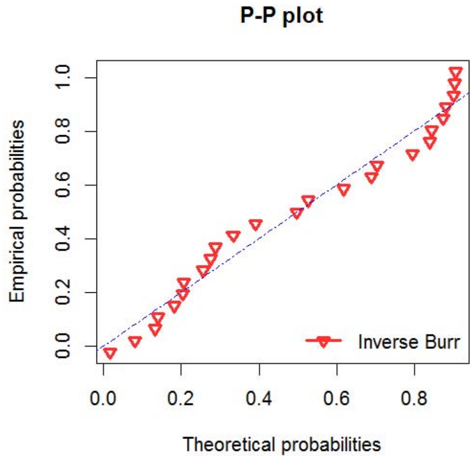


Figure 8: The P-P plot.

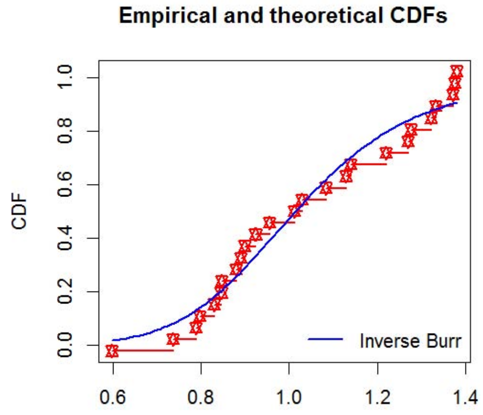


Figure 9: The CDF plot.

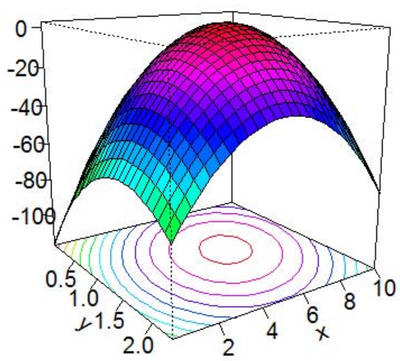


Figure 10: Contour plot and 3D profile plot of log likelihood for invert trap data (x=α and y=β).

We shall use these data to consider three different GIIPH censoring schemes:

Case I: $n = 25, m = 20$ $R = (5, 0 * 19)$,$T_1 = 1.4$ and $T_2 = 1.6$;

Case II: $n = 25, m = 20$ $R = (5, 0 * 19)$,$T_1 = 1.2$ and $T_2 = 1.6$;

Case III: $n = 25, m = 20$ $R = (5, 0 * 19)$,$T_1 = 1.2$ and $T_2 = 1.35$.

Based on the following censoring schemes, the MLEs and Bayes estimates of both the unknown parameters are reported in Table 5.

The length of approximate intervals (LAC) and HPD intervals (LHPD) are also calculated individually and presented in Table 5. For Bayesian aspect, we use non-informative Gamma priors ($a_1 = 0.0001; a_2 = 0.0001; b_1 = 0.0001; b_2 = 0.0001$) due to the lack prior information.

As seen in Table 5, two types of point estimates of parameters are observed: MLEs and Bayes estimates are quite similar. Comparing approximate and credible intervals derived from Bayesian method, the latter are noticeably smaller in interval lengths than the former.

Table 5: Different point and interval estimates of α and β for $(n, m) = (25, 20)$.

Cases	T_1	T_2	$\hat{\alpha}$	$\tilde{\alpha}_{LIN}$	$\tilde{\alpha}_{MH}$	*LAC*	*LHPD*
Case I	1.4	1.6	0.99817	1.03622	0.94491	1.2327	0.8574
Case II	1.2	1.6	1.20987	1.23875	1.15643	1.3734	0.8897
Case III	1.2	1.35	1.21124	1.24054	1.15991	1.3798	0.8687
Cases	T_1	T_2	$\hat{\beta}$	$\tilde{\beta}_{LIN}$	$\tilde{\beta}_{MH}$	*LAC*	*LHPD*
Case I	1.4	1.6	7.51460	7.74952	7.46857	2.7632	1.2864
Case II	1.2	1.6	7.66071	7.79034	7.65335	2.8395	1.3323
Case III	1.2	1.35	7.62534	7.75890	7.44881	2.8567	1.3682

6. CONCLUSIONS

In this paper, we derived the different point and interval estimators of the inverse Burr distribution based on a newly proposed censoring scheme known as generalized progressive hybrid censoring, where experimenters are allowed more flexibility in designing the test, leading to shorter experimental periods and higher efficiency. We obtained the maximum likelihood estimates using the EM algorithm. The observed Fisher information matrix is used to construct the asymptotic confidence intervals of the unknown parameters. Moreover, the Bayesian approach is investigated with a flexible prior distribution, since Bayesian estimation cannot be derived in closed form, two approximations say Lindley's approximation and Metropolis–Hastings algorithm are utilized to achieve approximate point estimates. Using these MCMC samples, the HPD credible intervals are also constructed. The numerical experiments are carried out to evaluate the performance of proposed point and interval estimators, and some conclusions can be drawn from the results that the Bayesian method is comparatively favorable compared to considered classical method. The applicability of the inverse Burr distribution in real situation has been illustrated based on the separation of sewer solids data and it was observed that the proposed distribution can be utilized for analyzing this data well.

ACKNOWLEDGMENTS

Authors are heartily thankful to the editor(s) and learned referees for their useful and insightful comments which led to improve the presentation of the paper.

REFERENCES

[1] Abd-Elfattah, A.M. and Alharbey, A.H. (2012). Bayesian estimation for Burr distribution type III based on trimmed samples, *International Scholarly Research Network*, 250–393.

[2] Altindag, O.; Cankaya, M.N.; Yalcinkaya, A. and Aydogdu, H. (2017). Statiatical infeence for the Burr type III distrribution under type II censored data, *Communications Faculty of Sciences University of Ankara*, **66**, 297–310.

[3] Ateya, Saieed F. (2015). Estimation under Inverse Weibull Distribution based on Balakrishnan's Unified Hybrid Censored Scheme, *Communications in Statistics Simulation and Computation*, **46**, 3645–3666.

[4] Banerjee, A. and Kundu, D. (2008). Inference based on type-II hybrid censored data from Weibull distribution, *IEEE Transactions on Reliability*, **57**, 369–378.

[5] Balakrishnan, N.; Rasouli, A.; Sanjari, N. and Farsipour, S. (2008). Exact likelihood inference based on an unified hybrid censored sample from the exponential distribution, *Journal of Statistical Computation and Simulation*, **78**, 475–488.

[6] Burr, I.W. (1942). Cumulative frequency functions, *The Annals of Mathematical Statistics*, **13**(2), 215–232.

[7] Buxton, A.; Tait, S.; Stovin, V. and Saul, A. (2002). Developments in a methodology for the design of engineered invert traps in combined sewer systems, *Water Science and Technology*, **45**(7), 133–142.

[8] Congdon, P. (2014). *Applied Bayesian Modelling*, Vol. 595, John Wiley and Sons.

[9] Chan, P.; Ng, H. and Su, F. (2015). Exact likelihood inference for the two-parameter exponential distribution under type II progressively hybrid censoring, *Metrika*, **6**, 747–770.

[10] Childs, A.; Chandrasekhar, B.; Balakrishnan, N. and Kundu, D. (2003). Exact likelihood inference based on type I and type II hybrid censored samples from the exponential distribution, *Annals of the Institute of Statistical Mathematics*, **55**, 319–330.

[11] Cho, Y.; Sun, H. and Lee, K. (2015b). Exact likelihood inference for an exponential parameter under generalized progressive hybrid censoring scheme, *Statistical Methodology*, **23**, 18–34.

[12] El-Sherpieny, E.A.; Almetwally, E.M. and Muhammed, H.Z., (2020). Progressive Type-II hybrid censored schemes based on maximum product spacing with application to Power Lomax distribution, *Physica A: Statistical Mechanics and its Applications*, **553**, 124251.

[13] Epstein, B. (1954). Truncated life tests in the exponential case, *Ann. Math. Stat.*, **25**, 555–564.

[14] Gupta, P.K. and Singh, B. (2012). Parameter estimation of Lindley distribution with hybrid censored data, *International Journal of System Assurance Engineering and Management*, **1**, 1–8.

[15] Gorny, J. and Cramer, E. (2016). Exact likelihood inference for exponential distributions under generalized progressive hybrid censoring schemes, *Statistical Methodology*, **29**, 70–94.

[16] Cho, Y.; Sun, H. and Lee, K. (2015). Exact likelihood inference for an exponential parameter under generalized progressive hybrid censoring scheme, *Statistical Methodology*, **23**, 18–34.

[17] Kundu, D. and Howlader, H. (2010). Bayesian inference and prediction of the inverse Weibull distribution for type-II censored data, *Computational Statistics and Data Analysis*, **54**, 1547–1558.

[18] Koley, A. and Kundu, D. (2017). On generalized progressive hybrid censoring in presence of competing risks, *Metrika*, **80**, 401–426.

[19] LINDLEY, D.V. (1980). Approximate Bayesian method, *Trabajos de Estadistica*, **31**, 223–245.

[20] METROPOLIS, N.; ROSENBLUTH, A.W.; ROSENBLUTH, M.N.; TELLER, A.H. and TELLER, E. (1953). Equations of state calculations by fast computing machines, *Chemmical Physics*, **21**, 1087–1091.

[21] MOHSIN, M. and KAUSHAL, D.R. (2017). Three-dimensional computational fluid dynamics (volume of fluid) modelling coupled with a stochastic discrete phase model for the performance analysis of an invert trap experimentally validated using field sewer solids, *Particuology*, **33**, 98–111.

[22] PANAHI, H. (2017). Estimation methods for the generalized inverted exponential distribution under type II progressively hybrid censoring with application to spreading of micro-drops data, *Communications in Mathematics and Statistics*, **5**, 159–174.

[23] PANAHI, H. and ASADI, S. (2021). On adaptive progressive hybrid censored Burr type III distribution: application to the nano droplet dispersion data, *Quality Technology and Quantitative Management*, **18**(2), 179–201.

[24] SINGH, D.P.; TRIPATHI, Y.M.; RASTOGI, M.K. and DABRAL, N. (2016). Estimation and prediction for a Burr III distribution with progressive censoring, *Communication in Statistics – Theory and Methods*, **46**, 44–59.

[25] THINGLAS, T. and KAUSHAL, D.R. (2008). Three-dimensional CFD modeling for optimization of invert trap configuration to be used in sewer solids management, *Particulate Science and Technology*, **26**(5), 507–519.

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 369–385
https://doi.org/10.57805/revstat.v22i3.516

Bayesian Sampling Plan for Weibull Distribution with Type II Hybrid Censoring under Random Decision Function

Authors: Hacène Belbachir
– USTHB, Faculty of Mathematics, RECITS Laboratory,
BP 32 El-Alia Bab-Ezzouar 16111, Algiers, Algeria
hacenebelbachir@gmail.com

Mohamed Benahmed ✉
– USTHB, Faculty of Mathematics, RECITS Laboratory,
BP 32 El-Alia Bab-Ezzouar 16111, Algiers, Algeria
– Higher National School of Statistics and Applied Economics,
Koléa 42003, Tipaza, Algeria
– CERIST, Research Center on Scientific and Technical Information,
Ben Aknoun 16028, Algiers, Algeria
krmbmed@gmail.com

Received: November 2020 Revised: August 2022 Accepted: August 2022

Abstract:

- This article studies the problem of conception of a Bayesian single variable sampling plan for Weibull distribution under type II hybrid censoring based on two-sided decision function with a linear random doubt zone. Based on an appropriate loss function, an explicit expression for the Bayes risk is determined. To find the optimal sampling plans, a simple algorithm based on the grid search method is described. Finally, simulation study is given to illustrate the proposed model.

Keywords:

- *Bayesian sampling; type II hybrid censoring; loss function; random decision function; grid search method.*

AMS Subject Classification:

- 62D05, 62F15, 62N05.

✉ Corresponding author

1. INTRODUCTION

Quality control is one of the most important issues of the modern industry, to determine whether the quality of the products or process is satisfactory according to certain criteria established in advance. We distinguish two types of control, the control during production: which is the one carried out at different stages during the production process, and the reception control: which is the one carried out by the producer or the consumer during the inspection of a finished product, which also requires taking sampling plans. There have been several criteria to construct sampling plans. Criteria based on decision theory are the most efficient for quality control, in the sense that the sampling plan is determined by taking an optimal decision. Numerous study have investigated along with this approach, we refer to [9, 11, 19, 10].

Recently, a number of studies have investigated Bayesian sampling plans based on the lifetime censored data. Readers are referred to the sampling plan based on type II censored sample [12] and [5], sampling plan based on type I censored sample [13] and [18], interval censored sample [6]. The type I hybrid censored sample was initially introduced in [8]. In [7] the exact distribution of the maximum likelihood estimator (MLE) of the expected lifetime is provided where the lifetime of components follows exponential distribution under type I and type II hybrid censoring. Reference [14] have studied sampling plans under type I and type II hybrid censoring for quadratic loss function based on the results of [7]. Furthermore, a Bayesian sampling plan based on type I hybrid censored samples has been developed in [15] using a conventional one-sided decision function. Modified type II hybrid censoring has been provided by [20]. For exponential distribution under type I censoring and type I hybrid censoring a new shrinkage estimator for the expected lifetime has been studied in [17], which always exists even if no failure occurs at the termination time. In addition, Reference [17] provided that the construction of the Bayes decision function (as in [20], [15]), which is based on the posterior expectation, becomes more difficult if the loss function is not polynomial.

In some industrial process, the quality characteristics data are derived from a complex production process or from an uncertain environment. Much acceptance sampling plans have been proposed under this situation, [2, 3] have developed acceptance sampling plan for variable and attribute using the neutrosophic statistics. [4] discussed a Bayesian sampling plan under two-sided decision function based on linear random doubt zone.

In this work, we develop a Bayesian single variable sampling plan for Weibull distribution based on the modified type II hybrid censored sample under random decision function. However, we generalize the work of [4] into two valuable issues. The first issue, the Weibull distribution, which is frequently used in life testing due to flexibility in term of hazard function (see e.g. [1]), and with the commonly used of other distributions as special cases, such as the exponential and Rayleigh distributions. The second issue, the type II hybrid censoring which is a generalization of type II censoring. The type II hybrid censoring has the advantage that at least m failures or more can be observed at the censoring time, which leads to significant efficiency of the model. The rest of this paper is organized in the following way. In Section 2, we provide the proposed random decision function and all necessary assumptions. In Sections 3 and 4, we obtain an explicit expression for the Bayes risk using a polynomial and non polynomial loss respectively. A simple algorithm based on the grid search method to obtain an optimal sampling plan is provided in Section 5. In Section 6, we give numerical examples for the polynomial and non polynomial loss functions followed by some remarks. We finish by a conclusion in Section 7.

2. FORMULATION OF THE PROBLEM

Suppose that we have a batch of items prepared for inspection. The lifetime of each item is a random variable X which follows a Weibull distribution $W(\lambda, \mu)$:

$$f(x|\lambda,\mu) = \begin{cases} \lambda\mu x^{\mu-1}\exp(-\lambda x^{\mu}), & \text{for } x \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$

with the shape parameter μ is known and the scale parameter λ is unknown. It is easy to show that X^{μ} follows an exponential distribution with expected lifetime $1/\lambda$. Further, We assume that λ has a prior distribution $\Gamma(\alpha, \beta)$ where α and β are known, with the pdf:

$$g(\lambda;\alpha,\beta) = \begin{cases} \lambda^{\alpha-1}\exp(-\beta\lambda)\beta^{\alpha}/\Gamma(\alpha), & \text{for } \lambda > 0, \\ 0, & \text{otherwise.} \end{cases}$$

Given a random sample of size n, taken from a batch for life testing. Assume that the modified type II hybrid censoring is adopted. Let $\underline{X} = (X_{(1)}, X_{(2)}, ..., X_{(n)})$ be the order statistic of sample $(X_1, X_2, ..., X_n)$, the life test terminates at the random time $\tau_{n,m} = \min\{max(X_{(m)}, t), X_{(n)}\}$ with $m \leq n$. The likelihood function in this case is given by:

$$l(\underline{X}|\lambda) = \begin{cases} \dfrac{n!(\lambda\mu)^m \prod_{i=1}^{m} X_{(i)}^{\mu}}{(n-m)!} e^{-\lambda\left(\sum_{i=1}^{m} X_{(i)}^{\mu} + (n-m)X_{(m)}^{\mu}\right)} & \text{for } D = 0, 1, ..., m-1, \\ \dfrac{n!(\lambda\mu)^D \prod_{i=1}^{D} X_{(i)}^{\mu}}{(n-D)!} e^{-\lambda\left(\sum_{i=1}^{D} X_{(i)}^{\mu} + (n-D)t^{\mu}\right)} & \text{for } D = m, m+1, ..., n \end{cases}$$

where D represents the number of observed failures that occur before time t.
Then, the MLE of $\theta = 1/\lambda$ is given by:

$$\hat{\theta} = \begin{cases} \frac{\sum_{i=1}^{m} X_{(i)}^{\mu} + (n-m)X_{(m)}^{\mu}}{m}, & \text{for } D = 0, 1, ..., m-1, \\ \frac{\sum_{i=1}^{D} X_{(i)}^{\mu} + (n-D)t^{\mu}}{D}, & \text{for } D = m, m+1, ..., n, \end{cases} \tag{2.1}$$

According to [7], the exact distribution of the MLE of θ:

$$f_{\hat{\theta}}(y) = \sum_{d=0}^{n}\sum_{j=0}^{d}(-1)^j \binom{n}{d}\binom{d}{j} e^{-\lambda t^{\mu}(n-d+j)} g(y - a_{j,M}; M, \lambda M). \tag{2.2}$$

where $a_{j,M} = (n-d+j)t^{\mu}/M$, and $M = max\{d, m\}$.

Let C_s, C_t and C_r be positive constants and represent respectively the unit inspection cost, the cost per unit of time used for the test and the loss due to rejection of the batch. Let $a_0 + a_1\lambda + \cdots + a_k\lambda^k$ denote the loss of accepting the batch and be positive and increasing in λ. When the life test is interrupted, the unfailures items can be reused and therefore have the salvage value v_s, where $0 < v_s < C_s$, then the loss function is defined as follows:

$$L(\lambda, \delta(\underline{x})) = \begin{cases} nC_s - (n - D_{n,m})v_s + C_t\tau_{n,m} + \sum_{i=0}^{k} a_i\lambda^i, & \text{for } \delta(\underline{x}) = d_0, \\ nC_s - (n - D_{n,m})v_s + C_t\tau_{n,m} + C_r, & \text{for } \delta(\underline{x}) = d_1, \end{cases} \tag{2.3}$$

where d_0 and d_1 represent the decisions of accepting and rejecting the batch respectively. The random variable $D_{n,m}$ denotes the number of failures that occur before the termination time $\tau_{n,m}$. $\delta(\underline{x})$ is the decision function which depends on the observation failures $\underline{x} = (x_{(1)}, x_{(2)}, ..., x_{(n)})$. We propose the following two-sided decision function:

$$\delta(\underline{x}) = \begin{cases} d_0, & \text{for } \hat{\theta} \geq T_0, \\ \begin{cases} d_1, & \text{with probability } p_\theta \\ d_0, & \text{with probability } 1 - p_\theta \end{cases} & \text{for } T_1 \leq \hat{\theta} < T_0, \\ d_1, & \text{for } \hat{\theta} < T_1, \end{cases} \tag{2.4}$$

where $p_\theta = \frac{T_0 - \hat{\theta}}{T_0 - T_1}$, and $0 < T_1 < T_0$. Note that, the decision function in Equation (2.4) is described similarly as in [4].

3. COMPUTATION OF THE BAYES RISK

Based on the decision function $\delta(\underline{x})$, the Bayes risk can be computed as follows:

$$\begin{aligned} R(n, m, t, T_0, T_1) &= E\{E[L(\lambda, \delta(\underline{x}))]\} \\ &= E\left\{E\left[nC_s + C_t\tau_{n,m} - (n - D_{n,m})v_s + d_1C_r + (1 - d_1)\sum_{i=0}^{k} a_i\lambda^i|\lambda\right]\right\} \\ &= n(C_s - v_s) + v_sE\{E[D_{n,m}|\lambda]\} + C_tE\{E[\tau_{n,m}|\lambda]\} + \sum_{i=0}^{k} a_i\gamma_i \\ &\quad + E\left\{E\left[d_1\sum_{i=0}^{k} \omega_i\lambda^i|\lambda\right]\right\} \\ &= n(C_s - v_s) + v_sE\{E[D_{n,m}|\lambda]\} + C_tE\{E[\tau_{n,m}|\lambda]\} + \sum_{i=0}^{k} a_i\gamma_i + r(n, m|d_1), \end{aligned}$$

here γ_i represents the i-th moment of λ, and

$$\omega_i = \begin{cases} C_r - a_0, & \text{for } i = 0, \\ -a_i & \text{for } i = 1, ..., k. \end{cases} \tag{3.1}$$

Such as

$$r(n,m|d_1) = E\Big\{E\Big[\sum_{i=0}^{k}\omega_i\lambda^i d_1|\lambda\Big]\Big\} = E\Big\{\sum_{i=0}^{k}\omega_i\lambda^i E\Big[I_{\hat\theta<T_1} + p_\theta I_{T_1\le\hat\theta<T_0}|\lambda\Big]\Big\}$$
$$= \sum_{i=0}^{k}\omega_i \int_0^\infty \tfrac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\lambda}\lambda^{\alpha+i-1}\Bigg[\int_0^{T_1} f_{\hat\theta}(y)dy + \int_{T_1}^{T_0}\tfrac{T_0-y}{T_0-T_1} f_{\hat\theta}(y)dy\Bigg]d\lambda$$
$$= \sum_{d=0}^{n}\sum_{j=0}^{d}\sum_{i=0}^{k}(-1)^j\omega_i\tbinom{n}{d}\tbinom{d}{j}\int_0^\infty\Bigg[\int_{a_{j,M}}^{T_1}\tfrac{\beta^\alpha M^M\left(y-a_{j,M}\right)^{M-1}}{\Gamma(\alpha)\Gamma(M)}e^{-(\beta+My)\lambda}\lambda^{\alpha+M+i-1}dy$$
$$+\int_{T_1}^{T_0}\tfrac{T_0-y}{T_0-T_1}\tfrac{\beta^\alpha M^M\left(y-a_{j,M}\right)^{M-1}}{\Gamma(\alpha)\Gamma(M)}e^{-(\beta+My)\lambda}\lambda^{\alpha+M+i-1}dy\Bigg]d\lambda$$
$$= \sum_{d=0}^{n}\sum_{j=0}^{d}\sum_{i=0}^{k}\tfrac{\beta^\alpha M^M\Gamma(M+\alpha+i)}{\Gamma(\alpha)\Gamma(M)}\Bigg[\int_0^{T_1-a_{j,M}}\tfrac{y^{M-1}}{\left(\beta+Ma_{j,M}+My\right)^{\alpha+M+i}}dy$$
$$+\int_{T_1-a_{j,M}}^{T_0-a_{j,M}}\tfrac{T_0-y-a_{j,M}}{T_0-T_1}\tfrac{y^{M-1}}{\left(\beta+Ma_{j,M}+My\right)^{\alpha+M+i}}dy\Bigg],$$

Using $z = \dfrac{My}{My+\beta+Ma_{j,M}}$ we obtain

$$r(n,m|d_1)$$
$$= \sum_{d=0}^{n}\sum_{j=0}^{d}\sum_{i=0}^{k}(-1)^j\omega_i\tbinom{n}{d}\tbinom{d}{j}\tfrac{\beta^\alpha\Gamma(M+\alpha+i)}{\Gamma(\alpha)\Gamma(M)\left(\beta+Ma_{j,M}\right)^{\alpha+i}}\Bigg[\int_0^{q_1}z^{M-1}(1-z)^{\alpha+i-1}dz$$
$$+\tfrac{T_0-a_{j,M}}{T_0-T_1}\int_{q_1}^{q_0}z^{M-1}(1-z)^{\alpha+i-1}dz - \tfrac{\beta+Ma_{j,M}}{T_0-T_1}\int_{q_1}^{q_0}z^{M-1}(1-z)^{\alpha+i-1}dz\Bigg]$$
$$= \sum_{d=0}^{n}\sum_{j=0}^{d}\sum_{i=0}^{k}\tfrac{(-1)^j\omega_i\binom{n}{d}\binom{d}{j}\beta^\alpha\Gamma(\alpha+i)}{\Gamma(\alpha)\left(\beta+Ma_{j,M}\right)^{\alpha+i}}\Big\{I_{q_1}(M,\alpha+i) + \tfrac{T_0-a_{j,M}}{T_0-T_1}[I_{q_0}(M,\alpha+i)-$$
$$I_{q_1}(M,\alpha+i)] - \tfrac{\beta+Ma_{j,M}}{(\alpha+i-1)(T_0-T_1)}[I_{q_0}(M+1,\alpha+i-1) - I_{q_1}(M+1,\alpha+i-1)]\Big\},$$

where $q_i = \frac{M\left(T_i-a_{j,M}\right)}{\beta+M\left(T_i-a_{j,M}\right)+Ma_{j,M}}$. $B_x(a,b)$ and $I_x(a,b)$ denote the incomplete Beta function and the cdf of Beta distribution respectively.

Hence, the Bayes risk $R(n,m,t,T_0,T_1)$ can be expressed as:

$$R(n,m,t,T_0,T_1) \tag{3.2}$$
$$= \sum_{d=0}^{n}\sum_{j=0}^{d}\sum_{i=0}^{k}\tfrac{(-1)^j\omega_i\binom{n}{d}\binom{d}{j}\beta^\alpha\Gamma(\alpha+i)}{\Gamma(\alpha)\left(\beta+Ma_{j,M}\right)^{\alpha+i}}\Big\{I_{q_1}(M,\alpha+i) + \tfrac{T_0-a_{j,M}}{T_0-T_1}[I_{q_0}(M,\alpha+i)-$$
$$I_{q_1}(M,\alpha+i)] - \tfrac{\beta+Ma_{j,M}}{(\alpha+i-1)(T_0-T_1)}[I_{q_0}(M+1,\alpha+i-1) - I_{q_1}(M+1,\alpha+i-1)]\Big\}$$
$$+ n(C_s-v_s) + v_s\sum_{d=0}^{n}\sum_{j=0}^{d}(-1)^{d-j}M\tbinom{n}{d}\tbinom{d}{j}\Big(\tfrac{\beta}{\beta+(n-j)t^\mu}\Big)^\alpha + \sum_{i=0}^{k}a_i\gamma_i + \tau^* C_t,$$

where, for $m < n$

$$\begin{aligned}\tau^* &= E\{E[\tau_{n,m}|\lambda]\}\\ &= m\binom{n}{m}\sum_{j=0}^{m-1}(-1)^{m-j-1}\binom{m-1}{j}\frac{\alpha\beta^{1/\mu}}{(n-j)^{1+1/\mu}}B_{1-q*}\Big(1+\tfrac{1}{\mu},\alpha-\tfrac{1}{\mu}\Big)\\ &\quad+\frac{tn!}{(m-1)!(n-m-1)!}\sum_{i=0}^{m-1}\sum_{j=0}^{n-m-1}\Big[(-1)^{n-i-j}\binom{m-1}{i}\binom{n-m-1}{j}\\ &\quad\times\frac{\beta^\alpha}{(m+j-i)(n-m-j)}\Big(\frac{1}{((n-m-j)t^\mu+\beta)^\alpha}-\frac{1}{((n-i)t^\mu+\beta)^\alpha}\Big)\Big]\\ &\quad+n\sum_{j=0}^{n-1}(-1)^{n-j-1}\binom{n-1}{j}\frac{\alpha\beta^{1/\mu}}{(n-j)^{1+1/\mu}}B_{q*}\Big(1+\tfrac{1}{\mu},\alpha-\tfrac{1}{\mu}\Big),\end{aligned}$$

and, for $m = n$

$$\tau^* = E\{E[\tau_{n,m}|\lambda]\} = n\alpha\beta^{1/\mu}B\Big(1+\tfrac{1}{\mu},\alpha-\tfrac{1}{\mu}\Big)\sum_{j=0}^{n-1}(-1)^j\binom{n-1}{j}\frac{1}{(j+1)^{1+1/\mu}},$$

with $q* = \frac{(n-j)t^\mu}{\beta+(n-j)t^\mu}$. The computation of $E\{E[D_{n,m}|\lambda]\}$ and $E\{E[\tau_{n,m}|\lambda]\}$ is provided in the appendix.

4. BAYES RISK FOR NON-POLYNOMIAL LOSS FUNCTION

In this section we provide an explicit expression for the Bayes risk under non-polynomial loss function, which can be written as:

$$L_{NP}(\lambda,\delta(\underline{x})) = \begin{cases} nC_s-(n-D_{n,m})v_s+C_t\tau_{n,m}+\exp(c\lambda)-c\lambda-1, & \text{for } \delta(\underline{x})=d_0,\\ nC_s-(n-D_{n,m})v_s+C_t\tau_{n,m}+C_r, & \text{for } \delta(\underline{x})=d_1,\end{cases} \tag{4.1}$$

where the loss of accepting the batch $\exp(c\lambda) - c\lambda - 1$ is of the form LINEX loss (see e.g. [1, 16]). The value of c must be positive for ensuring that, the loss of accepting the batch is increasing in λ.

$$\begin{aligned}&R_{NP}(n,m,t,T_0,T_1) = E\{E[L_{NP}(\lambda,\delta(\underline{x}))]\}\\ &= E\{E[nC_s+C_t\tau_{n,m}-(n-D_{n,m})v_s+d_1C_r+(1-d_1)(\exp(c\lambda)-c\lambda-1)|\lambda]\}\\ &= n(C_s-v_s)+v_sE\{E[D_{n,m}|\lambda]\}+C_tE\{E[\tau_{n,m}|\lambda]\}+\Big(\tfrac{\beta}{\beta-c}\Big)^\alpha-c\tfrac{\alpha}{\beta}-1\\ &\quad+E\{E[d_1(C_r+1+c\lambda-\exp(c\lambda))|\lambda]\}\\ &= n(C_s-v_s)+v_sE\{E[D_{n,m}|\lambda]\}+C_tE\{E[\tau_{n,m}|\lambda]\}+\Big(\tfrac{\beta}{\beta-c}\Big)^\alpha-c\tfrac{\alpha}{\beta}-1+r'(n,m|d_1),\end{aligned}$$

with

$$\begin{aligned}r'(n,m|d_1) &= E\{E[d_1(C_r+1+c\lambda-\exp(c\lambda))|\lambda]\}\\ &= E\Big\{(C_r+1+c\lambda-\exp(c\lambda))E\Big[I_{\hat\theta<T_1}+p_\theta I_{T_1\le\hat\theta<T_0}|\lambda\Big]\Big\}\\ &= \sum_{i=0}^{1}\omega_i'\int_0^\infty\frac{\beta^\alpha}{\Gamma(\alpha)}e^{-\beta\lambda}\lambda^{\alpha+i-1}\left[\int_0^{T_1}f_{\hat\theta}(y)dy+\int_{T_1}^{T_0}\frac{T_0-y}{T_0-T_1}f_{\hat\theta}(y)dy\right]d\lambda\\ &\quad-\int_0^\infty\frac{\beta^\alpha}{\Gamma(\alpha)}e^{-(\beta-c)\lambda}\lambda^{\alpha-1}\left[\int_0^{T_1}f_{\hat\theta}(y)dy+\int_{T_1}^{T_0}\frac{T_0-y}{T_0-T_1}f_{\hat\theta}(y)dy\right]d\lambda,\end{aligned}$$

where $\omega_0' = C_r + 1$, $\omega_1' = c$.

From the previous section, we have

$$
\begin{aligned}
&\int_0^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-(\beta-c)\lambda}\lambda^{\alpha-1}\left[\int_0^{T_1} f_{\hat\theta}(y)dy + \int_{T_1}^{T_0}\frac{T_0-y}{T_0-T_1} f_{\hat\theta}(y)dy\right]d\lambda\\
&=\sum_{d=0}^{n}\sum_{j=0}^{d}(-1)^j\binom{n}{d}\binom{d}{j}\int_0^\infty\left[\int_{a_{j,M}}^{T_1}\frac{\beta^\alpha M^M\left(y-a_{j,M}\right)^{M-1}}{\Gamma(\alpha)\Gamma(M)}e^{-(\beta-c+My)\lambda}\lambda^{\alpha+M-1}dy\right.\\
&\quad\left.+\int_{T_1}^{T_0}\frac{T_0-y}{T_0-T_1}\frac{\beta^\alpha M^M\left(y-a_{j,M}\right)^{M-1}}{\Gamma(\alpha)\Gamma(M)}e^{-(\beta-c+My)\lambda}\lambda^{\alpha+M-1}dy\right]d\lambda\\
&=\sum_{d=0}^{n}\sum_{j=0}^{d}\frac{\beta^\alpha M^M\Gamma(M+\alpha)}{\Gamma(\alpha)\Gamma(M)}\left[\int_0^{T_1-a_{j,M}}\frac{y^{M-1}}{\left(\beta-c+Ma_{j,M}+My\right)^{\alpha+M}}dy\right.\\
&\quad\left.+\int_{T_1-a_{j,M}}^{T_0-a_{j,M}}\frac{T_0-y-a_{j,M}}{T_0-T_1}\frac{y^{M-1}}{\left(\beta-c+Ma_{j,M}+My\right)^{\alpha+M+i}}dy\right]\\
&=\sum_{d=0}^{n}\sum_{j=0}^{d}\frac{(-1)^j\binom{n}{d}\binom{d}{j}\beta^\alpha}{\left(\beta-c+Ma_{j,M}\right)^\alpha}\left\{I_{q_1'}(M,\alpha)+\frac{T_0-a_{j,M}}{T_0-T_1}\left[I_{q_0'}(M,\alpha)-I_{q_1'}(M,\alpha)\right]\right.\\
&\quad\left.-\frac{\beta-c+Ma_{j,M}}{(\alpha-1)(T_0-T_1)}\left[I_{q_0'}(M+1,\alpha-1)-I_{q_1}(M+1,\alpha-1)\right]\right\},
\end{aligned}
$$

with $q_i' = \frac{M\left(T_i-a_{j,M}\right)}{\beta-c+M\left(T_i-a_{j,M}\right)+Ma_{j,M}}$.

Therefore, the Bayes risk expression under the loss function 4.1 is given by:

$$
\begin{aligned}
(4.2)\qquad & R_{NP}(n,m,t,T_0,T_1)\\
&=\sum_{d=0}^{n}\sum_{j=0}^{d}\sum_{i=0}^{1}\frac{(-1)^j\omega_i'\binom{n}{d}\binom{d}{j}\beta^\alpha\Gamma(\alpha+i)}{\Gamma(\alpha)\left(\beta+Ma_{j,M}\right)^{\alpha+i}}\left\{I_{q_1}(M,\alpha+i)+\frac{T_0-a_{j,M}}{T_0-T_1}[I_{q_0}(M,\alpha+i)-\right.\\
&\left.I_{q_1}(M,\alpha+i)]-\frac{\beta+Ma_{j,M}}{(\alpha+i-1)(T_0-T_1)}[I_{q_0}(M+1,\alpha+i-1)-I_{q_1}(M+1,\alpha+i-1)]\right\}\\
&-\sum_{d=0}^{n}\sum_{j=0}^{d}\frac{(-1)^j\binom{n}{d}\binom{d}{j}\beta^\alpha}{\left(\beta-c+Ma_{j,M}\right)^\alpha}\left\{I_{q_1'}(M,\alpha)+\frac{T_0-a_{j,M}}{T_0-T_1}\left[I_{q_0'}(M,\alpha)-I_{q_1'}(M,\alpha)\right]\right.\\
&\quad\left.-\frac{\beta-c+Ma_{j,M}}{(\alpha-1)(T_0-T_1)}\left[I_{q_0'}(M+1,\alpha-1)-I_{q_1}(M+1,\alpha-1)\right]\right\}+n(C_s-v_s)\\
&\quad+v_s\sum_{d=0}^{n}\sum_{j=0}^{d}(-1)^{d-j}M\binom{n}{d}\binom{d}{j}\left(\frac{\beta}{\beta+(n-j)t^\mu}\right)^\alpha+\left(\frac{\beta}{\beta-c}\right)^\alpha-c\frac{\alpha}{\beta}-1+\tau^*C_t.
\end{aligned}
$$

5. NUMERICAL APPROXIMATIONS

The expression of $R(n,m,t,T_0,T_1)$ and $R_{NP}(n,m,t,T_0,T_1)$ are quite complicated, so we cannot get the optimal sampling plan analytically. Using the grid search method we can obtain an optimal sampling plan numerically. As given in [17], we assume that T_0 has an upper bound since $0 < T_0 < T_0^*$, and for t as given in [13], we obtain a confidence interval $[t_L, t_U]$ such that $P(X > t_U) = \eta/2$ and $P(X < t_L) = \eta/2$ where:

$$P(X < t_L) = \int_0^\infty\int_0^{t_L}\frac{\beta^\alpha}{\Gamma(\alpha)}\lambda^{\alpha-1}e^{-\beta\lambda}\lambda x^{\mu-1}e^{-\lambda x^\mu}dxd\lambda = \eta/2,$$

and

$$P(X > t_U) = \int_0^\infty \int_{t_U}^\infty \frac{\beta^\alpha}{\Gamma(\alpha)} \lambda^{\alpha-1} e^{-\beta\lambda} \lambda x^{\mu-1} e^{-\lambda x^\mu} dx d\lambda = \eta/2,$$

hence

$$t_L = \left\{\beta\left[\left(1-\tfrac{\eta}{2}\right)^{-1/\alpha} - 1\right]\right\}^{\frac{1}{\mu}}$$
$$t_U = \left\{\beta\left[\left(\tfrac{\eta}{2}\right)^{-1/\alpha} - 1\right]\right\}^{\frac{1}{\mu}}.$$

5.1. An upper bound for the optimal size sample

To obtain the optimal sampling plan, we provide an upper bound for the optimal sample size, and then the optimal sampling plan can be obtained in a finite number of search steps.

Theorem 5.1. *The optimal sample is bounded by:*

$$N = \min\left\{ \left[\frac{C_r}{C_s - v_s}\right], \left[\frac{\sum_{i=0}^k a_i\gamma_i}{C_s - v_s}\right] \right\}, \tag{5.1}$$

where $[x]$ *is the integer part of* x.

Proof: Let $(0,0,0,0,0)$ and $(0,0,0,\infty,\infty)$ be the sampling plans that accepts and rejects the batch without taking sampling respectively. For (n', m', t', T_0', T_1') an optimal sampling plan, we have $R(n', m', t', T_0', T_1') \le R(0,0,0,0,0) = \sum_{i=0}^k a_i\gamma_i$. and $R(n', m', t', T_0', T_1') \le R(0,0,0,\infty,\infty) = C_r$.

As $n(C_s - v_s) \le R(n', m', t', T_0', T_1')$, therefore

$$n(C_s - v_s) \le \min\left\{ C_r, \sum_{i=0}^k a_i\gamma_i \right\}$$
$$n \le \min\left\{ \left[\frac{C_r}{C_s - v_s}\right], \left[\frac{\sum_{i=0}^k a_i\gamma_i}{C_s - v_s}\right] \right\}.$$

Hence the result. □

Algorithm 5.1. To derive an optimal sampling plan (n', m', t', T_0', T_1') based on the minimization of the Bayes risk, a finite algorithm is described in the following steps:

a) Start with $(n, m, t) = (0, 0, 0)$, compute N from (5.1) and compute $R(0, 0, 0, T_0, T_1)$ $= \min\left\{ R(0,0,0,\infty,\infty) = C_r, R(0,0,0,0,0) = \sum_{i=0}^k a_i\gamma_i \right\}$.

b) For fixed (n, m, t), compute the optimal $T'_{0,(n,m,t)}$ and $T'_{1,(n,m,t)}$ using grid search method, such that $R\left(n, m, t, T'_{0,(n,m,t)}, T'_{1,(n,m,t)}\right) = \min_{0<T_1<T_0\le T^*} R(n, m, t, T_1, T_0)$, with grid size 0.0125.

c) For fixed (n,m), compute the optimal $t'_{(n,m)}$ using grid search method, such that $R\Big(n,m,t'_{(n,m)},T'_{0,(n,m,t)},T'_{1,(n,m,t)}\Big)=\min_{t_L\le t\le t_U} R\Big(n,m,t,T'_{0,(n,m,t)},T'_{1,(n,m,t)}\Big)$, with grid size $\frac{t_U-t_L}{100}$.

d) For $0\le m\le n\le N$, choose (n',m',t',T'_0,T'_1) which corresponds to the smallest value of the Bayes risks $R\Big(n,m,t'_{(n,m)},T'_{0,(n,m,t)},T'_{1,(n,m,t)}\Big)$.

6. AN ILLUSTRATIVE EXAMPLE

To implement the Algorithm 5.1, we assume that the loss is a quadratic function with $(k=2)$. Then, we assume that the loss function is a quintic polynomial. Using the upper bound of sample size and the grid search method various numerical examples are presented in Tables 1–4. In each table we indicate the optimal Bayesian sampling plans by $S_0\equiv(n',m',t',T'_0,T'_1)$, and the corespondent Bayes risk by $R_0\equiv R(n',m',t',T'_0,T'_1)$. Also, we denote the expected number of observation failures by $E[D_0]$, and the expected termination time by $E[\tau_0]$. During computation and in some cases the optimal sampling plan is achieved when T_1 close to T_0. So, to make a sense to the sampling plan (n,m,t,T_0,T_1) we assume that $T_0-T_1\ge 0.05$, $T^*=T^*_0=2$ and $\eta=0.05$. As the true values of parameters and coefficients for the quadratic loss for which we made the calculations, we take $\mu=2$, $\alpha=2$, $\beta=1$, $a_0=a_1=a_2=3$, $C_s=0.5$, $v_s=0.2$, $C_t=2$, $C_r=30$. For the previous standard values, the optimal sampling plan is $(5,1,0.3104,0.7750,0.2000)$, which means, we put 5 items for life testing, and when $t=0.3104$ is less than the time of fifth failure $X_{(5)}$, the life test terminates after the maximum between the first failure and $t=0.3104$, otherwise the life test terminates at $X_{(5)}$. We accept the batch if the estimator of the average lifetime $\hat\theta$ is greater than or equal 0.7750, and we reject it if $\hat\theta$ is less than 0.2000. For $\hat\theta$ is between 0.7750 and 0.2000, the batch is rejected and accepted with probability $p_{\hat\theta}=\Big(0.7750-\hat\theta\Big)/(0.7750-0.2000)$ and $1-p_{\hat\theta}$ respectively, the corresponding Bayes risk is $R_0=23.9637$.

In Table 1, we observe that for α fixed and β decreases while $\mu=2$, $a_0=a_1=a_2=3$, $C_s=0.5$, $v_s=0.2$, $C_t=2$ and $C_r=30$, the Bayes risk R_0 increase. And for β fixed R_0 is increasing in α. On the other hand, we can see that the expected number of failure $E[D_0]$ is close to m' and the expected termination time $E[\tau_0]$ is always greater than t'. Furthermore, for each couple $(\alpha,\beta)=(1.5,0.2)$,$(2.0,0.4)$,$(2.5,0.6)$,$(3.0,0.8)$,$(3.5,0.8)$,$(3.5,1.0)$, the batch is rejected without any sample cost, and thus $R_0=C_r=30$. In Table 2, we can see that, the minimum Bayes risk R_0 significantly increases with the values of a_2, and the optimal sample size n' decreases for a_2 increasing. Furthermore, the optimal number of fixed failures m' is close to n' when a_2 increases. For $a_2\le 2$ and the other parameters and coefficients are fixed, the sampling plan $S_0=(0,0,0,0,0)$ with $R_0=a_0+a_1\alpha/\beta+a_2(\alpha^2+\alpha)/\beta^2$ where the batch is accepted for no sampling case. And, for $a_2\ge 15$ the optimal plan $S_0=(0,0,0,\infty,\infty)$ with $R_0=C_r=30$, the batch is rejected without taking sampling. In Table 3, it is observed that $E[D_0]\ge m'$ and $E[\tau_0]\ge t'$, this indicates that the sampling plan S_0 takes more time to better observe the lifetime components, and can obtain more information about the expected lifetime of items. Also, the number of fixed failures brings closer to the optimal sample size when C_t closes to 0. On the other hand, for C_t increases the optimal sample size increases and the minimum Bayes risk increases. From Table 4, it can be seen that, R_0 is increasing in C_r.

And, for $C_r \leq 17.5$, the batch will be rejected with $R_0 = Cr$. For $Cr \geq 45$, the batch will be accepted with minimum Bayes risk $R_0 = 27$.

Table 1: Optimal sampling plans and Bayes risks for α and β vary.

α	β	n'	m'	t'	T'_0	T'_1	$E[D_0]$	$E[\tau_0]$	R_0
1.5	0.2	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
1.5	0.4	3	2	0.0825	0.6000	0.5500	2.0000	0.6113	28.3413
1.5	0.6	4	1	0.4172	0.1100	0.3000	1.5853	0.5178	25.4861
1.5	0.8	5	1	0.3413	0.8500	0.3250	1.3627	0.4815	22.3640
2.0	0.4	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
2.0	0.6	3	2	0.0874	0.6000	0.5500	2.0000	0.5881	28.8383
2.0	0.8	4	3	0.1009	0.4750	0.4250	3.0000	0.7894	26.5777
2.0	1.0	5	1	0.3104	0.7750	0.2000	1.2956	0.4212	23.9637
2.5	0.6	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
2.5	0.8	3	2	0.0902	0.6000	0.5500	2.0000	0.5764	29.2431
2.5	1.0	4	3	0.1009	0.4750	0.4250	3.0000	0.7492	27.3505
2.5	1.2	4	1	0.3577	0.8250	0.2000	1.3065	0.4528	24.9906
3.0	0.8	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
3.0	1.0	3	2	0.0921	0.6000	0.5500	2.0000	0.5694	29.5670
3.0	1.2	4	3	0.1008	0.4750	0.4250	3.0000	0.7252	27.9490
3.5	0.8	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
3.5	1.0	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
3.5	1.2	3	2	0.0933	0.5750	0.5250	2.0000	0.5648	29.8291

Table 2: Optimal sampling plans and Bayes risks for a_2 varies.

a_2	n'	m'	t	T'_0	T'_1	$E[D_0]$	$E[\tau_0]$	R_0
2.0	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	21.0000
2.5	6	1	0.2884	0.7000	0.3000	1.3312	0.3873	22.8246
3.0	5	1	0.3104	0.7750	0.2000	1.2956	0.4212	23.9637
4.0	5	1	0.3762	1.0750	0.3500	1.5060	0.4607	25.4895
5.0	4	1	0.4421	1.2500	0.3250	1.5159	0.5278	26.0883
6.0	4	1	0.4860	1.5250	0.3750	1.6468	0.5575	27.4065
7.0	4	3	0.1129	0.7250	0.6750	3.0000	0.8827	28.0233
8.0	3	2	0.1129	0.8500	0.8000	2.0000	0.7592	28.5096
10.0	3	2	0.1129	1.0000	0.9500	2.0000	0.7592	29.1995
15.0	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000

Table 3: Optimal sampling plans and Bayes risks for C_t varies.

C_t	n'	m'	t	T'_0	T'_1	$E[D_0]$	$E[\tau_0]$	R_0
1.0	3	3	0.2226	0.4000	0.3500	3.0000	1.1436	23.2092
1.5	4	3	0.1129	0.4000	0.3500	3.0000	0.8827	23.6896
2.0	5	1	0.3104	0.7750	0.2000	1.2956	0.4212	23.9637
2.5	5	1	0.3104	0.7750	0.2000	1.2956	0.4212	24.1743
3.0	6	1	0.2884	0.7500	0.2500	1.3312	0.3873	24.3711
4.0	6	1	0.2884	0.7500	0.2500	1.3312	0.3873	24.7584
5.0	6	1	0.2884	0.7500	0.2500	1.3312	0.3873	25.1458
6.0	6	1	0.2884	0.7500	0.2500	1.3312	0.3873	25.5331
8.0	7	1	0.2665	0.7500	0.2500	1.3437	0.3583	26.3014
10.0	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	27.0000

Table 4: Optimal sampling plans and Bayes risks for C_r varies.

C_r	n'	m'	t	T_0'	T_1'	$E[D_0]$	$E[\tau_0]$	R_0
17.5	0	0	0.0000	∞	∞	0.0000	0.0000	17.5000
20.0	4	1	0.4640	1.4000	0.3250	1.5804	0.5425	19.2507
22.5	4	1	0.4421	1.2500	0.3250	1.5159	0.5278	20.6154
25.0	4	1	0.3982	1.0250	0.2500	1.3943	0.5000	21.8282
27.5	4	1	0.3982	0.9750	0.3000	1.3943	0.5000	22.9330
30.0	5	1	0.3104	0.7750	0.2000	1.2956	0.4212	23.9637
32.5	6	1	0.2884	0.7250	0.2750	1.3312	0.3873	24.7412
35.0	6	1	0.2884	0.7000	0.3000	1.3312	0.3873	25.4510
40.0	6	1	0.2884	0.6750	0.3250	1.3312	0.3873	26.1433
45.0	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	27.0000

6.1. Numerical examples for higher degree polynomal and non polynomial loss

To simulate the Bayes risk performance and obtain the optimal sampling plan under non polynomial loss, a similar algorithm to the one in Section 5 is considered:

a) Start with $(n,m,t)=(0,0,0)$, compute N from (5.1) and compute $R_{NP}(0,0,0,T_0'',T_1'')=\min\Big\{R_{NP}(0,0,0,\infty,\infty)=C_r, R_{NP}(0,0,0,0,0)=\Big(\frac{\beta}{\beta-c}\Big)^{\alpha}-c\frac{\alpha}{\beta}-1\Big\}$.

b) For fixed (n,m,t), compute the optimal $T'_{0,(n,m,t)}$ and $T'_{1,(n,m,t)}$ using grid search method, such that $R\Big(n,m,t,T''_{0,(n,m,t)},T''_{1,(n,m,t)}\Big)=\min\limits_{0<T_1<T_0\le T^*}R(n,m,t,T_1,T_0)$, with grid size 0.0125.

c) For fixed (n,m), compute the optimal $t'_{(n,m)}$ using grid search method, such that $R\Big(n,m,t''_{(n,m)},T''_{0,(n,m,t)},T''_{1,(n,m,t)}\Big)=\min\limits_{t_L\le t\le t_U}R\Big(n,m,t,T''_{0,(n,m,t)},T''_{1,(n,m,t)}\Big)$, with grid size $\frac{t_U-t_L}{100}$.

d) For $0\le m\le n\le N$, choose $(n'',m'',t'',T_0'',T_1'')$ which corresponds to the smallest value of the Bayes risks $R\Big(n,m,t''_{(n,m)},T''_{0,(n,m,t)},T''_{1,(n,m,t)}\Big)$.

Table 5 provides some optimal sampling plans for the polynomial loss with order $k=5$. Under setting: $\mu=2$, $a_1=a_2=a_4=0$, $a_0=a_3=1$, $C_s=0.5$, $v_s=0.2$, $C_t=2$ and $C_r=30$, while α, β and a_5 vary. It appears from this table that the minimum Bayes risk R_0 increases quickly when a_5 increases while α and β fixed are fixed. On the other hand, the values of $E[\tau_0]$ are significant comparing with Table 2, in this case we may observe more than m' failures and this will result in an efficient life testing procedure.

In Table 6, various optimal sampling plans and their minimum Bayes risk are depicted for different values of α, β and c while $\mu=2$, $C_s=0.5$, $v_s=0.2$, $C_t=2$, $Cr=30$. Such that $S_{NP}(n'',m'',t'',T_0'',T_1'')\equiv S_{NP}$ and $R_{NP}(n'',m'',t'',T_0'',T_1'')\equiv R_{NP}$ denote optimal sampling plan and its minimum Bayes risk respectively. As shown in Table 6, the Bayes risk R_{NP} decreases when c is close to 0 for α and β fixed. When c is close to β, R_{NP} and $E[\tau_0]$ are large. There are some optimal sampling plans under no sampling case. For instance see $(\alpha,\beta,c)=(2,1,0.5),(2,1.5,0.7),(4,2,1)$, the optimal sampling plan $S_{NP}=(0,0,0,0,0)$ and the batch is accepted without any sample cost. When $(\alpha, \beta, c) = (5, 3, 2.5)$, $S_{NP} = (0, 0, 0, \infty, \infty)$ and the batch must be rejected without any sample cost.

Table 5: Optimal sampling plans and Bayes risks under polynomial loss with order 5, for α, β and a_5 vary.

α	β	a_5	n'	m'	t	T_0'	T_1'	$E[D_0]$	$E[\tau_0]$	R_0
2	1.0	1	6	5	0.1129	0.7625	0.7125	5.0000	1.0373	26.6566
2	1.0	2	6	5	0.1129	0.8875	0.8375	5.0000	1.0373	28.0317
2	1.0	3	5	4	0.1129	1.1250	1.0750	4.0000	0.9701	29.1941
2	1.5	1	7	5	0.1382	0.6625	0.6125	5.0000	1.1045	21.1053
2	1.5	2	7	5	0.1382	0.7875	0.7375	5.0000	1.1045	23.0980
2	1.5	3	7	5	0.1382	0.8750	0.8250	5.0000	1.1045	24.2304
3	1.5	1	6	5	0.1127	0.7750	0.7250	5.0000	0.9528	28.3606
3	1.5	2	5	4	0.1127	0.9875	0.9375	4.0000	0.8911	29.7964
3	1.5	3	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000
3	2.0	1	7	6	0.1302	0.6625	0.6125	6.0000	1.1577	23.7820
3	2.0	2	7	6	0.1302	0.7875	0.7375	6.0000	1.1577	26.0688
3	2.0	2	6	5	0.1302	0.9125	0.8625	5.0000	1.1002	27.3013

Table 6: Optimal sampling plans and Bayes risks under non polynomial loss for α, β and c vary.

α	β	c	n''	m''	t''	T_0''	T_1''	$E[D_0]$	$E[\tau_0]$	R_{NP}
2	1.0	0.5	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	02.0000
2	1.0	0.8	5	4	0.1129	0.2250	0.1750	4.0000	0.9701	12.5445
3	1.5	0.7	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	04.1918
3	1.5	1.0	6	5	0.1127	0.2875	0.2375	5.0000	0.9528	15.8429
3	1.5	1.3	7	6	0.1127	0.4500	0.4000	6.0000	1.0026	21.8487
4	2.0	1.0	0	0	0.0000	0.0000	0.0000	0.0000	0.0000	13.0000
4	2.0	1.2	6	5	0.1127	0.3625	0.3125	5.0000	0.9168	19.9275
4	2.0	1.5	7	6	0.1127	0.5375	0.4875	6.0000	0.9647	25.4796
4	2.0	1.8	6	5	0.1127	0.7375	0.6875	5.0000	0.9168	28.9217
5	3.0	1.5	6	5	0.1234	0.4375	0.3875	5.0000	0.9825	20.9812
5	3.0	2.0	6	5	0.1234	0.7750	0.7250	5.0000	0.9825	28.5097
5	3.0	2.5	0	0	0.0000	∞	∞	0.0000	0.0000	30.0000

7. CONCLUSION

In [14], Bayesian sampling plans for exponential distribution based on type II hybrid censored samples under the quadratic loss have been discussed, since the time-consuming cost and the salvage value are not included in the loss function. However, Several single variables sampling plans have been improved in recent years, most improvements have been achieved by considering the one-sided decision function. Such that, these studies do not take into account that a doubt zone can be existed in the decision interval, e.g. this can be happened when the experimenter estimates that the minimum acceptable and the maximum rejectable surviving time are not equal. Nevertheless, there are still some interesting and relevant problems to be addressed in this situation. With this purpose, we have determined Bayesian sampling plans for Weibull distribution under type II hybrid censoring based on a two-sided decision function with a random doubt zone. We provided an explicit expression for the Bayes risk using a suitable polynomial loss, which includes the unit inspection cost, the time consuming-cost, the rejection cost, the salvage value, and the after-sales cost. Furthermore, we have expressed an explicit form for the Bayes risk under non polynomial loss with the LINEX form. It is noticed that, the Bayes risk under the polynomial loss (resp. non polynomial loss) is always quite complicated. So, we proposed an upper bound for the optimal size of the sample and a finite algorithm to simulate the risk function numerically based on the grid search method. Based on the results, it can be concluded that the Bayes risk based on the two-side decision function have robust behavior with considering the changes of the parameters and coefficients in the proposed sampling plan. However, in this paper we have considered Weibull distribution with known shape parameter. Further study of the issue is still required for completely Bayesian analysis to the two-parameter Weibull distribution. More research will be needed along with this issue for other censoring.

A. APPENDIX

A.1. Computation of $E\{E(D_{n,m}|\lambda)\}$

Let $F(x|\lambda,\mu)$ be the cdf of X. The probability function of $D_{n,m}$, such that $D_{n,m} = m, m+1, ..., n$ can be calculated as follows:
For $j = m+1, ..., n$

$$\begin{aligned} P(D_{n,m} = j) &= P(X_1 \le t, X_2 \le t, ..., X_j \le t, X_{j+1} > t, X_{j+2} > t, ..., X_n > t) \\ &= \binom{n}{j} F(t|\lambda,\mu)^j (1 - F(t|\lambda,\mu))^{n-j} = \binom{n}{j}\left(1 - e^{-\lambda t^\mu}\right)^j e^{-\lambda(n-j)t^\mu}, \\ P(D_{n,m} = m) &= 1 - P(D_{n,m} > m) = 1 - \sum_{d=m+1}^{n} \tbinom{n}{d}\left(1 - e^{-\lambda t^\mu}\right)^d e^{-\lambda(n-d)t^\mu} \\ &= \sum_{d=0}^{m} \tbinom{n}{d}\left(1 - e^{-\lambda t^\mu}\right)^d e^{-\lambda(n-d)t^\mu}, \end{aligned}$$

Then for $m \le n$

$$\begin{aligned} &E(D_{n,m}|\lambda) = \sum_{d=m}^{n} dP(D_{n,m} = d) \\ &= \sum_{d=m+1}^{n} d\tbinom{n}{d}\left(1 - e^{-\lambda t^\mu}\right)^d e^{-\lambda(n-d)t^\mu} + m\sum_{d=0}^{m} \tbinom{n}{d}\left(1 - e^{-\lambda t^\mu}\right)^d e^{-\lambda(n-d)t^\mu} \\ &= \sum_{d=m+1}^{n}\sum_{j=0}^{d} (-1)^{d-j} d\tbinom{n}{d}\tbinom{d}{j} e^{-\lambda(n-j)t^\mu} + m\sum_{d=0}^{m}\sum_{j=0}^{d} (-1)^{d-j}\tbinom{n}{d}\tbinom{d}{j} e^{-\lambda(n-j)t^\mu} \\ &= \sum_{d=0}^{n}\sum_{j=0}^{d} (-1)^{d-j} M\tbinom{n}{d}\tbinom{d}{j} e^{-\lambda(n-j)t^\mu}, \end{aligned}$$

it is easy to show that when $m = n$, $E(D_{n,m}|\lambda) = nP(D_{n,m} = n) = n$. Hence

$$\begin{aligned} E\{E(D_{n,m}|\lambda)\} &= \int_0^\infty E(D_{n,m}|\lambda) g(\lambda;\alpha,\beta) d\lambda \\ &= \sum_{d=0}^{n}\sum_{j=0}^{d} (-1)^{d-j} M\tbinom{n}{d}\tbinom{d}{j} \tfrac{\beta^\alpha}{\Gamma(\alpha)} \int_0^\infty e^{-\lambda(\beta+(n-j)t^\mu)} \lambda^{\alpha-1} d\lambda \\ &= \sum_{d=0}^{n}\sum_{j=0}^{d} (-1)^{d-j} M\tbinom{n}{d}\tbinom{d}{j} \left(\tfrac{\beta}{\beta+(n-j)t^\mu}\right)^\alpha. \end{aligned}$$

A.2. Computation of $E\{E(\tau_{n,m}|\lambda)\}$

The computation of $E\{E(\tau_{n,m}|\lambda)\}$ is similar as in [20]. Let I_A be the indicator function of a set A.

For $m < n$, when $X_{(m)} \geq t$, $\tau_{n,m} = X_{(m)}$, then

$$E\big(X_m I_{\{X_m \geq t\}}|\lambda\big) = \int\limits_t^\infty y f_{X_{(m)}}(y)dy$$
$$= m\binom{n}{m}\sum_{j=0}^{m-1}(-1)^{m-j-1}\tbinom{m-1}{j}\int\limits_t^\infty e^{-\lambda(n-j)y^\mu}\lambda y^\mu dy$$

Therefore

$$E\big\{E\big(X_m I_{\{X_m \geq t\}}|\lambda\big)\big\}$$
$$= m\binom{n}{m}\sum_{j=0}^{m-1}(-1)^{m-j-1}\tbinom{m-1}{j}\int\limits_0^\infty\int\limits_t^\infty \tfrac{\beta^\alpha}{\Gamma(\alpha)}e^{-\lambda(\beta+(n-j)y^\mu)}\lambda^\alpha y^\mu dy d\lambda$$
$$= m\binom{n}{m}\sum_{j=0}^{m-1}(-1)^{m-j-1}\tbinom{m-1}{j}\int\limits_t^\infty \alpha\beta^\alpha \tfrac{\mu y^\mu}{(\beta+(n-j)y^\mu)^{\alpha+1}}dy.$$

A simple transformation $z = (n-j)y^\mu/(\beta+(n-j)y^\mu)$ yields

$$E\big\{E\big[X_m I_{\{X_m \geq t\}}|\lambda\big]\big\}$$
$$= m\binom{n}{m}\sum_{j=0}^{m-1}(-1)^{m-j-1}\tbinom{m-1}{j}\tfrac{\alpha\beta^{1/\mu}}{(n-j)^{1+1/\mu}}B_{1-q*}\Big(1+\tfrac{1}{\mu}, \alpha-\tfrac{1}{\mu}\Big).$$

For $X_m < t < X_n$, $\tau_{n,m} = t$, then

$$E\big[tI_{\{X_m<t<X_n\}}|\lambda\big]$$
$$= \int\limits_0^t\int\limits_t^\infty \tfrac{tn!(\lambda\mu)^2(xy)^{\mu-1}e^{-\lambda(x^\mu+y^\mu)}}{(m-1)!(n-m-1)!}\big(1-e^{-\lambda x^\mu}\big)^{m-1}\big(e^{-\lambda x^\mu}-e^{-\lambda y^\mu}\big)^{n-m-1}dydx$$
$$= \tfrac{tn!}{(m-1)!(n-m-1)!}\int\limits_0^t\int\limits_t^\infty\sum_{i=0}^{m-1}(-1)^{m-i-1}\tbinom{m-1}{i}\lambda\mu x^{\mu-1}e^{-\lambda(m-i)x^\mu}$$
$$\times\sum_{j=0}^{n-m-1}(-1)^{n-m-j-1}\tbinom{n-m-1}{j}\lambda\mu y^{\mu-1}e^{-\lambda(n-m-j)y^\mu}e^{-\lambda j x^\mu}$$
$$= \tfrac{tn!}{(m-1)!(n-m-1)!}\sum_{i=0}^{m-1}\sum_{j=0}^{n-m-1}\Bigg[(-1)^{n-i-j}\tbinom{m-1}{i}\tbinom{n-m-1}{j}\int\limits_0^t\lambda\mu x^{\mu-1}e^{-\lambda(m+j-i)x^\mu}dx$$
$$\times\int\limits_t^\infty\lambda\mu y^{\mu-1}e^{-\lambda(n-m-j)y^\mu}dy\Bigg]$$
$$= \tfrac{tn!}{(m-1)!(n-m-1)!}\sum_{i=0}^{m-1}\sum_{j=0}^{n-m-1}(-1)^{n-i-j}\tbinom{m-1}{i}\tbinom{n-m-1}{j}\tfrac{e^{-\lambda(n-m-j)t^\mu}-e^{-\lambda(n-i)t^\mu}}{(m+j-i)(n-m-j)}.$$

Thus

$$E\big\{E\big[tI_{\{X_m<t<X_n\}}|\lambda\big]\big\} = \int\limits_0^\infty E\big[tI_{\{X_m<t<X_n\}}|\lambda\big]g(\lambda;\alpha,\beta)d\lambda$$
$$= \tfrac{tn!}{(m-1)!(n-m-1)!}\sum_{i=0}^{m-1}\sum_{j=0}^{n-m-1}\Big[(-1)^{n-i-j}\tbinom{m-1}{i}\tbinom{n-m-1}{j}$$
$$\times\tfrac{\beta^\alpha}{(m+j-i)(n-m-j)}\Big(\tfrac{1}{((n-m-j)t^\mu+\beta)^\alpha}-\tfrac{1}{((n-i)t^\mu+\beta)^\alpha}\Big)\Big].$$

For $X_{(n)} \leq t$, $\tau_{n,m} = X_{(n)}$, then

$$E\big\{E\big[X_{(n)}I_{\{X_n\leq t\}}|\lambda\big]\big\} = \int\limits_0^\infty\int\limits_0^t y f_{X_{(n)}}(y)g(\lambda;\alpha,\beta)dyd\lambda$$
$$= n\sum_{j=0}^{n-1}(-1)^{n-j-1}\tbinom{n-1}{j}\tfrac{\alpha\beta^{1/\mu}}{(n-j)^{1+1/\mu}}B_{q*}\Big(1+\tfrac{1}{\mu}, \alpha-\tfrac{1}{\mu}\Big).$$

Hence, for $m < n$

$$
\begin{aligned}
E\{E[\tau_{n,m}|\lambda]\} &= m\binom{n}{m}\sum_{j=0}^{m-1}(-1)^{m-j-1}\tbinom{m-1}{j}\tfrac{\alpha\beta^{1/\mu}}{(n-j)^{1+1/\mu}}B_{1-q*}\Big(1+\tfrac{1}{\mu},\alpha-\tfrac{1}{\mu}\Big)\\
&\quad+\tfrac{tn!}{(m-1)!(n-m-1)!}\sum_{i=0}^{m-1}\sum_{j=0}^{n-m-1}\Big[(-1)^{n-i-j}\tbinom{m-1}{i}\tbinom{n-m-1}{j}\\
&\quad\times\tfrac{\beta^{\alpha}}{(m+j-i)(n-m-j)}\Big(\tfrac{1}{((n-m-j)t^{\mu}+\beta)^{\alpha}}-\tfrac{1}{((n-i)t^{\mu}+\beta)^{\alpha}}\Big)\Big]\\
&\quad+n\sum_{j=0}^{n-1}(-1)^{n-j-1}\tbinom{n-1}{j}\tfrac{\alpha\beta^{\frac{1}{\mu}}}{(n-j)^{1+1/\mu}}B_{q*}\Big(1+\tfrac{1}{\mu},\alpha-\tfrac{1}{\mu}\Big)
\end{aligned}
$$

For $m = n$, $\tau_{n,m} = X_{(n)}$

$$
\begin{aligned}
E\{E[\tau_{n,m}|\lambda]\} &= \int_0^\infty\int_0^\infty y f_{X_{(n)}}(y)g(\lambda;\alpha,\beta)dyd\lambda\\
&= n\alpha\beta^{1/\mu}B\Big(1+\tfrac{1}{\mu},\alpha-\tfrac{1}{\mu}\Big)\sum_{j=0}^{n-1}(-1)^{j}\tbinom{n-1}{j}\tfrac{1}{(j+1)^{1+1/\mu}}.
\end{aligned}
$$

ACKNOWLEDGMENTS

The authors would like to thank the two referees and Associate Editor for their valuable comments and suggestions which lead to the improvement of the presentation of the work. This work was partially supported by DGRSDT Grant C0656701 from Algeria.

REFERENCES

[1] Asar, Y. and Belaghi, R.A. (2021). Estimation in Weibull distribution under progressively type-I hybrid censored data, *REVSTAT – Statistical Journal* (Forthcoming papers).

[2] Aslam, M. (2019). A variable acceptance sampling plan under neutrosophic statistical interval method, *Symmetry*, **11**(1), 114.

[3] Aslam, M. (2019). A new attribute sampling plan using neutrosophic statistical interval method, *Complex & Intelligent Systems*, **5**(4), 365–370.

[4] Belbachir, H. and Benahmed, M. (2022). Two-sided sampling plan for exponential distribution under type II censored samples, *Hacettep Journal of Mathematics & Statistics*, **51**(1), 327–337.

[5] Chen, L.S.; Yang, M.C. and Liang, T. (2017). Curtailed Bayesian sampling plans for exponential distributions based on Type-II censored samples, *Journal of Statistical Computation and Simulation*, **87**(6), 1160–1178.

[6] Chen, L.S.; Yang, M.C. and Liang, T. (2015). Bayesian sampling plans for exponential distributions with interval censored samples, *Naval Research Logistics*, **62**(7), 604–616.

[7] Childs, A.; Chandrasekar, B.; Balakrishnan, N. and Kundu, D. (2003). Exact likelihood inference based on Type-I and Type-II hybrid censored samples from the exponential distribution, *Annals of the Institute of Statistical Mathematics*, **55**(2), 319–330.

[8] Epstein, B. (1954). Truncated life tests in the exponential case, *Annals of Mathematical-Statistics*, **25**(3), 555–564.

[9] Guenther, W.C. (1971). On the determination of single sampling attribute plans based upon a linear cost model and a prior distribution, *Technometrics*, **13**(3), 483–498.

[10] Hald, A. (1967). Asymptotic properties of Bayesian single sampling plans, *Journal of the Royal Statistical Society Series B*, **29**(1), 162–173.

[11] Hald, A. (1981). *Statistical Theory of Sampling Inspection by Attributes*, Academic Press, New York.

[12] Lam, Y. (1990). An optimal single variable sampling plan with censoring, *The Statistician*, **39**(1), 53–67.

[13] Lam, Y. (1994). Bayesian variable sampling plans for the exponential distribution with type I censorin, *Annals of Statistics*, **22**(2), 696–711.

[14] Lin, C.T.; Huang, Y.L. and Balakrishnan, N. (2008). Exact Bayesian variable sampling plans for the exponential distribution based on type-I and type-II censored samples, *Communications in Statistics – Simulation*, **37**(6), 1101–1116.

[15] Liang, T. and Yang, M.C. (2013). Optimal Bayesian sampling plans for exponential distributions based on hybrid censored samples, *Journal of Statistical Computation and Simulation*, **83**(5), 920–938.

[16] Nagy, M. and Alrasheedi, A.F. (2022). The lifetime analysis of the Weibull model based on Generalized Type-I progressive hybrid censoring schemes, *Mathematical Biosciences and Engineering*, **19**(3), 2330–2354.

[17] Prajapati, D.; Mitra, S. and Kundu, D. (2020). A new decision theoretic sampling plan for exponential distribution under Type-I censoring, *Communications in Statistics – Simulation and Computation*, **49**(2), 453–471.

[18] Tsai, T.R.; Chiang, J.Y.; Liang, T. and Yang, M.C. (2014). Efficient Bayesian sampling plans for exponential distributions with Type-I censored samples, *Journal of Statistical Computation and Simulation*, **84**(5), 964–981.

[19] Wetherill, G.B. (1977). *Sampling Inspection and Quality Control*, second ed., Chapman and Hall, London.

[20] Yang, M.C.; Chen, L.S. and Liang, T. (2017). Optimal Bayesian variable sampling plans for exponential distributions based on modified type-II hybrid censored samples, *Communications in Statistics – Simulation and Computation*, **46**(6), 4722–4744.

REVSTAT – Statistical Journal
Volume 22, Number 3, July 2024, 387–410
https://doi.org/10.57805/revstat.v22i3.528

Stochastic Comparison for Extreme Order Statistics Arising from PHR, PRHR or Location Model

Authors: Madhurima Datta ✉
– Department of Mathematics, Indian Institute of Technology Kharagpur, India-721302
– Centre for Data Science, Institute of Technical Education and Research, Bhubaneswar, Odisha, India-751030
madhurima92.datta@gmail.com

Nitin Gupta
– Department of Mathematics, Indian Institute of Technology Kharagpur, India-721302
nitin.gupta@maths.iitkgp.ac.in

Received: March 2020 Revised: September 2022 Accepted: September 2022

Abstract:

- We have considered series and parallel systems with independent set of random variables belonging from PHR, PRHR family of distributions to study dispersive and star ordering between two systems such that the number of components in both the system are different. This will help us in identifying the deviation of lifetime of a product from the warranty of the product. Moreover, we have also considered series and parallel systems with dependent set of random variables each belonging from location based models, such that the baseline distribution for both the sets are different. The Archimedean copula generators used here are ψ_1 and ψ_2 such that the condition "$\phi_1 \cdot \psi_2$ or $\phi_2 \cdot \psi_1$ is super-additive" holds. Earlier researchers have studied location-scaled or resilience-scaled models for independent or dependent set of random variables. Our study is an addition to the existing research.

Keywords:

- *Archimedean copula; dispersive order; proportional hazard rate distribution; proportional reversed hazard rate distribution; star order.*

AMS Subject Classification:

- 62N05.

✉ Corresponding author

1. INTRODUCTION

The series and parallel systems are the most frequent and maximum encountered systems in nature. These systems are statistically referred to as the minimum and the maximum order statistic respectively. Let $X_1, X_2, ..., X_n$ be n independent and non-identical random variables from a particular population. Then arranging the random variables according to their magnitude or strength we observe that $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$, where $X_{k:n}$ is known as the k-th order statistic. $X_{k:n}$ represents the lifetime of a $(n-k+1)$-out-of-n system. In this paper, we focus only on the minimum and maximum order statistic. A great deal of literature is available on the stochastic relationship among the order statistics for various probability distributions.

In particular, our problem deals with the Proportional hazard rate (PHR) model and the proportional reversed hazard rate (PRHR) model. We consider $X_1, X_2, ..., X_n$ as independent random variables, where the survival function of each random variable X_i follows the PHR model for $i = 1, ..., n$. Then reliability or survival probability of X_i is:

$$P(X_i > x) = \overline{F}_i(x) = [\overline{F}_0(x)]^{\lambda_i}, \lambda_i > 0, i = 1, 2, ..., n,$$

where λ_i is the proportionality parameter. Here let X_0 be the baseline random variable with the baseline distribution $F_0(x)$ and the baseline survival function $\overline{F}_0(x) = 1 - F_0(x)$. Exponential, Weibull, Pareto, Lomax, Kumaraswamy's distributions are some examples of PHR model distribution. [24] pioneered the study of stochastic ordering (details about stochastic ordering are given in the next section) for k-out-of-n systems which included usual stochastic ordering results for PHR model. [25] studied dispersive and star ordering for general distributions in detail. Later on, many researchers have continued the study and found many results for PHR model. [5] discussed that the existing results for exponential distribution which can be extended for PHR models, this was possible as the random variable corresponding to the cumulative hazard rate function of a PHR family of distribution follows exponential distribution with the proportionality constant as the parameter i.e., if X follows $[\overline{F}(x)]^{\lambda}$, then the cumulative hazard rate function follows Exp(λ) distribution. [15] demonstrated dispersive ordering between the maximum order statistics of two PHR populations. [23] and [28] observed dispersive ordering between the 2^{nd} order statistics (also known as fail-safe systems) from two different populations and derived bounds on the corresponding parameters. Considering the parallel systems having PHR distributed components, [16] studied the dispersive ordering between them. A comprehensive review of the various stochastic ordering between the order statistics for random variables belonging from the PHR model has been done by [5]. Recently [11] observed stochastic ordering for series and parallel systems with Kumaraswamy's and Frechet distributed components. Now we shall observe what is meant by a multiple outlier model.

Let $X_1, X_2, ..., X_n$ and $Y_1, Y_2, ..., Y_n$ be n-independent PHR samples having the same baseline distribution but the parameter vectors are given by $(\underbrace{\alpha_1, ..., \alpha_1}_{p}, \underbrace{\alpha_2, ..., \alpha_2}_{q})$ and $(\underbrace{\beta_1, ..., \beta_1}_{p}, \underbrace{\beta_2, ..., \beta_2}_{q})$ respectively, where $p + q = n$. Such an arrangement is described as the multiple-outlier model. [4] and [29] discussed the hazard rate and the likelihood ratio ordering for parallel systems with multiple-outlier PHR model. For a similar model, [2] derived conditions on the distribution function for the dispersive ordering of k-th order statistic where the parameter vectors follow majorization relation. [9] found the necessary and sufficient conditions for the hazard rate ordering among the second order statistics. [30] examined the stochastic comparison between series and parallel systems where the component lifetimes are dependent, heterogeneous and resilience scaled. [13] and [14] found several conditions for stochastic ordering of maximum and minimum order statistics from a location-scale family of distributions. [3] observed stochastic ordering between the sample ranges where component lifetimes (number of components are different) are independent and follows multiple-outlier exponential distribution and PHR models.

In contrast to the PHR model, proportional reversed hazard rate (PRHR) model was developed. Let X_i follows PRHR model then the distribution function of X_i is given by

$$P(X_i < x) = F_i(x) = [F_0(x)]^{\theta_i}, \theta_i > 0, i = 1, 2, ..., n,$$

where θ_i is the proportionality constant. Some known examples of PRHR model are exponentiated Weibull, exponentiated exponential, exponentiated Gamma, etc. [1] observed dispersive ordering for the series systems with components following the PRHR model.

Here we consider two sets of independent PHR and PRHR models where the baseline distribution of both the sets are different and the sample sizes are also different i.e., the first set of random variables $X_i \sim \overline{F}_i(x) = (\overline{F}_0(x))^{\alpha_i}$ for $i = 1, 2, ..., n_1$ and the second set $Y_i \sim \overline{G}_i(x) = (\overline{G}_0(x))^{\beta_i}$ for $i = 1, 2, ..., n_2$. Considering the same baseline distribution $F_0 = G_0$, we study dispersive and star ordering for series/parallel models. A similar kind of study being conducted for series/parallel system made up of PRHR distributed components.

We have also considered a general model as, $X_1, ..., X_{p_1}$ that has survival function $[\overline{F}(x)]^{\alpha_i}$ and $X_{p_1+1}, ..., X_{n_1}$ has survival function $[\overline{G}(x)]^{\alpha_i}$. And $Y_1, ..., Y_{p_2}$ has survival function as $[\overline{F}(x)]^{\beta_i}$ whereas the components $Y_{p_2+1}, ..., Y_{n_2}$ has survival function $[\overline{G}(x)]^{\beta_i}$. We have proved that the hazard rate ordering for sample minimum exists for such models, analogously reversed hazard rate ordering for sample maximums exist for PRHR model. A reversed hazard rate ordering for sample maximum (with equal sample sizes) for Pareto distributed random variables has been observed when only the shape parameter varies.

Lastly, we study some results for series system having dependent components, where the dependence among components has been considered as having Archimedean type of copula. These studies include the results when the location parameter is varied along with a comparison between two generating functions (super-additive property) and usual stochastic ordering among baseline distributions.
The paper has been constructed as follows: Section 2 includes all the definitions used in the paper, Section 3 contains results and discussion where Subsection 3.1 contains dispersive ordering results for PHR and PRHR model with unequal sample sizes, Subsection 3.2 contains star ordering result for unequal sample sizes and Subsection 3.2 contains result for the dependent model. The various well-known lemmas that have been used in proving the results are discussed under Section 2.

2. DEFINITIONS

Let X and Y be two absolutely continuous random variables with distribution functions $F(x)$ and $G(x)$; reliability functions as $\overline{F}(x)$ and $\overline{G}(x)$; probability density functions as $f(x)$ and $g(x)$; hazard rate functions as $r(x) = \dfrac{f(x)}{\overline{F}(x)}$ and $s(x) = \dfrac{g(x)}{\overline{G}(x)}$; reversed hazard rate functions as $\tilde{r}(x) = \dfrac{f(x)}{F(x)}$ and $\tilde{s}(x) = \dfrac{g(x)}{G(x)}$, respectively. Let F^{-1} and G^{-1} be the right continuous quantiles of X and Y respectively. A real valued function ψ is super-additive when $\psi(x_1+x_2) \geq \psi(x_1)+\psi(x_2)$ for all $x_1, x_2 \in Domain(\psi)$. This concept is valid even when the summation is over n-variables. For details on the above definitions we refer the reader to [6]. One can note that a random variable has decreasing reversed hazard rate(DRHR) if and only if the distribution function is log-concave. It is known that there exists no distribution which is log convex or increasing reversed hazard rate(IRHR) over the entire domain $[0,\infty)$. An IRHR distribution can be constructed if the domain is taken as $(-\infty,\alpha)$ for some finite α (see [8]), an example of that is Truncated Normal distribution with domain as $(-\infty,0]$. Next we discuss some of the various stochastic orders available in literature. We refer the reader to [26] for the detail of these orderings.

Definition 2.1. X is smaller than Y in:

a) Usual stochastic order $(X \leq_{st} Y)$ if and only if

$$\overline{F}(x) \leq \overline{G}(x),\ \forall x \in (-\infty,\infty).$$

b) Hazard rate order $(X \leq_{hr} Y)$ if $r(x) \geq s(x), x \in \mathbb{R}$. Equivalently, if $\dfrac{\overline{G}(x)}{\overline{F}(x)}$ is increasing in x over the union of the supports of X and Y.

c) Reversed hazard rate order $(X \leq_{rh} Y)$ if $\tilde{r}(x) \leq \tilde{s}(x), x \in \mathbb{R}$. Equivalently, if $\dfrac{G(x)}{F(x)}$ is increasing in x over the union of the supports of X and Y.

d) Likelihood ratio order $(X \leq_{lr} Y)$ if $\dfrac{g(x)}{f(x)}$ is increasing in x over the union of the supports of X and Y.

e) Dispersive order $(X \leq_{\rm disp} Y)$ if

$$F^{-1}(\alpha_2) - F^{-1}(\alpha_1) \leq G^{-1}(\alpha_2) - G^{-1}(\alpha_1) \text{ whenever } 0 < \alpha_1 \leq \alpha_2 < 1.$$

Equivalently, $(X \leq_{\rm disp} Y)$ if and only if

$$G^{-1}(\alpha) - F^{-1}(\alpha) \text{ increases in } \alpha \in (0,1).$$

f) Star order $(X \leq_{*} Y)$ if $\dfrac{G^{-1}(t)}{F^{-1}(t)}$ increases in $t \in (0,1)$.

Here,

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{st} Y.$$

Similarly,

$$X \leq_{lr} Y \Rightarrow X \leq_{rh} Y \Rightarrow X \leq_{st} Y.$$

The detailed description about the inter-relationship between each of the stochastic orders can be seen from the book [26].

Definition 2.2. Majorization: Let $\underline{a} = (a_1, ..., a_n)$ and $\underline{b} = (b_1, ..., b_n)$ be two real valued vectors. Then:

- $\underline{a}$ is majorized by $\underline{b}$ ($\underline{a} \prec \underline{b}$) if $\sum_{i=1}^{n} a_{i:n} = \sum_{i=1}^{n} b_{i:n}$ and $\sum_{i=1}^{k} a_{i:n} \geq \sum_{i=1}^{k} b_{i:n}\ \forall\ k = 1, ..., n-1$;
- $\underline{a}$ is weakly submajorized by $\underline{b}$ ($\underline{a} \prec_w \underline{b}$) if $\sum_{i=1}^{k} a_{n-i+1:n} \leq \sum_{i=1}^{k} b_{n-i+1:n}\ \forall\ k = 1, ..., n$;
  $\underline{a}$ is weakly supermajorized by $\underline{b}$ ($\underline{a} \prec^w \underline{b}$) if $\sum_{i=1}^{k} a_{i:n} \geq \sum_{i=1}^{k} b_{i:n}\ \forall\ k = 1, ..., n$;
  where $a_{1:n} \leq ... \leq a_{n:n}$ $(b_{1:n} \leq ... \leq b_{n:n})$ is the increasing arrangement of $a_1, ..., a_n$ $(b_1, ..., b_n)$.

For $\underline{a}$ and $\underline{b}$, we have $\underline{a} \prec^w \underline{b} \Leftarrow \underline{a} \prec \underline{b} \Rightarrow \underline{a} \prec_w \underline{b}$.

Definition 2.3. Schur-convexity (Schur-concavity): A real valued function ψ defined on a subset of $\mathbb{R}^n$ is *Schur-convex (Schur-concave)* if

$$\underline{a} \prec \underline{b} \Rightarrow \psi(\underline{a})\ \leq (\geq)\ \psi(\underline{b}), \tag{2.1}$$

where $\underline{a} = (a_1, ..., a_n)$ and $\underline{b} = (b_1, ..., b_n)$ are two real valued vectors.

Throughout the paper, the notation $a \stackrel{\text{sgn}}{=} b$ has been used to represent sign of a is same as b. The results and lemmas that are used in obtaining the proofs are mentioned in the following subsection.

2.1. Useful results

Lemma 2.1 (Theorem 3.A.4, see [19]). *Let*

$$\Delta = (a_i - a_j)\left(\frac{\partial \psi(\underline{a})}{\partial a_i} - \frac{\partial \psi(\underline{a})}{\partial a_j}\right),$$

for an open interval $\mathbb{A} \subset \mathbb{R}$, a continuously differentiable function $\psi : \mathbb{A}^n \rightarrow \mathbb{R}$ is Schur-convex (Schur-concave) if and only if it is symmetric on $\mathbb{A}^n$ and for all $i \neq j$, $\Delta \geq (\leq) 0$.

Lemma 2.2 (Proposition 3.C.1, see [19]). *If $\mathbb{A} \subset \mathbb{R}$ is an interval and $h : \mathbb{A} \rightarrow \mathbb{R}$ is convex (concave), then $\psi(\underline{a}) = \sum_{i=1}^{n} h(a_i)$ is Schur-convex (Schur-concave) on $\mathbb{A}^n$, where $\underline{a} = (a_1, ..., a_n)$.*

Lemma 2.3 (Theorem 3.A.8, see [19]). *Let $S \subset \mathbb{R}^n$, a function $f : S \rightarrow \mathbb{R}$ satisfying*

$$\underline{a} \prec_w \underline{b}\ (\underline{a} \prec^w \underline{b})\ \text{on}\ S\ \Rightarrow f(\underline{a}) \leq f(\underline{b})$$

if and only if f is increasing (decreasing) and Schur-convex on S.

Lemma 2.4 (see [25]). *Let F_α, $\alpha \in \mathbb{R}$ be a class of distribution functions such that the support of F_α is given by some interval $(x_0, x_1) \subset \mathbb{R}^+$ and has a non-vanishing density $f_\alpha(x)$ on any subinterval of (x_0, x_1), where x_0 and x_1 are the left and right end points respectively. Then*

$$F_\alpha \leq_{\text{disp}} F_{\alpha^*},\ \ \alpha, \alpha^* \in \mathbb{R},\ \ \alpha \leq \alpha^*, \tag{2.2}$$

if and only if $\dfrac{F_\alpha'(x)}{f_\alpha(x)}$ is decreasing in x, where F_α' is the derivative of F_α with respect to α. And

$$F_\alpha \leq_* F_{\alpha^*},\ \ \alpha, \alpha^* \in \mathbb{R},\ \ \alpha \leq \alpha^*, \tag{2.3}$$

if and only if $\dfrac{F_\alpha'(x)}{x f_\alpha(x)}$ is decreasing in x, where F_α' is the derivative of F_α with respect to α. The first inequalities in (2.2) and (2.3) reverses as the quantity $\dfrac{F_\alpha'(x)}{f_\alpha(x)}$ and $\dfrac{F_\alpha'(x)}{x f_\alpha(x)}$ respectively increases in x.

3. RESULTS AND DISCUSSION

3.1. Dispersive ordering results for unequal sample sizes

In this section we compare minimum and maximum order statistics arising from taking random variables having general proportional hazard rate and proportional reversed hazard rate distribution. As a corollary some results for multiple-outlier models has also been obtained. The multiple-outlier model has been explained in [4, 29] as an independent set of random variables $X_1, X_2, ..., X_n$, where $F_{X_i} = F_X$ for $i = 1, ..., p$ and $F_{X_i} = F_Y$ for $i = p+1, ..., n$, necessarily $1 \leq p < n$. When the value of $p = n-1$, this becomes a single-outlier model. Earlier many researchers have studied various results for the comparison of order statistics from multiple-outlier models. [2] considered the following model

$$(X_1, X_2, ..., X_n) \sim (\underbrace{(\overline{F}(x))^{\alpha_1}, ..., (\overline{F}(x))^{\alpha_1}}_{p}, \underbrace{(\overline{F}(x))^{\alpha_2}, ..., (\overline{F}(x))^{\alpha_2}}_{q})$$

and

$$(Y_1, Y_2, ..., Y_n) \sim (\underbrace{(\overline{F}(x))^{\alpha_1^*}, ..., (\overline{F}(x))^{\alpha_1^*}}_{p}, \underbrace{(\overline{F}(x))^{\alpha_2^*}, ..., (\overline{F}(x))^{\alpha_2^*}}_{q}).$$

They observed star and dispersive ordering for the k-th order statistic by imposing majorization properties over the parameters. In the following paper they primarily discussed hazard rate ordering for exponentially distributed components and derived similar hazard rate ordering results for maximum order statistic with some additional conditions over the parameters. Under the same conditions [9] observed hazard rate ordering for second order statistic. Moreover they found hazard rate orderings when the number of components and number of outliers were different. Whereas [21] studied maximum order statistic for PHR model (survival function of X_i is $\overline{F}_{X_i}(x) = (\overline{F}(x))^{\alpha_i}$ for $i = 1, ..., n$) such that the distribution function of $\max_{i \in P} X_i$, $P \subset \{1, 2, ..., n\}$ is

$$F_{\max}(x) = Q_P(F(x)),$$

where Q_P is a distortion function (continuous and increasing in [0,1], also Q(0)=0, Q(1)=1) and it depends on the underlying copula and the proportionality parameters. Few results were observed for different subsets of $\{1, 2, ..., n\}$. Further they have also discussed some results corresponding to multiple outlier model, PHR distributions using the aforementioned distortion function and results for the independent cases. We have considered various models in our study which includes model where the baseline distributions are same but the shape parameter varies, the baseline distributions are different and the shape parameters are also different. Several researchers have studied multiple-outlier models extensively as it helps in dealing with outliers. Recently, [31] studied some results where the n-component lifetimes of both the systems are dependent with multiple-outlier proportional hazard rates. [10] studied stochastic ordering for two types of models: Modified proportional hazard rate scale model and Modified proportional reversed hazard rate scale model. In our present study we first observe results for series systems where the component lifetimes are independent and follows different proportional hazard rates (the number of components in both the systems are not necessarily same) and the results for multiple-outlier models can be derived subsequently.

The following theorem has been observed for series systems with components following PHR family of distributions such that the baseline distribution for both the sets are different.

Theorem 3.1. *Let $X_1, X_2, ..., X_{n_1}$ be a set of n_1-independent random variables each belonging from a particular PHR family with parameters $(\alpha_1, \alpha_2, ..., \alpha_{n_1})$. We assume that $X_i \sim \overline{F}_i(x) = (\overline{F}_0(x))^{\alpha_i}$ for $i = 1, 2, ..., n_1$. Also, let $Y_1, Y_2, ..., Y_{n_2}$ be another set of n_2-independent random variables each following PHR family of distributions with a different distribution function and the parameter set is $(\beta_1, \beta_2 ..., \beta_{n_2})$. Let $Y_i \sim \overline{G}_i(x) = (\overline{G}_0(x))^{\beta_i}$ for $i = 1, 2, ..., n_2$. Under the assumption that $\sum_{i=1}^{n_2} \beta_i \geq \sum_{i=1}^{n_1} \alpha_i$, the baseline distribution function F_0 is DFR, and $G_0 \leq_{hr} F_0$ then $Y_{1:n_2} \leq_{\text{disp}} X_{1:n_1}$.*

Proof: The distribution function of $X_{1:n_1}$ and $Y_{1:n_2}$ are

$$\overline{F}_{1:n_1}(x) = \left(\overline{F}_0(x)\right)^{\sum_{i=1}^{n_1} \alpha_i}, \tag{3.1}$$

and,

$$\overline{G}_{1:n_2}(x) = \left(\overline{G}_0(x)\right)^{\sum_{i=1}^{n_2} \beta_i} \tag{3.2}$$

respectively. For simplicity we replace $\sum_{i=1}^{n_1} \alpha_i$ by α and $\sum_{i=1}^{n_2} \beta_i$ by β. Let

$$\begin{aligned} \psi_1(y) &= F_{1:n_1}^{-1}(y) - G_{1:n_2}^{-1}(y) \\ &= \overline{F}_0^{-1}\left((1-y)^{1/\alpha}\right) - \overline{G}_0^{-1}\left((1-y)^{1/\beta}\right) \end{aligned}$$

We are required to prove $Y_{1:n_2} \leq_{\text{disp}} X_{1:n_1}$, i.e., $\psi_1(y)$ is increasing in $y \in (0, 1)$. Hence $Y_{1:n_2} \leq_{\text{disp}} X_{1:n_1}$ if and only if $\phi_1(t) = \overline{F}_0^{-1}(t) - \overline{G}_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right)$

is decreasing in $t \in (0,1)$, where $t = (1-y)^{1/\alpha}$. Note that

$$\phi_1'(t) = -\frac{1}{f_0(\overline{F}_0^{-1}(t))} + \frac{\alpha}{\beta}\frac{t^{\frac{\alpha}{\beta}-1}}{g_0\left(\overline{G}_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right)\right)}. \tag{3.3}$$

We need to show that $\phi_1'(t) \leq 0$, i.e.,

$$\frac{t}{f_0(\overline{F}_0^{-1}(t))} \geq \frac{\alpha}{\beta}\frac{t^{\frac{\alpha}{\beta}}}{g_0\left(\overline{G}_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right)\right)}. \tag{3.4}$$

Let $\overline{F}_0^{-1}(t) = z_1$ and $\overline{G}_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right) = z_2$,

$$\begin{aligned} \frac{\overline{F}_0(z_1)}{f_0(z_1)} &\geq \frac{\alpha}{\beta}\frac{\overline{G}_0(z_2)}{g_0(z_2)} \\ \Rightarrow s_0(z_2)\frac{\beta}{\alpha} &\geq r_0(z_1), \end{aligned} \tag{3.5}$$

where $r_0(z_1) = \dfrac{f_0(z_1)}{\overline{F}_0(z_1)}$ and $s_0(z_2) = \dfrac{g_0(z_2)}{\overline{G}_0(z_2)}$. Under the hypothesis of the theorem

$$\begin{aligned} &\beta \geq \alpha \\ \Rightarrow t = \overline{F}_0(z_1) &\leq t^{\left(\frac{\alpha}{\beta}\right)} = \overline{G}_0(z_2). \end{aligned}$$

and $G_0 \leq_{hr} F_0$ implies that $z_2 \leq z_1$ ($G_0 \leq_{st} F_0$ follows from $G_0 \leq_{hr} F_0$ subsequently we can derive that $\overline{G}_0(z_2) \geq \overline{F}_0(z_1) \geq \overline{G}_0(z_1)$. Finally the implication is possible as $\overline{G}_0$ is a decreasing function) and $s_0(z_2) \geq r_0(z_2)$. Also F_0 is DFR then, $z_2 \leq z_1 \Rightarrow r_0(z_2) \geq r_0(z_1)$. Combining all these we find that (3.5) holds true. Hence the result. □

The above result provides a general outlook over the PHR distributions. Apart from the fact that the component lifetimes are independent, the result can be compared with Theorem 3.11 from [31]. Here the baseline distributions are different, also the number of components are not same. The theorem holds true when we encounter a multiple-outlier model. An example has been provided here that satisfies the condition given in the above theorem.

Example 3.1. Let (X_1, X_2, X_3) and (Y_1, Y_2, Y_3) be independent Transformed Pareto distributed random variables. The survival function of X_i is $\overline{F}_{k_i}(t)$ and corresponding to Y_i is $\overline{G}_{k_i^*}(t)$ for $i = 1, 2, 3$. Consider $k_1 = 1.7, k_2 = 2, k_3 = 0.9$ and $k_1^* = 1, k_2^* = 3, k_3^* = 2.3$, here $\sum_{i=1}^{3} k_i = 4.6$ and $\sum_{i=1}^{3} k_i^* = 6.3$.

Let us consider $\overline{F}_{k_i}(t) = (\overline{F}_0(t))^{k_i}$, where $\overline{F}_0(t) = \dfrac{1}{(1+t)^2}, t > 0$; $\overline{G}_{k_i^*}(t) = (\overline{G}_0(t))^{k_i^*}$, where

$\overline{G}_0(t) = \dfrac{1}{(1+t)^3}, t > 0.$

Here, F_0 is DFR and the ratio $\dfrac{\overline{F}_0(t)}{\overline{G}_0(t)} = 1 + t$ is increasing in t $\forall t > 0$. Thus, $F_0 \geq_{hr} G_0$.

We observe that $\overline{F}_0^{-1}(u) = \left(\dfrac{1}{u}\right)^{\frac{1}{2}} - 1, 0 < u < 1$ and $\overline{G}_0^{-1}(u) = \left(\dfrac{1}{u}\right)^{\frac{1}{3}} - 1, 0 < u < 1$. Hence, as mentioned in Theorem 3.1, the expression

$$\begin{aligned}\psi_1(y) &= \overline{F}_0^{-1}((1-y)^{1/\alpha}) - \overline{G}_0^{-1}((1-y)^{1/\beta}) \\ &= \frac{1}{(1-y)^{1/2\alpha}} - \frac{1}{(1-y)^{1/3\beta}}, \alpha = \sum_{i=1}^{3} k_i \text{ and } \beta = \sum_{i=1}^{3} k_i^*. \qquad (3.6)\end{aligned}$$

Plotting (3.6) with respect to y, for $0 < y < 1$, we observe that $\psi_1(y)$ is increasing in y, i.e. the theorem holds true in this case.

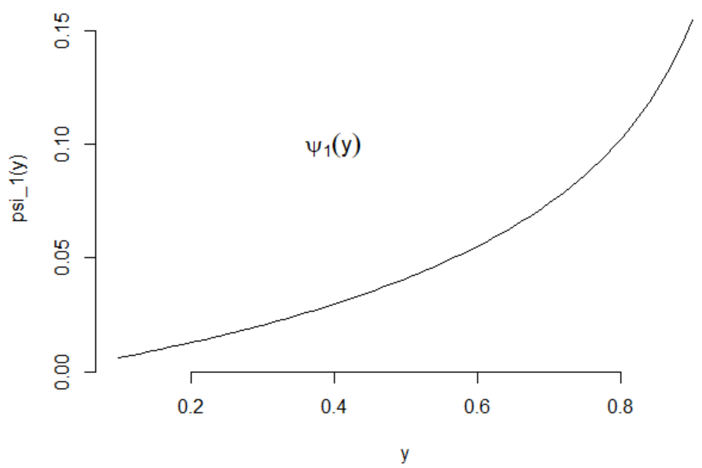


Figure 1: $\psi_1(y)$ is increasing for $0 < y < 1$.

The conditions "DFR" and hr order necessary in Theorem 3.1. Let us consider

$$\overline{F}_{1:n_1}(x) = \left(\overline{F}_0(x)\right)^{\sum_{i=1}^{n_1} \alpha_i}$$

and

$$\overline{G}_{1:n_2}(x) = \left(\overline{G}_0(x)\right)^{\sum_{i=1}^{n_2} \beta_i}$$

where $\overline{F}_0(x) = \exp(-x^2), x > 0$ and $\overline{G}_0(x) = \exp(-x), x > 0$; $\sum_{i=1}^{n_1} \alpha_i = 2$ and $\sum_{i=1}^{n_2} \beta_i = 2.5$.

Here $\dfrac{\overline{F}_0(x)}{\overline{G}_0(x)}$ is non-monotone and $\psi_1(y) = \overline{F}_0^{-1}\left((1-y)^{1/\sum_{i=1}^{n_1} \alpha_i}\right) - \overline{G}_0^{-1}\left((1-y)^{1/\sum_{i=1}^{n_2} \beta_i}\right)$ is also non-monotone. The condition for DFR and hr order are not satisfied here (see Figure 2 and Figure 3), also the dispersive order does not hold in this situation even though the conditions for the parameters are satisfied. The plots are shown below:

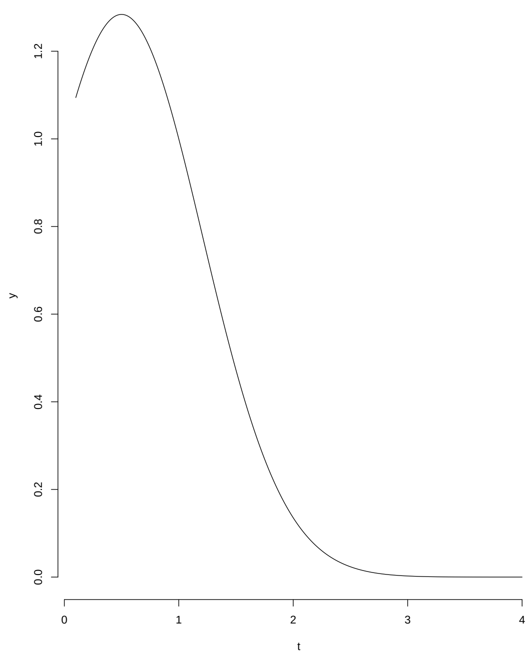


Figure 2: $\dfrac{\overline{F}_0(x)}{\overline{G}_0(x)}$ is non-monotone.

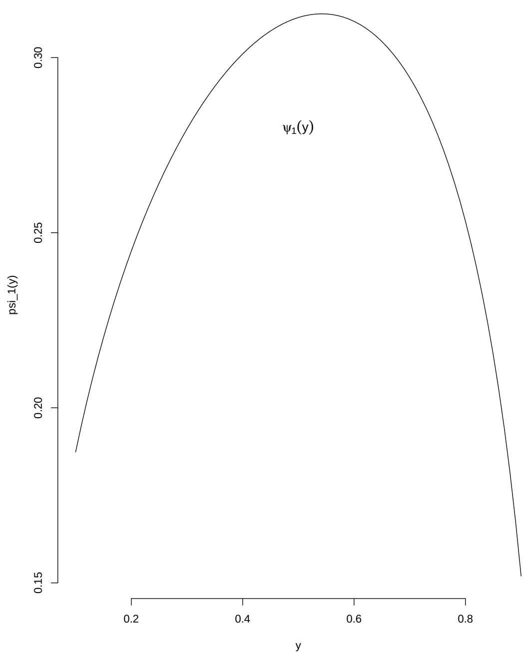


Figure 3: $\psi_1(y)$ is non-monotone.

Thus the conditions mentioned in Theorem 3.1 are necessary. Here we discuss the following result for a parallel system with the PRHR distributed components.

Theorem 3.2. *Let $X_1, X_2, ..., X_{n_1}$ be a n_1-independent set of random variables each belonging from PRHR family of distributions with parameters $(\alpha_1, \alpha_2..., \alpha_{n_1})$, such that $X_i \sim F_i(x) = (F_0(x))^{\alpha_i}$ for $i = 1, 2, ..., n_1$. Also $Y_1, Y_2, ..., Y_{n_2}$ be another set of n_2-independent random variables each following PRHR family of distributions with a different distribution function and the parameter set is $(\beta_1, \beta_2, ..., \beta_{n_2})$. Let $Y_i \sim G_i(x) = (G_0(x))^{\beta_i}$ for $i = 1, 2, ..., n_2$. Then $Y_{n_2:n_2} \leq_{\text{disp}} X_{n_1:n_1}$ if the baseline distribution F_0 follows IRHR model, $\sum_{i=1}^{n_2} \beta_i \geq \sum_{i=1}^{n_1} \alpha_i$ and $F_0 \leq_{rh} G_0$.*

Proof: The distribution function of $X_{n_1:n_1}$ and $Y_{n_2:n_2}$ are

$$F_{n_1:n_1}(x) = [F_0(x)]^{\sum_{i=1}^{n_1} \alpha_i}, \text{ and,}$$
$$G_{n_2:n_2}(x) = [G_0(x)]^{\sum_{i=1}^{n_2} \beta_i}$$

respectively. Similar to the previous theorem, we take $\sum_{i=1}^{n_1} \alpha_i = \alpha$ and $\sum_{i=1}^{n_2} \beta_i = \beta$. Let

$$\begin{aligned}\psi_2(y) &= F^{-1}_{n_1:n_1}(y) - G^{-1}_{n_2:n_2}(y)\\ &= F_0^{-1}(y^{1/\alpha}) - G_0^{-1}(y^{1/\beta}).\end{aligned}$$

We are required to prove that $Y_{n_2:n_2} \leq_{\text{disp}} X_{n_1:n_1}$, i.e., $\psi_2(y)$ is increasing in $y \in (0, 1)$. Hence $Y_{n_2:n_2} \leq_{\text{disp}} X_{n_1:n_1}$ if and only if $\phi_2(t) = F_0^{-1}(t) - G_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right)$ is increasing in $t \in (0, 1)$, where $t = y^{1/\alpha}$. Note that

$$\phi_2'(t) = \frac{1}{f_0(F_0^{-1}(t))} - \frac{\alpha}{\beta} \frac{t^{\frac{\alpha}{\beta}-1}}{g_0\left(G_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right)\right)}.$$

We need to show that $\phi_2'(t) \geq 0$, i.e.,

$$\frac{t}{f_0(F_0^{-1}(t))} \geq \frac{\alpha}{\beta} \frac{t^{\frac{\alpha}{\beta}}}{g_0\left(G_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right)\right)}. \tag{3.7}$$

Put $F_0^{-1}(t) = z_1$ and $G_0^{-1}\left(t^{\frac{\alpha}{\beta}}\right) = z_2$. From (3.7) it is sufficient to show

$$\begin{aligned}&\frac{F_0(z_1)}{f_0(z_1)} \geq \frac{\alpha}{\beta}\frac{G_0(z_2)}{g_0(z_2)}\\ &\Leftrightarrow \tilde{s}_0(z_2)\frac{\beta}{\alpha} \geq \tilde{r}_0(z_1).\end{aligned} \tag{3.8}$$

As

$$\frac{\beta}{\alpha} \geq 1$$
$$\Rightarrow t = F_0(z_1) \leq t^{\frac{\alpha}{\beta}} = G_o(z_2).$$

Since $F_0 \leq_{rh} G_0$ implies $F_0 \leq_{st} G_0$, hence $G_0(z_1) \leq F_0(z_1) \leq G_0(z_2)$ i.e., $z_1 \leq z_2$. Again F_0 follows increasing reversed hazard rate (IRHR) model hence $z_1 \leq z_2 \Rightarrow \tilde{r}_0(z_1) \leq \tilde{r}_0(z_2)$. Lastly, $F_0 \leq_{rh} G_0 \Rightarrow \tilde{r}_0(x) \leq \tilde{s}_0(x)$ for all x, thus $\tilde{r}_0(z_2) \leq \tilde{s}_0(z_2)$. Combining these inequalities we obtain the required result. □

Example 3.2. We have observed an example of IRHR distribution from Example 3.4 of [4]. Let X be a random variable following Truncated Normal(μ, σ^2) distribution with distribution function as

$$F(x) = \frac{\Phi\left(\frac{x-\mu}{\sigma}\right)}{\Phi\left(\frac{-\mu}{\sigma}\right)}, x \in (-\infty, 0]$$

Let us consider a set of 3 independent random variables X_1, X_2, X_3 such that $X_i \sim F_i(x) = [F_0(x)]^{\alpha_i}, i = 1, 2, 3$ where F_0 corresponds to Truncated Normal(0, 4) i.e., $F_0(x) = \frac{\Phi\left(\frac{x}{2}\right)}{0.5}$. Let us consider another set of 2 independent random variables Y_1, Y_2 such that $Y_i \sim G_i(x) = [G_0(x)]^{\beta_i}, i = 1, 2, 3$ where G_0 corresponds to Truncated Normal(0, 1) i.e., $G_0(x) = \frac{\Phi(x)}{0.5}$. We can observe that the reversed hazard rate function of the baseline distributions F_0 and G_0 are

$$\tilde{h}_{F_0}(x) = \frac{1}{2}\frac{\phi(x)}{\Phi(x)},$$
$$\tilde{h}_{G_0}(x) = \frac{\phi(x)}{\Phi(x)}.$$

Thus, $F_0 \leq_{rh} G_0$. The distribution function of $X_{3:3}$ and $Y_{2:2}$ are

$$F_{3:3}(x) = \left(\frac{\Phi\left(\frac{x}{2}\right)}{0.5}\right)^{\sum_{i=1}^{3}\alpha_i} \text{ and } G_{2:2}(x) = \left(\frac{\Phi(x)}{0.5}\right)^{\sum_{i=1}^{2}\beta_i}.$$

Taking $\sum_{i=1}^{3}\alpha_i = \alpha = 2$ and $\sum_{i=1}^{2}\beta_i = \beta = 3$. Here all the conditions of Theorem 3.2 are satisfied, further we observe that

$$\begin{aligned}\psi_2(y) &= F_{3:3}^{-1}(y) - G_{2:2}^{-1}(y) \\ &= F_0^{-1}(y^{1/\alpha}) - G_0^{-1}(y^{1/\beta}) \\ &= 2\Phi^{-1}(0.5y^{1/2}) - \Phi^{-1}(0.5y^{1/3})\end{aligned}$$

is increasing in $y \in (0, 1)$.

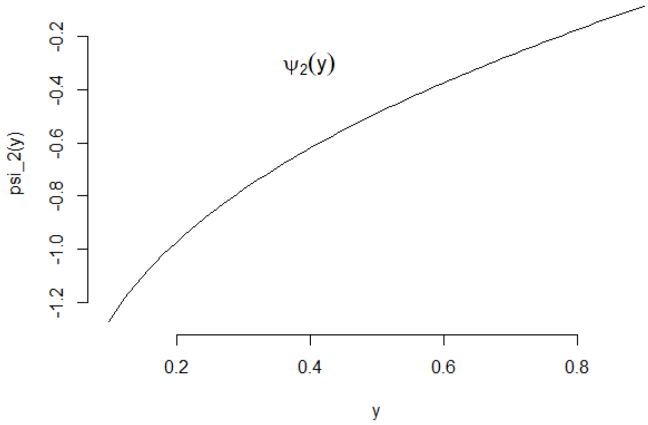


Figure 4: $\psi_2(y)$ is increasing for $0 < y < 1$.

As a corollary, we can obtain a result for multiple-outlier model from PRHR distributions. The condition IRHR is necessary here, we can understand this through an example. Let (X_1, X_2, X_3) and (Y_1, Y_2, Y_3) be independent random variables such that the distribution function of X_i is $(F_0(x))^{\alpha_i}$, where $F_0(x) = 1 - e^{-3x}, x > 0$ and distribution function of Y_i is $(G_0(x))^{\beta_i}$, where $G_0(x) = \dfrac{1}{(1+x)^2}, x > 0$. $F_0(x)$ is DRHR.

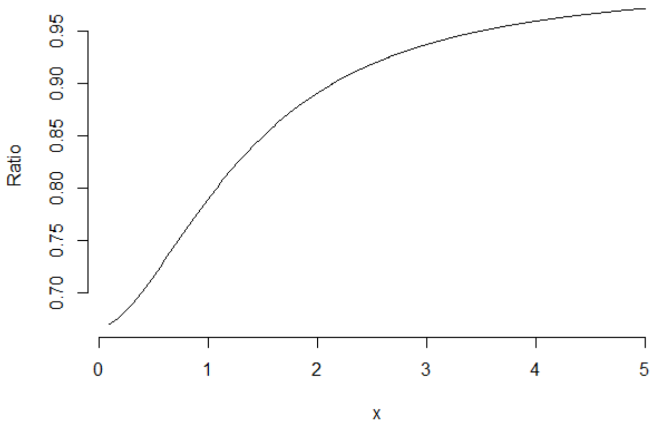


Figure 5: $\dfrac{G_0(x)}{F_0(x)}$ is increasing for $x > 0$.

Figure 5 represents that $F_0 \leq_{rh} G_0$. Let $\alpha = \sum_{i=1}^{n} \alpha_i$ and $\beta = \sum_{i=1}^{n} \beta_i$, then

$$\begin{aligned}\psi_2(y) &= F_0^{-1}(y^{1/\alpha}) - G_0^{-1}(y^{1/\beta}) \\ &= 1 + \frac{1}{3} ln(1 - y^{1/4.6}) - \frac{1}{(1 - y^{1/6.3})^{1/2}}, 0 < y < 1\end{aligned}$$

Thus from Figure 6, we observe that $Y_{3:3} \geq_{\text{disp}} X_{3:3}$, i.e. the inequality reverses.

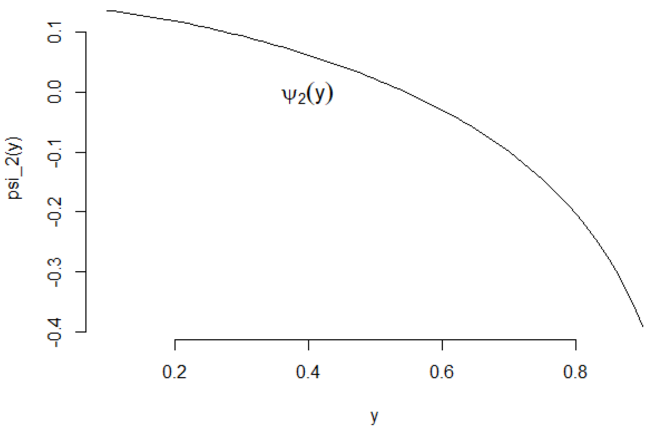


Figure 6: $\psi_2(y)$ is decreasing for $0 < y < 1$.

In the next theorem, we provide a result for series systems with unequal number of components following PHR models with different baseline distributions. It can be noted that there is some relationship between hazard rate ordering and dispersive ordering. If X and Y are two non negative random variables then:

1. If $X \leq_{\text{hr}} Y$ and X or Y is DFR, then $X \leq_{\text{disp}} Y$;
2. If $X \leq_{\text{disp}} Y$ and X or Y is IFR, then $X \leq_{\text{hr}} Y$;

from theorem 3.B.20 of [26] and Corollary 4.3 of [7].

Theorem 3.3. *Consider a system of n_1 components, where the lifetime of each component is represented by the random variable $X_1, X_2, ..., X_{n_1}$ respectively such that each of $X_1, ..., X_{p_1}$ has survival function $[\overline{F}(x)]^{\alpha_i}$, $i = 1, 2, ..., p_1$ and $X_{p_1+1}, ..., X_{n_1}$ has survival function $[\overline{G}(x)]^{\alpha_i}$, $i = p_1 + 1, p_1 + 2, ..., n_1$. Similarly another system with n_2 components is considered where the components $Y_1, ..., Y_{p_2}$ has survival function as $[\overline{F}(x)]^{\beta_i}$, $i = 1, 2, ..., p_2$ whereas the components $Y_{p_2+1}, ..., Y_{n_2}$ has survival function $[\overline{G}(x)]^{\beta_i}$, $i = p_2 + 1, p_2 + 2, ..., n_2$. Then $X_{1:n_1} \leq_{hr} Y_{1:n_2}$ whenever $\sum_{i=1}^{p_1} \alpha_i > \sum_{i=1}^{p_2} \beta_i$ and $\sum_{i=p_1+1}^{n_1} \alpha_i > \sum_{i=p_2+1}^{n_2} \beta_i$.*

Proof: The survival function of $X_{1:n_1}$ is

$$\overline{F}_{1:n_1}(x) = [\overline{F}(x)]^{\sum_{i=1}^{p_1} \alpha_i} [\overline{G}(x)]^{\sum_{i=p_1+1}^{n_1} \alpha_i},$$

and the survival function of $Y_{1:n_2}$ is

$$\overline{G}_{1:n_2}(x) = [\overline{F}(x)]^{\sum_{i=1}^{p_2} \beta_i} [\overline{G}(x)]^{\sum_{i=p_2+1}^{n_2} \beta_i}.$$

Consider the ratio

$$\frac{\overline{F}_{1:n_1}(x)}{\overline{G}_{1:n_2}(x)} = [\overline{F}(x)]^{\left(\sum_{i=1}^{p_1}\alpha_i - \sum_{i=1}^{p_2}\beta_i\right)} [\overline{G}(x)]^{\left(\sum_{i=p_1+1}^{n_1}\alpha_i - \sum_{i=p_2+1}^{n_2}\beta_i\right)} \tag{3.9}$$

Differentiating (3.9) with respect to x,

$$\frac{d}{dx}\left(\frac{\overline{F}_{1:n_1}(x)}{\overline{G}_{1:n_2}(x)}\right) = -\frac{\overline{F}_{1:n_1}(x)}{\overline{G}_{1:n_2}(x)}\left(\left(\sum_{i=1}^{p_1}\alpha_i - \sum_{i=1}^{p_2}\beta_i\right)\frac{f(x)}{\overline{F}(x)} + \left(\sum_{i=p_1+1}^{n_1}\alpha_i - \sum_{i=p_2+1}^{n_2}\beta_i\right)\frac{g(x)}{\overline{G}(x)}\right)$$
$$< 0,$$

whenever $\sum_{i=1}^{p_1}\alpha_i > \sum_{i=1}^{p_2}\beta_i$ and $\sum_{i=p_1+1}^{n_1}\alpha_i > \sum_{i=p_2+1}^{n_2}\beta_i$. Hence the result follows. □

When the random variables $X_{1:n_1}$ or $Y_{1:n_2}$ is DFR then $X_{1:n_1} \leq_{\text{disp}} Y_{1:n_2}$.

If we consider a similar problem wherein the random variables follows a PRHR distribution, we arrive at the following theorem.

Theorem 3.4. *Consider an independent set of n_1 random variables $X_1, X_2, ..., X_{p_1}$, $X_{p_1+1}, ..., X_{n_1}$ such that the distribution function of X_i, $F_{X_i}(x) = [F(x)]^{\alpha_i}$ for $i = 1, 2, ..., p_1$ and $F_{X_i}(x) = [G(x)]^{\alpha_i}$ for $i = p_1+1, ..., n_1$. Another set of n_2 independent components $Y_1, Y_2, ..., Y_{p_2}, Y_{p_2+1}, ..., Y_{n_2}$ are such that the distribution function of Y_i, $F_{Y_i}(x) = [F(x)]^{\beta_i}$, $i = 1, 2, ..., p_2$ and $F_{Y_i}(x) = [G(x)]^{\beta_i}$ for $i = p_2+1, ..., n_2$. Then $X_{n_1:n_1} \geq_{rh} Y_{n_2:n_2}$ whenever $\sum_{i=1}^{p_1}\alpha_i > \sum_{i=1}^{p_2}\beta_i$ and $\sum_{i=p_1+1}^{n_1}\alpha_i > \sum_{i=p_2+1}^{n_2}\beta_i$.*

Proof: The distribution functions of $X_{n_1:n_1}$ and $Y_{n_2:n_2}$ are

$$F_{n_1:n_1}(x) = [F(x)]^{\sum_{i=1}^{p_1}\alpha_i}[G(x)]^{\sum_{i=p_1+1}^{n_1}\alpha_i},$$

$$G_{n_2:n_2}(x) = [F(x)]^{\sum_{i=1}^{p_2}\beta_i}[G(x)]^{\sum_{i=p_2+1}^{n_2}\beta_i}$$

respectively. Differentiating the ratio $\frac{F_{n_1:n_1}(x)}{G_{n_2:n_2}(x)}$ with respect to x, we observe,

$$\frac{d}{dx}\left(\frac{F_{n_1:n_1}(x)}{G_{n_2:n_2}(x)}\right) = \frac{F_{n_1:n_1}(x)}{G_{n_2:n_2}(x)}\left(\left(\sum_{i=1}^{p_1}\alpha_i - \sum_{i=1}^{p_2}\beta_i\right)\frac{f(x)}{F(x)} + \left(\sum_{i=p_1+1}^{n_1}\alpha_i - \sum_{i=p_2+1}^{n_2}\beta_i\right)\frac{g(x)}{G(x)}\right)$$
$$> 0,$$

whenever $\sum_{i=1}^{p_1}\alpha_i > \sum_{i=1}^{p_2}\beta_i$ and $\sum_{i=p_1+1}^{n_1}\alpha_i > \sum_{i=p_2+1}^{n_2}\beta_i$ and the result follows. □

The above theorem deals with a different set of parameters and baseline distributions as compared to that of theorem 3.7 from [31] where the component lifetimes are dependent but the parameters are restricted and the baseline distributions are all same. If $X_{n_1:n_1}$ or $Y_{n_2:n_2}$ is IRFR then from the above theorem we can observe that $X_{n_1:n_1} \leq_{\text{disp}} Y_{n_2:n_2}$.
We can observe the inter-relationship between reversed hazard rate ordering and dispersive ordering, as mentioned in Corollary 4.4 of [7]. For two random variables X and Y,

1. If $X \leq_{\text{rh}} Y$ and X or Y is IRFR, then $Y \leq_{\text{disp}} X$;
2. If $X \leq_{\text{disp}} Y$ and X or Y is DRFR, then $Y \leq_{\text{rh}} X$.

It is interesting to note that in [21], Proposition 4.4 can be realized from Theorem 3.3 and 3.4. Such as, if $p_1 = n$ and $p_2 = n$ in theorem 3.3 then

$$X_{1:n} \leq_{hr} Y_{1:n} \text{ whenever } \sum_{i=1}^{n} \alpha_i > \sum_{i=1}^{n} \beta_i,$$

and if $p_1 = n$ and $p_2 = n$ in theorem 3.4 then

$$X_{n:n} \geq_{rh} Y_{n:n} \text{ whenever } \sum_{i=1}^{n} \alpha_i > \sum_{i=1}^{n} \beta_i.$$

Next we consider a reversed hazard rate ordering result for the parallel system having Pareto distributed components such that the sample sizes are equal. Pareto distribution is DRFR hence we have obtained a reversed hazard rate ordering for $X_{n:n}$ and $Y_{n:n}$.

Theorem 3.5. *Let $X_1, X_2, ..., X_n$ and $Y_1, Y_2, ..., Y_n$ be two sets of n-independent Pareto distributed random variables such that the survival function of X_i is $\overline{F}_i(x) = \left(1 + \frac{x}{\theta}\right)^{-\alpha_i}$, $x > 0$, $\theta > 0$, $\alpha_i > 0$ and that of Y_i is $\overline{G}_i(x) = \left(1 + \frac{x}{\theta}\right)^{-\alpha_i^*}$, $x > 0$, $\theta > 0$, $\alpha_i^* > 0$. Let $\underline{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)$, $\underline{\alpha}^* = (\alpha_1^*, \alpha_2^*, ..., \alpha_n^*)$, then $\underline{\alpha} \prec^w \underline{\alpha}^* \Rightarrow X_{n:n} \leq_{rh} Y_{n:n}$.*

Proof: The distribution function of $X_{n:n}$ is

$$F_{X_{n:n}}(x) = \prod_{i=1}^{n} \left[1 - \left(1 + \frac{x}{\theta}\right)^{-\alpha_i}\right], \tag{3.10}$$

and the corresponding reversed hazard rate function is

$$\tilde{r}_{X_{n:n}}(x) = \frac{1}{x + \theta} \sum_{i=1}^{n} g(\alpha_i), \tag{3.11}$$

where $g(\alpha) = \dfrac{\alpha}{\left(\frac{x}{\theta} + 1\right)^{\alpha} - 1}$. Let $u = \left(\frac{x}{\theta} + 1\right)^{\alpha}$ and $u > 1$ such that $g(\alpha) = \dfrac{\alpha}{u^{\alpha} - 1}$. Now,

$$g'(\alpha) = \frac{u^{\alpha}(1 - \alpha \ln u) - 1}{(u^{\alpha} - 1)^2} \tag{3.12}$$

and

$$g''(\alpha) = \frac{u^{\alpha} \ln u((u^{\alpha} \ln u + \ln u)\alpha - 2u^{\alpha} + 2)}{(u^{\alpha} - 1)^3}. \tag{3.13}$$

$g''(\alpha) \overset{sgn}{=} u^{\alpha}(\ln u)\phi(u)$, where $\phi(u) = (u^{\alpha}\ln u + \ln u)\alpha - 2u^{\alpha} + 2$, $\phi(1) = 0$. Also,

$$\begin{aligned}\phi'(u) &= \alpha^2 u^{\alpha-1}\ln u + \frac{\alpha}{u} - \alpha u^{\alpha-1}\\ &= \frac{\alpha}{u}\phi_1(u),\end{aligned}$$

such that $\phi_1(u) = \alpha u^{\alpha}\ln u + 1 - u^{\alpha}$ and $\phi_1(1) = 0$. And

$$\begin{aligned}\phi_1'(u) &= \alpha^2 u^{\alpha-1}\ln u\\ &> 0.\end{aligned}$$

Hence it is observed that $g''(\alpha) > 0$ for $x > 0$ $(u > 1)$, i.e., $g(\alpha)$ is convex in α. Hence, using Lemma 2.2 we obtain, $\tilde{r}_{X_{n:n}}(x)$ is Schur convex w.r.t $\underline{\alpha}$. Moreover,

$$\begin{aligned}g'(\alpha) &\overset{sgn}{=} u^{\alpha}(1-\alpha\ln u) - 1,\ u > 1\\ &= h(u) \text{ say,}\end{aligned}$$

then $h'(u) = -\alpha^2 u^{\alpha-1}\ln u$. Also $h(1) = 0$, then $g'(\alpha) < 0$ for $x > 0$ $(u > 1)$. Thus $\tilde{r}_{X_{n:n}}(x)$ is decreasing in $\underline{\alpha}$ and Schur convex w.r.t $\underline{\alpha}$. Using Lemma 2.3, we infer that $\underline{\alpha} \prec^{w} \underline{\alpha}^*$ $\Rightarrow$ $\tilde{r}_{X_{n:n}}(x) \le \tilde{r}_{Y_{n:n}}(x)$. Hence the result follows. □

3.2. Star ordering result for unequal sample sizes

In this section we present a comparison between two systems based on star ordering. Consider a series system with components following PHR model and have unequal sample sizes.

Theorem 3.6. *Let $X_1, X_2, ..., X_{n_1}$ be a n_1-independent set of non-negative random variables such that $X_i \sim [\overline{F}(x)]^{\alpha_i}$ for $i = 1, 2, ..., n_1$ and $Y_1, Y_2, ..., Y_{n_2}$ be another n_2-independent set of non-negative random variables such that $Y_i \sim [\overline{F}(x)]^{\beta_i}$ for $i = 1, 2, ..., n_2$, where n_1 and n_2 may or may not be the same. Then*

$$\sum_{i=1}^{n_1}\alpha_i \le \sum_{i=1}^{n_2}\beta_i \Rightarrow X_{1:n_1} \ge_* Y_{1:n_2}, \quad \text{whenever } xr(x) \text{ is decreasing.}$$

Proof: The survival function of $X_{1:n_1}$ is

$$\overline{F}_{1:n_1}(x) = [\overline{F}(x)]^{\sum_{i=1}^{n_1}\alpha_i}. \tag{3.14}$$

Let $\sum_{i=1}^{n_1}\alpha_i = \alpha$, then $\overline{F}_{1:n_1}(x) = [\overline{F}(x)]^{\alpha} = \overline{F}_{\alpha}(x)$ (say).
The corresponding probability density function is

$$\begin{aligned}f_{X_{1:n_1}}(x) &= \alpha f(x)[\overline{F}(x)]^{\alpha-1}\\ &= f_{\alpha}(x).\end{aligned}$$

Note that the ratio

$$\frac{F'_\alpha(x)}{f_\alpha(x)} = -\frac{1}{\alpha}\frac{\ln \overline{F}(x)}{r(x)}, \tag{3.15}$$

where $F_\alpha(x) = 1 - [\overline{F}(x)]^\alpha$ and $F'_\alpha(x) = \frac{d}{d\alpha}F_\alpha(x)$.
The theorem follows by differentiating the ratio $\dfrac{F'_\alpha(x)}{xf_\alpha(x)}$ with respect to x.
Note that

$$\frac{d}{dx}\left(\frac{F'_\alpha(x)}{xf_\alpha(x)}\right) = \frac{1}{\alpha}\left(\frac{x(r(x))^2 + (xr'(x) + r(x))\ln \overline{F}(x)}{(xr(x))^2}\right)$$
$$> 0,$$

whenever $xr(x)$ is decreasing in x. Now using Lemma 2.4, we obtain $X_{1:n_1} \geq_* Y_{1:n_2}$ whenever $\sum_{i=1}^{n_1}\alpha_i \leq \sum_{i=1}^{n_2}\beta_i$. □

We observe here that the hazard rate functions of $X_{1:n_1}$ and $Y_{1:n_2}$ are $r_{X_{1:n_1}}(x) = \sum_{i=1}^{n_1}\alpha_i r(x)$ and $r_{Y_{1:n_2}}(x) = \sum_{i=1}^{n_2}\beta_i r(x)$ respectively, where $r(x)$ is the hazard rate function of baseline distribution $F(x)$. Then

$$r_{X_{1:n_1}}(x) \leq r_{Y_{1:n_2}}(x) \text{ whenever } \sum_{i=1}^{n_1}\alpha_i \leq \sum_{i=1}^{n_2}\beta_i.$$

The class of decreasing proportional hazard rate has been studied by [22] where several examples are also provided. The above result is applicable for multiple-outlier models. Moreover this theorem can be considered as a more general form of theorem 3.9 from [31]. Here the parameters are all different and only a simple inequality exists between them.

3.3. Dependent model

In this section we have considered a dependent set of random variables instead of independent random variables as discussed in the earlier sections. [12] studied scaled samples with proportional hazard and proportional reversed hazard rate models whereas [30] studied stochastic ordering results of Resilience-scaled(RS) models ($X \sim RS(\alpha, \lambda)$ if $F_X(x) = F^\alpha(\lambda x),\ \alpha > 0,\ \lambda > 0$) for series and parallel systems with dependent set of components. Moreover, [13] and [14] have discussed about the stochastic ordering between two systems where the component lifetimes are independent and each belongs from a location-scale family, necessarily with the same baseline distribution function. [18] discussed stochastic ordering results for series system from dependent and independent random variables following location-scale family of distributions. Thus it might be interesting to study the conditions under which a series (parallel) system can be compared with another series (parallel) system, where all the component lifetimes are dependent and each belonging from location family of distributions, the baseline distribution functions for both the sets are also different.
Hence we shall observe few definitions required especially to study the dependent models.

Definition 3.1. Survival copula: Let $(X_1, ..., X_n)$ be a n-dimensional random vector defined on a probability space $(\Omega, \mathbb{F}, \mathbb{P})$, the multivariate survival function is defined as

$$\begin{aligned}\overline{F}(x_1, ..., x_n) &= P[X_1 > x_1, ..., X_n > x_n] \\ &= \tilde{C}(\overline{F}_1(x_1), ..., \overline{F}_n(x_n)),\ x_1, ..., x_n \in \mathbb{R},\end{aligned}$$

where $\tilde{C}$ is the n-dimensional survival copula of the random vector $(X_1, ..., X_n)$. $\tilde{C}$ is a continuous function defined over the n-dimensional space as $\tilde{C} : [0, 1]^n \mapsto [0, 1]$, to develop multivariate survival functions from the marginal survival functions.

Archimedean copula is a very widely used class of survival copula because of its analytical tractability.

Definition 3.2. Archimedean copula

A n-dimensional *Archimedean copula* $\tilde{C} : [0, 1]^n \mapsto [0, 1]$ is represented as

$$\tilde{C}(u_1, ..., u_n) = \psi(\psi^{-1}(u_1) + ... + \psi^{-1}(u_n)),\ u_k \in [0, 1] \text{ for } k = 1, ..., n,$$

where the survival copula $\tilde{C}$ is generated by the generator function (also known as Archimedean generator function) $\psi : [0, \infty) \mapsto [0, 1]$, ψ is n -monotone $(n \geq 2)$ over an open interval $I \subset \mathbb{R}$ (where the end points of the interval I belongs to the limit point of $\mathbb{R}$) if ψ has derivatives upto order $n - 2$ and

$$(-1)^r \psi^{(r)}(x) \geq 0 \ \text{ for } r = 0, 1, 2, ..., n - 2$$

for any $x \in I$ and also $(-1)^{(n-2)}\psi^{(n-2)}$ is non-increasing and convex over I. $\phi = \psi^{-1}$ is the corresponding inverse function. Clayton copula, Frank copula are few archimedean copulas studied in the literature.

For a detailed discussion on Archimedean Copula one can refer to [20]. Recently, [27] have published results for systems with heterogeneous, dependent and distribution-free components. The following two propositions are mentioned here, the proofs of these propositions can be easily derived from the proof of propositions 3.16 and 3.7 from [27].

Proposition 3.1. *Let $Y_1, Y_2, ..., Y_n$ be n random variables such that $Y_i = X - \mu_i$, $(P[X > x] = \overline{F}(x))$ where μ_i for $i = 1, 2, ..., n$ are the corresponding location parameters respectively, then the survival function of the minimum of $Y_1, Y_2, ..., Y_n$ $(P[\min\{Y_1, Y_2, ..., Y_n\} > x])$ is given by*

$$J_1(\underline{\mu}; \overline{F}(x), \psi_1) = \psi_1(\sum_{k=1}^{n} \phi_1(\overline{F}(x + \mu_k))),$$

ψ_1 is log-convex (log-concave) and F is IFR (DFR) distribution. If there exists another set of n random variables $Z_1, Z_2, ..., Z_n$ $(Z_i = W - \mu_i^*$ and $P[W > x] = \overline{G}(x))$ such that the survival function for the minimum of $Z_1, Z_2, ..., Z_n$ is

$$J_1(\underline{\mu}^*; \overline{G}(x), \psi_2) = \psi_2(\sum_{k=1}^{n} \phi_2(\overline{G}(x + \mu_k^*))),$$

then as $(\mu_1, \mu_2, ..., \mu_n) \prec_w (\prec^w)$ $(\mu_1^*, \mu_2^*, ..., \mu_n^*)$ we obtain $Y_{1:n} \geq_{st} (\leq_{st}) Z_{1:n}$ as ψ is log-convex (log-concave), $X \geq_{st} W$ and F is IFR (DFR) distribution and $\phi_1 \cdot \psi_2 (\phi_2 \cdot \psi_1)$ is super-additive.

Proposition 3.2. *Let $Y_1, ..., Y_n$ and $Z_1, ..., Z_n$ be two n-dimensional random variables such that $Y_i = X - \mu_i$ and $Z_i = W - \mu_i*$, $i = 1, 2, ..., n$. Then*

$$\underline{\mu} \prec_w \underline{\mu}^* \ (\underline{\mu} \prec^w \underline{\mu}^*) \ \Rightarrow Y_{n:n} \geq_{st} (\leq_{st}) Z_{n:n}$$

whenever $\psi_1$or ψ_2 is log-convex (log-concave), F is IRFR(DRFR) distribution, $X \geq_{st} (\leq_{st}) W$ and $\phi_2 \cdot \psi_1 (\phi_1 \cdot \psi_2)$ is super additive.

Here the condition "$\phi_2 \cdot \psi_1$ is super-additive" is necessary. Let us consider 2 Archimedean copula generators as

$$\phi_1(t) = (-\ln t)^2, \ t \in (0, 1] \text{ and } \phi_2(t) = (1 - t)^3, \ t \in (0, 1].$$

The corresponding inverses are

$$\psi_1(t) = \exp(-t^{1/2}) \text{ and } \psi_2(t) = 1 - t^{1/3}.$$

We can observe that ψ_1 is log-convex. We are ineterested in finding the sign of the difference term

$$\phi_2 \cdot \psi_1(t_1 + t_2) - \phi_2 \cdot \psi_1(t_1) - \phi_2 \cdot \psi_1(t_2).$$

Figure 7 shows that the generators are chosen such that $\phi_2 \cdot \psi_1$ is not super-additive.

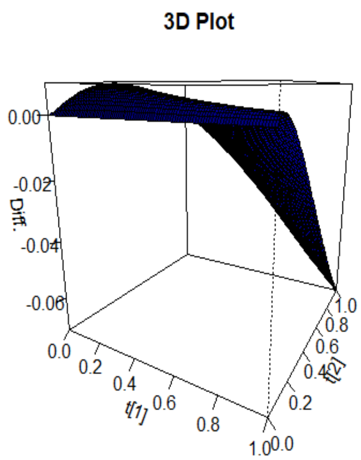


Figure 7: $\phi_2 \cdot \psi_1$ is not super additive.

As mentioned in the proposition, we shall consider $Y_i = X - \mu_i$ and $Z_i = W - \mu_i^*$ for $i = 1, 2, 3$. The location parameters are $\underline{\mu} = (0.5, 1, 2)$ and $\underline{\mu}^* = (1, 2, 3)$, thus $\underline{\mu} \prec_w \underline{\mu}^*$. The cdf of X and W are respectively given by

$$F(x) = \frac{\Phi\left(\frac{x}{2}\right)}{0.5}, x \in (-\infty, 0] \text{ and } G(x) = \frac{\Phi(x)}{0.5}, x \in (-\infty, 0].$$

The cdf of $Y_{3:3}$ is

$$J_2(\underline{\mu}; F(x), \psi_1) = 1 - \psi_1\left(\sum_{k=1}^{3}\left(-\ln \frac{\Phi\left(\frac{x + \mu_k}{2}\right)}{0.5}\right)^2\right).$$

The cdf of $Z_{3:3}$ is

$$J_2(\underline{\mu}^*; G(x), \psi_2) = 1 - \psi_2\left(\sum_{k=1}^{3}\left(1 - \frac{\Phi(x+\mu_k^*)}{0.5}\right)^3\right).$$

We shall observe the difference between the above 2 terms in Figure 8. Thus when $\phi_2 \cdot \psi_1$ is not super additive, usual stochastic ordering does not exist between $Y_{3:3}$ and $Z_{3:3}$.

$$\begin{aligned} y_1 &= J_2(\underline{\mu}; F(x), \psi_1) - J_2(\underline{\mu}^*; G(x), \psi_2) \\ &= 1 - \left(\sum_{k=1}^{3}\left(1 - \frac{\Phi(x+\mu_k^*)}{0.5}\right)^3\right)^{1/3} - \exp\left(-\left(\sum_{k=1}^{3}\left(-\ln\frac{\Phi\left(\frac{x+\mu_k}{2}\right)}{0.5}\right)^2\right)^{1/2}\right). \end{aligned}$$

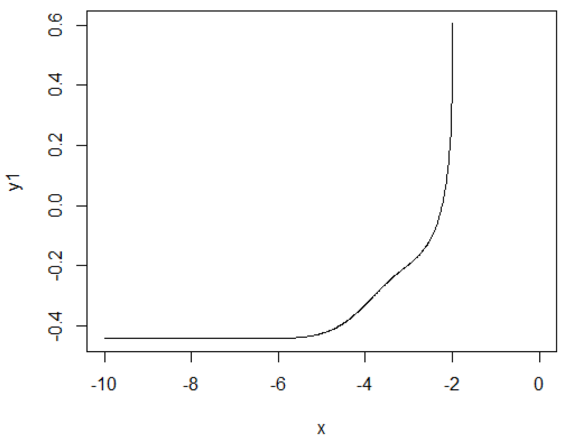


Figure 8: Usual stochastic ordering does not exist between $Y_{3:3}$ and $Z_{3:3}$.

When we take the generator function $\psi(x) = \exp(-x)$, $\phi(x) = -\ln x$. This generator indicates the independence copula (when the random variables are independent). Subsequently one can obtain the usual stochastic ordering between two sets of independent random variables.

Consider the Clayton copula generator function as

$$\psi_\theta(x) = \max((1+\theta x)^{-1/\theta}, 0), \;\; \theta > \;\; 0.$$

The above Archimedean generator is completely monotone (n-monotone for every $n \in \mathbb{N}$) for $\theta > 0$, and hence generates an Archimedean Copula. Here ψ_θ is a log-convex function, and hence the above theorems hold for this archimedean generator.

Examples: Let us consider $\phi(t) = (-\ln t)^\theta, \theta > 1, t \in (0, 1]$, the corresponding inverse function is $\psi(t) = e^{-t^{1/\theta}}, 0 \le t < \infty$. $\ln \psi(t)$ and its corresponding derivatives with respect to t

are

$$\begin{aligned}
\ln\psi(t) &= -t^{1/\theta}\\
\frac{d}{dt}(\ln\psi(t)) &= -\frac{1}{\theta}t^{-1+1/\theta}\\
\frac{d^2}{dt^2}(\ln\psi(t)) &= \frac{\theta-1}{\theta^2}t^{-2+1/\theta}.
\end{aligned}$$

We can observe that $\frac{d^2}{dt^2}(\ln\psi(t))$ is non-negative. Hence $\psi(t)$ is log-convex.

Let us consider $\phi(t) = \ln(1-\theta\ln t), \theta > 0, t \in (0,1]$, the corresponding inverse function is $\psi(t) = e^{\frac{1-e^t}{\theta}}, 0 \le t < \infty$. $\ln\psi(t)$ and its corresponding derivatives with respect to t are

$$\begin{aligned}
\ln\psi(t) &= \frac{1-e^t}{\theta}\\
\frac{d}{dt}(\ln\psi(t)) &= -\frac{e^t}{\theta}\\
\frac{d^2}{dt^2}(\ln\psi(t)) &= -\frac{e^t}{\theta}.
\end{aligned}$$

We can observe that $\frac{d^2}{dt^2}(\ln\psi(t))$ is non-positive. Hence $\psi(t)$ is log-concave.

4. CONCLUSION

Electronic devices, mechanical or electrical system consists of various units that are linked with one another either in series, parallel or any other combination, all of them are prone to failure at a certain point. We often refer to the warranty of the product to understand which system to purchase. Obviously any system which does not fail early is worth purchasing. If we are able to understand the dispersion of such a system compared to any other then we can compare two products. In order to understand the lifetime of any series or parallel system, we considered the random variables corresponding to the components. The results discussed in this paper can be divided into 3 subpart as Proportional Hazard rate (PHR) model, Proportional Reversed Hazard rate (PRHR) model, Dependent model. For PHR model we considered different models, a generalized situation where we consider two sets of independent PHR random variables and the baseline distribution for both the sets are different ($X_1, X_2, ..., X_{n_1}$ such that $X_i \sim \overline{F}_i(x) = (\overline{F}_0(x))^{\alpha_i}$ for $i = 1, 2, ..., n_1$ and another set $Y_1, Y_2, ..., Y_{n_2}$, $Y_i \sim \overline{G}_i(x) = (\overline{G}_0(x))^{\beta_i}$ for $i = 1, 2, ..., n_2$). We have obtained conditions over the parameters and the baseline distributions so that a dispersive ordering exist between the minimum order statistics. Whereas when both the baseline distributions are same, star ordering occurs between these minimum order statistics provided $xr(x)$ is decreasing. Since Pareto distribution is also PHR model, a reversed hazard rate ordering occurs between the sample maximums (also known as parallel systems) when the shape parameter varies. Proceeding similarly we have observed a result for PRHR model too. Here the two sets of random variables follow different baseline distributions and the number of samples are also unequal ($X_i \sim F_i(x) = (F_0(x))^{\alpha_i}$ for $i = 1, 2, ..., n_1$ and $Y_i \sim G_i(x) = (G_0(x))^{\beta_i}$ for $i = 1, 2, ..., n_2$). All of these results are true for multiple-outlier models.

Another form of generalized model has been studied where $X_1, ..., X_{p_1}$ has survival function $[\overline{F}(x)]^{\alpha_i}$ and $X_{p_1+1}, ..., X_{n_1}$ has survival function $[\overline{G}(x)]^{\alpha_i}$. Similarly another system with n_2 components is considered where the components $Y_1, ..., Y_{p_2}$ has survival function as $[\overline{F}(x)]^{\beta_i}$ whereas the components $Y_{p+1}, ..., Y_{n_2}$ has survival function $[\overline{G}(x)]^{\beta_i}$ and hazard rate ordering results has been observed for series systems. A reversed hazard rate ordering result with PRHR components has been observed.
In the last section, dependent random variables have been studied. Here we obtained usual stochastic ordering results between two sample minimums and two sample maximums such that the location parameter corresponding to the random variables from two sets obeys a weak majorization ordering while the baseline distribution obeys a usual stochastic ordering and the generating functions follows super-additive property.

ACKNOWLEDGMENTS

We would like to share our vote of thanks to all the reviewers for their valuable comment on our manuscript. The first author M.D. would like to thank IIT Kharagpur for research assistantship and ITER SOA Bhubaneswar for the necessary support.

REFERENCES

[1] Ali, D.; Towhidi, M. and Shekari, M. (2011). Stochastic and dependence comparisons between extreme order statistics in the case of proportional reversed hazard model, *Journal of the Iranian Statistical Society*, **10**(1), 29–43.

[2] Amini-Seresht, E.; Qiao, J.; Zhang, Y. and Zhao, P. (2016). On the skewness of order statistics in multiple-outlier PHR models, *Metrika*, **79**(7), 817–836.

[3] Balakrishnan, N.; Chen, J; Zhang, Y. and Zhao, P. (2019). Comparisons of sample ranges arising from multiple-outlier models: In memory of Moshe Shaked, *Probability in the Engineering and Informational Sciences*, 28–49.

[4] Balakrishnan, N. and Torrado, N. (2016). Comparisons between largest order statistics from multiple-outlier models, *Statistics* **50**(1), 176–189.

[5] Balakrishnan, N. and Zhao, P. (2013). Ordering properties of order statistics from heterogeneous populations: a review with an emphasis on some recent developments, *Probability in the Engineering and Informational Sciences*, **27**(4), 403–443.

[6] Barlow, R.E. and Proschan, F. (1975). *Statistical Theory of Reliability and Life Testing: Probability Models*, Florida State Univ. Tallahassee.

[7] Bartoszewicz, J. (1997). Dispersive functions and stochastic orders, *Applicationes Mathematicae*, **24**, 429–444.

[8] Block, H.W.; Savits, T.H. and Singh, H. (1998). The reversed hazard rate function, *Probability in the Engineering and Informational Sciences*, **12**(1), 69–90.

[9] Cai, X.; Zhang, Y. and Zhao, P. (2017). Hazard rate ordering of the second-order statistics from multiple-outlier PHR samples, *Statistics*, **51**(3), 615–626.

[10] Das, S. and Kayal, S. (2021). Some ordering results for the Marshall and Olkin's family of distributions, *Communications in Mathematics and Statistics*, **9**(2), 153–179.

[11] Datta, M. and Gupta, N. (2017). Stochastic comparison study of series and parallel systems having Kumaraswamy's and Frechet distributed components, *Journal of the Indian Statistical Association*, **55**(2), 103–126.

[12] Fang, R.; Li, C. and Li, X. (2018). Ordering results on extremes of scaled random variables with dependence and proportional hazards, *Statistics*, **52**(2), 458–478.

[13] Hazra, N.K.; Kuiti, M.R.; Finkelstein, M. and Nanda, A.K. (2017). On stochastic comparisons of maximum order statistics from the location-scale family of distributions, *Journal of Multivariate Analysis*, **160**, 31–41.

[14] Hazra, N.K.; Kuiti, M.R.; Finkelstein, M. and Nanda, A.K. (2018). On stochastic comparisons of minimum order statistics from the location-scale family of distributions, *Metrika*, **81**(2), 105–123.

[15] Jeon, J.; Kochar, S. and Park, C.G. (2006). Dispersive ordering – Some applications and examples, *Statistical Papers*, **47**(2), 227–247.

[16] Kochar, S.C. and Xu, M. (2007). Some recent results on stochastic comparisons and dependence among order statistics in the case of PHR model, *Journal of Reliability and Statistical Studies*, **6**(2), 125–140.

[17] Kochar, S. and Xu, M. (2009). Comparisons of parallel systems according to the convex transform order, *Journal of Applied Probability*, **46**(2), 342–352.

[18] Kundu, A. and Chowdhury, S. (2020). On stochastic comparisons of series systems with heterogeneous dependent and independent location-scale family distributed components, *Operations Research Letters*, **48**(1), 40–47.

[19] Marshall, A.W.; Olkin, I. and Arnold, B.C. (1979). *Inequalities: Theory of Majorization and its Applications*, Springer.

[20] McNeil, A.J. and Neslehova, J. (2009). Multivariate Archimedean copulas, d-monotone functions and l_1-norm symmetric distributions, *The Annals of Statistics*, **37**(5B), 3059–3097.

[21] Navarro, J.; Torrado, N. and del Aguila, Y. (2018). Comparisons between largest order statistics from multiple-outlier models with dependence, *Methodology and Computing in Applied Probability*, **20**(1), 411–433.

[22] Oliveira, P.E. and Torrado, N. (2015). On proportional reversed failure rate class, *Statistical Papers*, **56**(4), 999–1013.

[23] Păltănea, E. (2008). On the comparison in hazard rate ordering of fail-safe systems, *Journal of Statistical Planning and Inference*, **138**(7), 1993–1997.

[24] Pledger, P. and Proschan, F. (1971). Comparisons of order statistics and of spacings from heterogeneous distributions, *Optimizing Methods in Statistics*, 89–113.

[25] Saunders, I.W. and Moran, P.A.P. (1978). On the quantiles of the Gamma and F distributions, *Journal of Applied Probability*, **15**(2), 426–432.

[26] Shaked, M. and Shanthikumar, J.G. (2007). *Stochastic Orders*, Springer Science & Business Media.

[27] Torrado, N. (2021). Comparing the reliability of coherent systems with heterogeneous, dependent and distribution-free components, *Quality Technology & Quantitative Management*, **18**(6), 740–770.

[28] Zhao, P. and Balakrishnan, N. (2011). Dispersive ordering of fail-safe systems with heterogeneous exponential components, *Metrika*, **74**(2), 203–210.

[29] Zhao, P. and Balakrishnan, N. (2012). Stochastic comparisons of largest order statistics from multiple-outlier exponential models, *Probability in the Engineering and Informational Sciences*, **26**(2), 159–182.

[30] Zhang, Y.; Cai, X.; Zhao, P. and Wang, H. (2019). Stochastic comparisons of parallel and series systems with heterogeneous resilience-scaled components, *Statistics*, **53**(1), 126–147.

[31] Zhang, Y.; Ding, W. and Zhao, P. (2019). On variability of series and parallel systems with heterogeneous components, *Probability in the Engineering and Informational Sciences*, 1–20.

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.

 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 — *Journal Citation Reports - JCR (Clarivate); DOAJ-Directory of Open Access Journals; 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 email-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, in .pdf format;
- 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, 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.