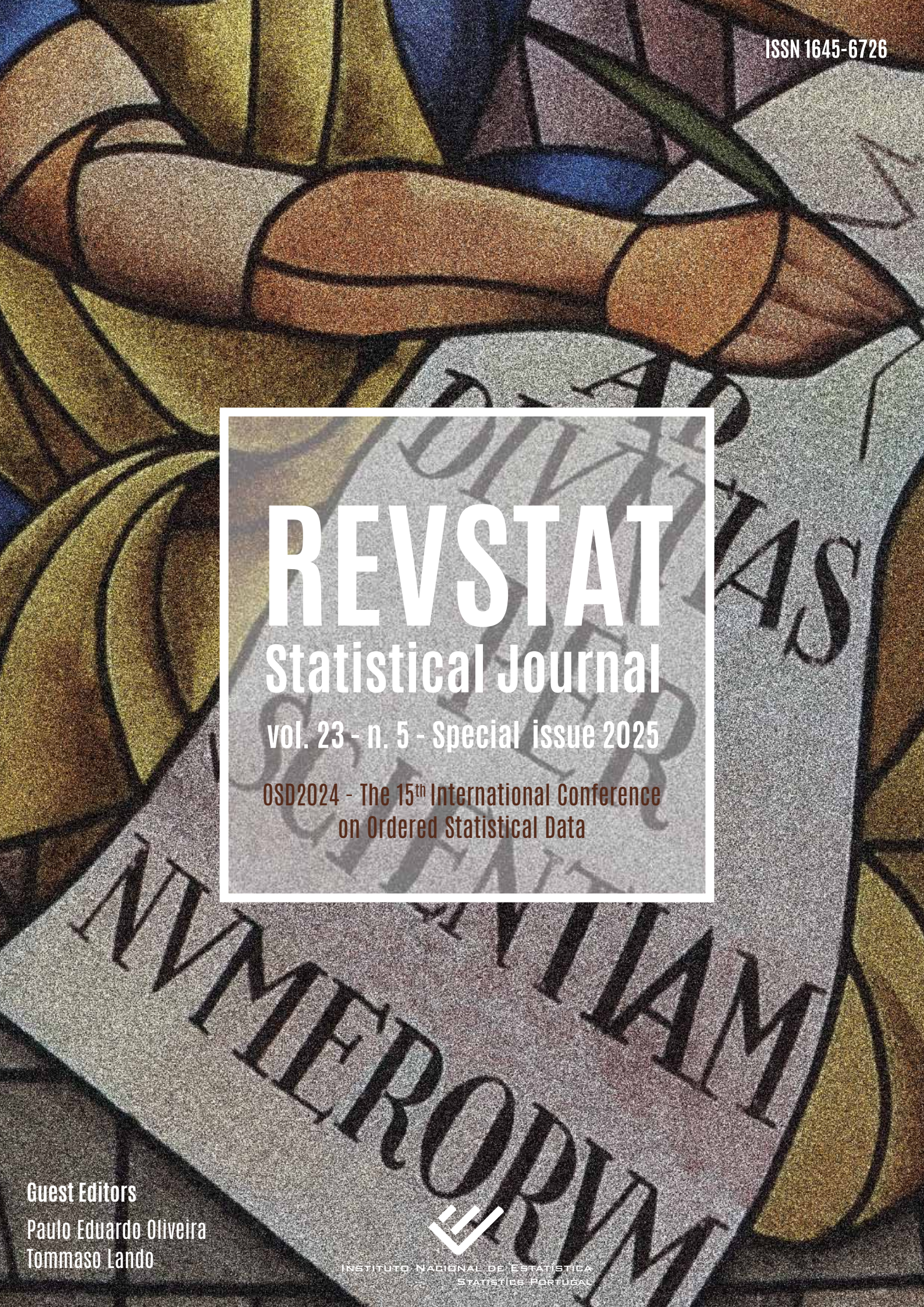
ISSN 1645-6726



REVSTAT

Statistical Journal

vol. 23 - n. 5 - Special issue 2025

OSD2024 - The 15th International Conference on Ordered Statistical Data

Guest Editors

Paulo Eduardo Oliveira

Tommaso Lando

Instituto Nacional de Estatística
Statistics Portugal

REVSTAT — Statistical Journal, vol.23, n. 5 (Special Edition 2025)

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

Continues: Revista de Estatística = ISSN 0873-4275.

ISSN 1645-6726 ; e-ISSN 2183-0371

Editorial Board (2024-2026)

Publisher – Statistics Portugal

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

Digital edition

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

Index

Editorial

The 15th International Conference on Ordered Statistical Data (OSD2024) was held at the Department of Mathematics of the University of Coimbra, Portugal, from June 11-14, 2024, continuing a series of biennial meetings that began in Mysore in 2000 and has since moved around the globe, fostering the sharing of results and new methodologies among an increasing community focused on the mathematical theory and applications of ordered or censored statistical data.

Ordered statistical data appear commonly in various applied problems in statistics, with survival analysis and reliability problems being the most popular ones, while shape-constrained estimation gained recent interest, aiming to incorporate broader geometric properties into approximation and testing procedures. These applications raise quite challenging theoretical questions, unveiling properties, suggesting methods to conveniently approximate most of the unknown probabilistic characterisations, and helping to tackle the subtleties of the real-life procedures and the way they translate into mathematical models.

The OSD2024 programme included plenary lectures on topics in stochastic order problems and their applications, and nonparametric or semiparametric estimation problems. A number of contributed sessions showcased a variety of rather interesting talks ranging from the classical, for this conference series, discussions on stochastic orders, properties of order statistics, and the applications of these topics to coherent systems, to the treatment of complex censoring schemes, dealing with theoretical characterisations and estimation difficulties, with a special interest in nonparametric methods.

The Ordered Statistics Data conference series and, in particular, OSD2024 offer compatibility with a wide range programme while avoiding parallel sessions, thus contributing to broadening the knowledge on statistical theory and methodologies. The OSD2024 also profited from being held at a UNESCO World Heritage site, the University of Coimbra, complementing the scientific programme with cultural and historical visits to the University Palace, offering a glimpse of the student's life throughout time in one of the oldest universities in the world.

This special issue of *REVSTAT – Statistical Journal* includes a number of research articles issued from the talks presented at OSD2024. We are grateful to the authors for submitting their work. More generally, we are very grateful to the invited speakers and contributors who made this conference such an interesting event.

The organization of OSD2024 was only possible with the support of Centre for Mathematics of the University of Coimbra (CMUC) UID/MAT/00324/2020, funded by the Portuguese Government through FCT/MCTES and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020, to which we express our most sincere gratitude. Moreover, we received full support from *REVSTAT – Statistical Journal*, warmly welcoming the idea of a special edition issued from OSD2024, which is now coming to life.

January 2026

Guest Editors

Paulo Eduardo Oliveira

Tommaso Lando

REVSTAT – Statistical Journal
Volume 23, Number 5, Special Edition 2025, 544–568
https://doi.org/10.57805/revstat.v23i5.696

On Progressively Type-II Censored Inverse Lomax Distribution: Characterizations, Estimation and Application to Cancer Data

Authors: Abhimanyu S. Yadav
– Department of Statistics, Banaras Hindu University, Varanasi, India.
asybhu10@gmail.com

Shivanshi Shukla
– Department of Statistics, Central University of Rajasthan, Ajmer, India.
speak2shivanshi@gmail.com

Mahendra Saha
– Department of Statistics, University of Delhi, Delhi, India.
drmahendrasaha1981@gmail.com

Debayan Koley ✉
– Department of Statistics, Banaras Hindu University, Varanasi, India.
devstat96@gmail.com

Neha Jaiswal
– Department of Statistics, Banaras Hindu University, Varanasi, India.
nehajaiss11@gmail.com

Received: November 2023 Revised: December 2024 Accepted: December 2024

Abstract:

- This paper discusses some additional distributional characteristics of inverse Lomax distribution and explores the classical and Bayesian estimation procedures of its survival characteristics. Moreover, the estimation techniques are developed under the progressive type-II censoring scheme. In classical setup, maximum likelihood estimators for the distribution parameters and the considered survival characteristics are obtained. The Bayesian estimation techniques have been discussed with informative priors under a generalized asymmetric loss function. Further, different interval estimators such as asymptotic confidence interval, bootstrap confidence interval, and the Bayes credible interval have also been constructed for the parameters and survival characteristics of the inverse Lomax distribution under progressive type-II censoring. A Monte Carlo simulation has been carried out to evaluate the performances of the point and interval estimators of the considered survival characteristics. To illustrate the practical applicability of the proposed procedures, two data sets pertaining to cancer patients have been explored.

Keywords:

- *Bayes estimation; different characterization; inverse Lomax distribution; survival properties.*

AMS Subject Classification:

- 60E05, 62F10, 62F15.

✉ Corresponding author

1. INTRODUCTION

In survival analysis, the estimation of survival characteristics plays an important role in gaining deeper insights about the observed time-to-event data. Analyzing such data using appropriate probability models are of special interest for several researchers. However, choosing a suitable lifetime model out of thousands is a mammoth task, but can be tackled skillfully if one decides to select the models based on the hazard nature of the data. Apart from the constant hazard rate, there exist monotone and non-monotone hazard rate which can occur naturally in many survival studies. The Weibull, gamma, generalized exponential distribution, etc. are usually the primary choices to analyze any survival data with monotone hazard rate. But in many cancer or tumor recurrence-related studies, the hazard associated with the events of interest can initially increase, and then can gradually decrease after a certain time point (see Boucher and Kerber, 2001, for instance). Such types of hazard pattern is known as hump type or non-monotone hazard rate and arises naturally in many demographic or pharmaceutical studies where the initial failures are very high, reaching a peak and then decline slowly. Child mortality data can be taken as another example. The inverse exponential, inverse Weibull, inverse Gaussian, generalized inverse exponential distributions, etc. can be chosen as suitable models for these types of situations.

The inverse Lomax distribution (ILD) was initially explored by Kleiber and Kotz (2003) in the context of stochastic modelling in economics and actuarial science. Later, it was studied by Rahman et al. (2013) for estimation and prediction purposes through a Bayesian approach. Having both decreasing and non-monotone hazard rates, ILD can be used to model lifetime data that exhibits a non-monotone hazard nature. Classical and the Bayesian estimation procedures for the parameters and reliability functions of ILD using type-II censored observations have been discussed by Singh et al. (2016). Yadav et al. (2016) explored the parametric estimation procedure for ILD parameters under type-I and type-II hybrid censoring setup. Kumar et al. (2021) proposed the transmuted version of ILD and discussed its distributional properties. The probability density function (PDF) and cumulative distribution function (CDF) of ILD are given as

$$f(x;\zeta,\vartheta) = \frac{\zeta\vartheta}{x^2}\left(1+\frac{\vartheta}{x}\right)^{-(1+\zeta)} \quad |_{x,\,\zeta,\,\vartheta\in\mathcal{R}^+} \tag{1.1}$$

and

$$F(x;\zeta,\vartheta) = \left(1+\frac{\vartheta}{x}\right)^{-\zeta} \quad |_{x,\,\zeta,\,\vartheta\in\mathcal{R}^+}, \tag{1.2}$$

respectively. Recently, Yadav et al. (2019) proposed the Bayesian estimation of stress-strength reliability parameter for ILD using observations under a progressive type-II censoring (PCS-II) setup. Analyzing survival data under a progressive censoring strategy has garnered significant attention over a few decades because as opposed to usual type-I and type-II censoring, it offers a convenient flexibility of removing experimental subjects during the experiment. Progressive censoring was properly discussed by Cohen (1963) to deal with more sophisticated life testing experiments. PCS-II can be described as follows; suppose n units are subjected to a certain life test and the experimenter needs to obtain only m failures. At the time of first failure $X_{1:m:n}$, R_1 of the surviving units are randomly removed from the remaining $(n-1)$ units and the experiment continues further. When the second failure $X_{2:m:n}$ is observed, R_2 of the surviving units are randomly removed from the remaining $(n-2-R_1)$ units. The experiment continues in a similar fashion until the m^{th} failure $X_{m:m:n}$ is observed, and all

the remaining R_m surviving units are removed from the experiment. Such a sample of size m is known as progressive type-II censored sample with censoring scheme $(R_1, R_2, ..., R_m)$ provided $n = m + \sum_{i=1}^{m} R_i$. If all $R_i; i = 1, 2, ..., m-1$ are zero and $R_m = n - m$, then the above censoring scheme reduces to conventional type-II censoring scheme and the case of complete sample is obtained by considering all $R_i; i = 1, 2, ..., m$ equal to zero. The theoretical aspects along with the applications of the PCS-II in life testing experiments are discussed by several statisticians(see Cohen, 1963; Balakrishnan and Aggarwala, 2000; Balakrishnan, 2007, for details). Various parametric inferential procedures for different lifetime models under PCS-II are explored by several researchers, viz., Kundu (2008), Pradhan and Kundu (2009), Rastogi and Tripathi (2014), Singh et al. (2015a), Singh et al. (2015b), Dey et al. (2016), Kayal et al. (2017), Maurya et al. (2017) etc. For recent accounts of the estimation of model parameters for different probability distributions under PCS-II, the readers may go through Lodhi et al. (2021), Maurya et al. (2019), Yadav et al. (2019), and Yadav et al. (2022).

The main objective of the present article is twofold. First; to derive some additional survival and distributional properties not considered by Rahman et al. (2013), and second; to address the classical and the Bayes estimation techniques of the survival and hazard function of ILD under PCS-II. Further, the asymptotic confidence interval (ACI) and different bootstrap confidence intervals (BCIs), namely, standard bootstrap (s-boot), percentile bootstrap (p-boot), and student-t bootstrap (t-boot), and the Bayes credible intervals are constructed for the survival and hazard function for the same design of censoring parameters. A Monte Carlo simulation study has been performed to investigate the performances of the obtained estimators and corresponding intervals in terms of mean squared errors (MSEs) and width of the intervals, respectively. To the best of our knowledge, no attempt has been made to estimate the survival characteristics for ILD under PCS-II till now. Hence, the present article has been molded to fill up this research gap.

The sectional bifurcation of the study is as follows; the introduction of the proposed study and model is given in Section 1. Section 2 describes some distributional properties of ILD. The likelihood function under PCS-II and maximum likelihood estimator (MLE) are discussed in Section 3. Section 4 is devoted to the estimation of survival characteristics. Different bootstrap confidence intervals (s-boot, p-boot, t-boot) under PCS-II are obtained in Section 5. Section 6 describes the Bayes estimation procedure along with the Bayes computation under symmetric and asymmetric loss functions using informative gamma prior. Comparative study among the proposed estimators is performed by conducting Monte Carlo simulation in Section 7. The practical application of the study has been illustrated through two medical data sets in Section 8. Finally, the concluding remarks of the paper are presented in Section 9.

2. DIFFERENT CHARACTERIZATIONS OF ILD

In this section, various survival and distributional properties of ILD such as survival and hazard rate, residual life function, moments and inverse moment based characterizations etc have been discussed, and important expressions related to them are derived.

2.1. Survival and hazard functions

The survival and hazard functions are the two vital characteristics of any lifetime distribution. Hazard rate (better known as instantaneous failure rate) is the dynamic speed with which a system or component fails, expressed in failures per unit of time, whereas the survival function gives the probability that an object of interest will survive beyond any specified time. So if the phenomenon under consideration resembles the characteristics of ILD then the survival and hazard of ILD for $t > 0$ are expressed as

$$S(t;\zeta,\vartheta) = 1 - F(t;\zeta,\vartheta) = 1 - \left(1 + \frac{\vartheta}{t}\right)^{-\zeta} \tag{2.1}$$

and

$$h(t;\zeta,\vartheta) = \frac{f(t;\zeta,\vartheta)}{S(t;\zeta,\vartheta)} = \frac{\frac{\zeta\vartheta}{t^2}(1+\frac{\vartheta}{t})^{-(\zeta+1)}}{1-(1+\frac{\vartheta}{t})^{-\zeta}}, \tag{2.2}$$

respectively.

2.2. Aging intensity

Aging is an interesting phenomenon and an inherent property of a system(may be living or non-living) in life tests. It is obtained by using the hazard rate and is a measure of deterioration of a unit (system) over time. It plays a vital role in depicting the failure pattern of an unit. Jiang et al. (2003) defined aging intensity (AI) function as a quantitative measure of aging. Larger value of the AI function indicates increasing tendency of aging of the random variable under consideration. It is interesting to note that for a given failure rate there exists a unique AI function but not vice versa. Using Equation (2.1) and (2.2), the AI function for ILD is

$$L_x(t;\zeta,\vartheta) = \frac{-tf(t;\zeta,\vartheta)}{S(t;\zeta,\vartheta)\log S(t;\zeta,\vartheta)} = \frac{\frac{-(\zeta\vartheta)}{t}(1+\frac{\vartheta}{t})^{-(\zeta+1)}}{1-\left(1+\frac{\vartheta}{t}\right)^{-\zeta}\log\left(1-\left(1+\frac{\vartheta}{t}\right)^{-\zeta}\right)}. \tag{2.3}$$

Singh et al. (2016) demonstrated the shape of hazard rate of ILD to be inverted bathtub shaped. A comparative study illustrates that the monotonicity of hazard rate is not, in general, transmitted to the monotonicity of the AI function (see Nanda et al., 2007).

2.3. Characterization of (reversed) residual lifetime distribution

The residual lifetime $R(t)$ can be interpreted as the additional life beyond which an unit survives under the constraint that it has survived up to time x i.e; $R(t) = X - x|X > x, \quad x \geq 0$. As a dual of this concept, the reverse residual lifetime $\bar{R}(t)$ is defined as the time glided to reach the failure time given that the unit has lifetime less than or equal to x, i.e, $\bar{R}(t) = x - X|X \leq x, \quad x \geq 0$. The survival function of $R(t)$ for ILD is given by

$$S_{R(t)} = P(X - x > t|X > x) = \frac{S(x+t;\zeta,\vartheta)}{S(x;\zeta,\vartheta)} = \frac{1-\left(1+\frac{\vartheta}{t+x}\right)^{-\zeta}}{1-\left(1+\frac{\vartheta}{x}\right)^{-\zeta}} \tag{2.4}$$

and corresponding PDF of $R(t)$ is obtained as

(2.5)
$$f_{R(t)} = \frac{\zeta\vartheta\left[\frac{\left(1+\frac{\vartheta}{t+x}\right)^{-(\zeta+1)}}{(t+x)^2} - \frac{\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}}{(x)^2}\right] - \zeta\vartheta\left(1+\frac{\vartheta}{t+x}\right)^{-(\zeta+1)}\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}\left[\frac{\left(1+\frac{\vartheta}{x}\right)}{(t+x)^2} - \frac{\left(1+\frac{\vartheta}{x+t}\right)}{(x)^2}\right]}{\left[1-\left(1+\frac{\vartheta}{x}\right)^{-\zeta}\right]^2}.$$

The hazard function of $R(t)$ denoted by $h_{R(t)}$ is given as

(2.6)
$$h_{R(t)} = \frac{\zeta\vartheta\left[\frac{\left(1+\frac{\vartheta}{t+x}\right)^{-(\zeta+1)}}{(t+x)^2} - \frac{\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}}{(x)^2}\right] - \zeta\vartheta\left(1+\frac{\vartheta}{t+x}\right)^{-(\zeta+1)}\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}\left[\frac{\left(1+\frac{\vartheta}{x}\right)}{(t+x)^2} - \frac{\left(1+\frac{\vartheta}{x+t}\right)}{(x)^2}\right]}{\left[1-\left(1+\frac{\vartheta}{x}\right)^{-\zeta}\right]\left[1-\left(1+\frac{\vartheta}{x+t}\right)^{-\zeta}\right]}.$$

Similarly, the expressions of survival function, PDF and hazard function for $\bar{R}(t)$, $t \geq 0$ are given as

$$S_{\bar{R}(t)} = P(x - X > t | X \leq x) = \frac{F(x-t;\zeta,\vartheta)}{F(x;\zeta,\vartheta)} = \frac{\left(1+\frac{\vartheta}{x-t}\right)^{-\zeta}}{\left(1+\frac{\vartheta}{x}\right)^{-\zeta}}, \tag{2.7}$$

$$f_{\bar{R}(t)} = \zeta\vartheta\left(1+\frac{\vartheta}{x-t}\right)^{-(\zeta+1)}\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}\left[\frac{\left(1+\frac{\vartheta}{x}\right)}{(x-t)^2} - \frac{\left(1+\frac{\vartheta}{x-t}\right)}{x^2}\right], \tag{2.8}$$

and

$$h_{\bar{R}(t)} = \zeta\vartheta\left(1+\frac{\vartheta}{x-t}\right)^{-1}\left(1+\frac{\vartheta}{x}\right)^{-(2\zeta+1)}\left[\frac{\left(1+\frac{\vartheta}{x}\right)}{(x-t)^2} - \frac{\left(1+\frac{\vartheta}{x-t}\right)}{x^2}\right], \tag{2.9}$$

respectively.

2.4. Moment based characterization

Moments are crucial measures for determining the shape and nature of any distribution. The expression for r^{th} raw moment of ILD is obtained as follows:

$$\mu_r = E(x^r) = \int_0^\infty \zeta\vartheta\, x^{r-2}\left(1+\frac{\vartheta}{x}\right)^{-(1+\zeta)} dx. \tag{2.10}$$

The integral stated above involves the term $1/x^2$ which is divergent, therefore moments of ILD do not exist for $r \geq 1$. To overcome this drawback, truncation can be considered as a viable option. Sometimes truncating the distribution at both ends holds significance when finite mean (first order moment) of the distribution does not exist. Truncation plays a crucial role in understanding the population characteristics when some units are lost or deliberately removed. A truncated distributions are the conditional distributions with a restricted domain, and it is categorized as doubly truncated (restricted at both ends), right truncated (restricted at the right end), and left truncated (restricted at the left end). The PDF of both end truncated ILD in $(a,b); a < b$ is given as

$$f_{TILD}(x;\zeta,\vartheta) = \frac{\vartheta\zeta\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}}{x^2\left[\left(1+\frac{\vartheta}{b}\right)^{-\zeta} - \left(1+\frac{\vartheta}{a}\right)^{-\zeta}\right]} \quad ; a < x < b. \tag{2.11}$$

Hence, the expression for r^{th} moments of TILD can be obtained as follows:

$$E(X^r_{TILD}) = \frac{\vartheta\zeta}{\left[\left(1+\frac{\vartheta}{b}\right)^{-\zeta} - \left(1+\frac{\vartheta}{a}\right)^{-\zeta}\right]} \int_a^b x^{r-2}\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)} dx.$$

However, after careful consideration it can be seen that even after truncating, the finite moments do not exist in closed form.

2.5. Characterization based on inverse moments

Several studies have asserted that for an inverse family of distributions, moments do not exist. The probable reason is that the integral under consideration for the calculation of moments might not be absolutely convergent. Therefore, the need to evaluate inverse moments and subsequently inverse moment generating function arises. The r^{th} inverse moment about origin is calculated as follows:

$$\mu'_{r^{-1}} = E\left(\frac{1}{x^r}\right) = \int_0^\infty \frac{1}{x^r}\frac{\zeta\vartheta}{x^2}\left(1+\frac{\vartheta}{x}\right)^{-(1+\zeta)} dx.$$

after simplification of the above integral, we get

$$\mu'_{r^{-1}} = \frac{\zeta}{\vartheta^r}\,\mathfrak{B}(r+1, \zeta-r), \tag{2.12}$$

where $\mathfrak{B}(\cdot,\cdot)$. For a particular value of $r = 1$, inverse moment converts into the harmonic mean H.

$$H = \frac{1}{(\zeta-1)\vartheta}, \qquad \zeta > 1. \tag{2.13}$$

Inverse moment generating function is derived as follows and denoted by $G_x(t)$. The other generating function can be obtained in a similar fashion.

$$G_x(t) = \sum_0^\infty \frac{t^r}{r!}\,\mu'_{r^{-1}}.$$

which yields

$$G_x(t) = \sum_0^\infty \frac{t^r}{r!}\frac{\zeta}{\vartheta^r}\,\mathfrak{B}(r+1, \zeta-r). \tag{2.14}$$

2.6. Entropy

Renyi entropy is a common measure for quantifying information that provides a generalization to various notions of entropy. It is particularly useful in the field of statistical inference, econometrics, and pattern recognition in computer science. For ILD, the entropy function is given by

$$\eta = \frac{1}{1-\delta_1}\log\int_0^\infty f(x;\zeta,\vartheta)^{\delta_1}\,dx = \frac{1}{1-\delta_1}\log\int_0^\infty \left(\frac{\zeta\vartheta}{x^2}\left(1+\frac{\vartheta}{x}\right)^{-(\zeta+1)}\right)^{\delta_1} dx.$$

After simplifications, the above integral becomes

$$\eta = \frac{1}{1-\delta_1}\log\left(\frac{\zeta^{\delta_1}}{\vartheta^{\delta_1-1}}\,\mathfrak{B}[2\delta_1-1, (\zeta-1)\delta_1+1]\right); \qquad \delta_1 \neq 1. \tag{2.15}$$

2.7. Stochastic ordering

A random variable X is said to be stochastically greater than another random variable Y ($Y \leq_{st} X$) if $F_Y(x) \leq F_X(x)$ for all x. Similar assertions can be firmly stated in

- hazard rate order ($Y \leq_{hr} X$) if $h_Y(x) \leq h_X(x)$ for all x.
- mean residual life order ($Y \leq_{mlr} X$) if $e_Y(x) \leq e_X(x)$ for all x.
- likelihood ratio order ($Y \leq_{lr} X$) if $\frac{f_X(x)}{f_Y(x)}$ is an increasing function of x.

The results due to Shaked and Shanthikumar (1994) are well known for establishing the implications regarding the stochastic ordering of distributions. The inference drawn using the likelihood ratio implies that the same will be obtained if we consider either hazard rate or mean residual life.

Theorem 2.1. *Let X and Y be two independent random variables that follow ILD with shape parameters ζ_1 and ζ_2 and scale parameters ϑ_1 and ϑ_2 respectively. If $\zeta_1 > \zeta_2$ and $\vartheta_1 > \vartheta_2$ then $(Y \leq_{lr} X)$ for all x.*

Proof: For given X $\sim$ ILD(ζ_1, ϑ_1) and Y $\sim$ ILD(ζ_2,ϑ_2), we have

$$\psi = \frac{f(x;\zeta_1,\vartheta_1)}{f(x;\zeta_2,\vartheta_2)}$$

$$\psi = \frac{\zeta_1\vartheta_1}{\zeta_2\vartheta_2}\left(1+\frac{\vartheta_1}{x}\right)^{-(\zeta_1+1)}\left(1+\frac{\vartheta_2}{x}\right)^{\zeta_2+1}.$$

$$\frac{d}{dx}\psi = \left(\frac{\zeta_1\vartheta_1}{\zeta_2\vartheta_2}\right)\frac{\left(1+\frac{\vartheta_1}{x}\right)^{-(\zeta_1+2)}\left(1+\frac{\vartheta_2}{x}\right)^{\zeta_2}}{x^2}\left[\left(1+\frac{\vartheta_2}{x}\right)\vartheta_1(\zeta_1+1)-\left(1+\frac{\vartheta_1}{x}\right)\vartheta_2(\zeta_2+1)\right].$$

the above equation increases in x for all $\zeta_1 > \zeta_2$ and $\vartheta_1 > \vartheta_2$. □

Corollary 2.1. *Let $X \sim ILD(\zeta_1, \vartheta_1)$ and $Y \sim ILD(\zeta_2,\vartheta_2)$. If $\zeta_1 > \zeta_2$ and $\vartheta_1 > \vartheta_2$ then $(Y \leq_{lr} X)$. Hence, $(Y \leq_{hr} X)$ and $(Y \leq_{mrl} X)$ and $(Y \leq_{st} X)$.*

2.8. Characterization based on order statistics

2.8.1. Distribution of extreme order statistics

Let $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$ be the ordered sample taken from ILD with PDF in Equation (1.1) and CDF in Equation (1.2). Then, the PDF of minimum order statistics for the ILD, denoted by $f_{X_{(1)}}(x_{(1)})$, is given as

$$f_{X_{(1)}}(x_{(1)}) = \begin{cases} \frac{n\zeta\vartheta}{x_{(1)}^2}\left(1-\left(1+\frac{\vartheta}{x_{(1)}}\right)^{-\zeta}\right)^{n-1}\left(1+\frac{\vartheta}{x_{(1)}}\right)^{-(\zeta+1)} & x_{(1)} > 0 \\ 0 \quad \text{otherwise.} \end{cases} \tag{2.16}$$

and the PDF of maximum order statistics for ILD, denoted by $f_{X_{(n)}}(x_{(n)})$, is given as

$$f_{X_{(n)}}(x_{(n)}) = \begin{cases} \frac{n\zeta\vartheta}{x_{(n)}^2}(1+\frac{\vartheta}{x_{(n)}})^{-(\zeta n+1)} & x_{(n)} > 0 \\ 0 \quad \text{otherwise.} \end{cases}, \tag{2.17}$$

respectively.

2.8.2. Distribution of range

Range is defined as the difference between maximum order statistics and minimum order statistics, given as

$$r = x_{(n)} - x_{(1)}.$$

Let, $s = x_{(1)}$. Now the joint distribution of (r, s) is given as follows:

$$\mathfrak{h}(r,s) = \frac{n(n-1)\zeta^2\vartheta^2}{(r+s)s}\left[\left(1+\frac{\vartheta}{r+s}\right)^{-\zeta} - \left(1+\frac{\vartheta}{s}\right)^{-\zeta}\right]^{n-2}\left[\left(1+\frac{\vartheta}{r+s}\right)\left(1+\frac{\vartheta}{s}\right)\right]^{-(\zeta+1)}. \tag{2.18}$$

To obtain the distribution of range r, integrate $h(r,s)$ with respect to s and we get following as the PDF of range r and is given as

$$\mathfrak{h}(r) = \zeta^2\vartheta^2 n(n-1)\frac{\sum_{i=0}^{n-2}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}(-1)^{n-2-i+l-j-k+\zeta(i+1-n)}\binom{-(\zeta(i+1)+1)}{j}\binom{-(\zeta(n-i-1)+1)}{k}(r+j)^j\vartheta^j r^{n-2-i-j-k+\zeta(i+1-n)}}{\sum_{l=0}^{\infty}\binom{-\zeta(i+1)}{l}(l-j-k-\zeta(n-i-1)-1)}. \tag{2.19}$$

2.8.3. Distribution of median

Let $x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)}$ be the ordered sample taken from ILD with PDF (1.1) and CDF (1.2). Then median is defined as the value corresponding to that observation that exceeds and is exceeded by the same number of observations. There are two cases to deal with:

First case is when the number of observations (n) is odd i.e., $n = 2m+1$ where m is any positive integer. In this case median is given by $(n+1)/2$ th observation and the PDF for the median is as follows:

$$\mathfrak{h}(x_{med}) = \frac{\zeta\vartheta(2m+1)!}{(m!)^2 x_{med}^2}\left[1+\frac{\vartheta}{x_{med}}\right]^{-\zeta(m+1)+1}\left[1-\left(1+\frac{\vartheta}{x_{med}}\right)^{-\zeta}\right]^m. \tag{2.20}$$

The Second case, when the number of observations (n) is even i.e., $n = 2m$ where m is any positive integer. In this case median is given as the average value of $(n/2)^{th}$ observation and $(n/2)^{th+1}$ observation. For the PDF of median in this case we first evaluate the joint pdf of $u = \left(X_{(n)} + X_{(n+1)}\right)/2$ and $v = X_{(n)}$ which is

$$\mathfrak{h}(u,v) = \frac{2m!\zeta\vartheta}{(m-1)^2!}\frac{1}{u^2(2u-v)^2}\int_0^u\left[\left(1+\frac{\vartheta}{v}\right)^{-\zeta}\left(1-\left(1+\frac{\vartheta}{2u-v}\right)^{-\zeta}\right)\right]^{m-1}\left[\left(1+\frac{\vartheta}{v}\right)\left(1-(1+\frac{\vartheta}{2u-v})^{-}\zeta\right)\right]^{-(\zeta+1)}. \tag{2.21}$$

Now the PDF of median (obtained by integrating with respect to v) is as follows:

$$\mathfrak{h}'(u) = \frac{(2m)!\sum_{r=0}^{m-1}\sum_0^{\infty}\sum_0^{\infty}(-1)^{i+k+j-1}\binom{m-1}{i}\binom{-(2\zeta+1)}{j}\binom{-\zeta i}{k}\vartheta^{-(2\zeta+1+i)}(2u+\vartheta)^{-(\zeta i+k)}u^{j+k+(3+i)\zeta-1}}{\sum_0^{\infty}\binom{-\zeta(i+1)-1}{l}2^{-(\zeta(i+1)+l+1)}}. \tag{2.22}$$

3. LIKELIHOOD FUNCTION AND MLE OF THE PARAMETERS

In this section, the method of maximum likelihood estimation has been discussed to obtain the estimate of the $\mathcal{S}$ for $t > 0$, where $\mathcal{S} = [S(t), h(t)]$. Let us assume that (X, R, m, n) is the progressively censored data observed from ILD, i.e.,

$$(X, R, m, n) \Rightarrow (X_{1:m:n}, R_1), (X_{2:m:n}, R_2), ..., (X_{m:m:n}, R_m)$$

and the likelihood function of (ζ, ϑ) for observed data is written as:

$$\begin{aligned} L(\zeta, \vartheta|\underline{X}) &= \zeta_1 \prod_{i=1}^{m} f(x_i; \zeta, \vartheta)[S(x_i; \zeta, \vartheta)]^{R_i} \\ &= \zeta_1 \zeta^m \vartheta^m \prod_{i=1}^{m} \frac{1}{x_i^2} \left(1 + \frac{\vartheta}{x_i}\right)^{-(\zeta+1)} \left[1 - \left(1 + \frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i} \end{aligned} \tag{3.1}$$

where

$$\zeta_1 = n(n - 1 - R_1)(n - 2 - R_2 - R_1)...(n - m+1 - R_1... - R_{m-1}).$$

The log-likelihood function is

$$\begin{aligned} \ln L(\zeta, \vartheta) &= \ln \zeta + m \ln \zeta + m \ln \vartheta - (1 + \zeta) \sum_{i=1}^{m} \ln \left(1 + \frac{\vartheta}{x_i}\right) - 2 \sum_{i=1}^{m} \ln x_i \\ &+ \sum_{i=1}^{m} R_i \ln \left[1 - \left(1 + \frac{\vartheta}{x_i}\right)^{-\zeta}\right]. \end{aligned} \tag{3.2}$$

The MLEs of the parameters ζ and ϑ are computed by differentiating Equation (3.2) with respect to the parameters and equate it to zero, i.e,

$$\frac{\partial \ln L(\zeta, \vartheta)}{\partial \zeta} = \frac{m}{\zeta} - \sum_{i=1}^{m} \ln \left(1 + \frac{\vartheta}{x_i}\right) + \sum_{i=1}^{m} \frac{R_i \left(1 + \frac{\vartheta}{x_i}\right)^{-\zeta} \ln \left(1 + \frac{\vartheta}{x_i}\right)}{\left[1 - \left(1 + \frac{\vartheta}{x_i}\right)^{-\zeta}\right]} = 0 \tag{3.3}$$

and

$$\frac{\partial \ln L}{\partial \vartheta} = \frac{m}{\vartheta} - (1 + \zeta) \sum_{i=1}^{m} \frac{1}{x_i + \vartheta} + \sum_{i=1}^{m} \frac{\zeta R_i \left(1 + \frac{\vartheta}{x_i}\right)^{-(\zeta+1)}}{x_i \left[1 - \left(1 + \frac{\vartheta}{x_i}\right)^{-\zeta}\right]} = 0. \tag{3.4}$$

From the obtained likelihood equations, given in Equations (3.3) and (3.4), it is observed that the MLEs can not be computed in closed form due to the implicit form of the likelihood equation. Therefore, one may use any iterative procedure like Newton-Raphson (N-R) to obtain the MLEs of the parameter.

4. ESTIMATION OF THE SURVIVAL CHARACTERISTICS

In parametric inferential theory, the estimation of the survival characteristics for the complete sample case is somehow manageable due to the explicit form of the associated

distribution, but the same is not true in case of censored sample. Therefore, the invariance property of MLE may be used to obtain the estimates of the survival characteristics. Once, the MLEs $(\hat{\zeta}, \hat{\vartheta})$ of the parameters (ζ, ϑ) are obtained, the MLEs of the survival and hazard function for any $t > 0$ can be calculated by simply plugging the MLEs of the parameters into the corresponding functions. Hence, the MLEs of $S(t)$, $h(t)$ are given as

$$\hat{S}(t) = 1 - \left(1 + \frac{\hat{\vartheta}}{t}\right)^{-\hat{\zeta}} \tag{4.1}$$

and

$$\hat{h}(t) = \frac{\hat{\zeta}\,\hat{\vartheta}\left(1 + \frac{\hat{\vartheta}}{t}\right)^{-(1+\hat{\zeta})}}{t^2\left[1 - \left(1 + \frac{\hat{\vartheta}}{t}\right)^{-\hat{\zeta}}\right]}, \tag{4.2}$$

respectively.

4.1. Asymptotic distribution of survival characteristics

From the previous section, it is observed that the exact distribution of MLE of the survival characteristics $\mathcal{S} = [S(t), h(t)]$ are not available. Hence, in this section, the asymptotic distribution of $\mathcal{S}$ with the help of asymptotic distribution of (ζ, ϑ) based on asymptotic properties and general condition of MLEs are derived. Generally, for large samples, the asymptotic distribution of the parameters approximately follow a normal distribution. i.e.

$$[\sqrt{n_1}(\hat{\zeta} - \zeta), \sqrt{n_2}(\hat{\vartheta} - \vartheta)] \to N_2\left(0, \frac{1}{I(\hat{\Theta})}\right)$$

where $\Theta = (\zeta, \vartheta)$ and

$$I(\hat{\Theta}) = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix}. \tag{4.3}$$

The matrix $I(\hat{\Theta})$ is called the Fisher information matrix.

$$\phi_{11} = -E\left(\frac{\partial^2 \ln L}{\partial \zeta^2}\right), \phi_{12} = -E\left(\frac{\partial^2 \ln L}{\partial \zeta \partial \vartheta}\right), \phi_{21} = -E\left(\frac{\partial^2 \ln L}{\partial \zeta \partial \vartheta}\right), \phi_{22} = -E\left(\frac{\partial^2 \ln L}{\partial \vartheta^2}\right)$$

where,

$$\begin{aligned}
\frac{\partial^2 \ln L}{\partial \zeta^2} &= \frac{m}{\zeta^2} + \sum_{i=1}^{m} R_i \log\left(1 + \frac{\vartheta}{x_i}\right)^2 \frac{(1 + \frac{\vartheta}{x_i})^{\zeta}}{\left[\left(1 + \frac{\vartheta}{x_i}\right)^{\zeta} - 1\right]^2}, \\
\frac{\partial^2 \ln L}{\partial \vartheta^2} &= \frac{m}{\vartheta^2} + \sum_{i=0}^{m}\left(\frac{\zeta + 1}{(x_i + \vartheta)^2}\right) + \zeta \sum_{i=0}^{m} R_i \left(\frac{(\zeta + 1)(1 + \frac{\vartheta}{x_i})^{\zeta} - 1}{x_i^2\left[(1 + \frac{\vartheta}{x_i})^{\zeta+1} - (1 + \frac{\vartheta}{x_i})\right]^2}\right), \\
\frac{\partial^2 \ln L}{\partial \zeta \partial \vartheta} &= \frac{\partial^2 \ln L}{\partial \vartheta \partial \zeta} \\
&= \sum_{i=1}^{m} \frac{1}{(\vartheta + x_i)} - \sum_{i=1}^{m} \frac{R_i\left[\frac{1}{x_i + \vartheta}\left[\left(1 + \frac{\vartheta}{x_i}\right)^{\zeta} - 1\right] - \zeta \log(1 + \frac{\vartheta}{x_i})(1 + \frac{\vartheta}{x_i})^{\zeta - 1}\frac{1}{x_i}\right]}{(1 + \frac{\vartheta}{x_i})^{\zeta} - 1}.
\end{aligned}$$

We are interested in constructing the confidence interval for $\mathcal{S}$. Since the explicit distribution of $\mathcal{S}$ is not easily obtained, thus the concept of large sample theory has been used to construct an asymptotic confidence interval. For large samples, it can be easily verified that,

$$Z = \frac{\hat{\mathcal{S}} - \mathcal{S}}{\sqrt{Var(\hat{\mathcal{S}})}} \to N(0,1) \; when \; n_i \to \infty; i = 1, 2. \tag{4.4}$$

From Equation (4.4), it is clear that an estimated variance of $\mathcal{S}$ is needed. Thus, the concept of delta method is used to obtain the variance of $\mathcal{S}$.

$$Var(\mathcal{S}) = \frac{1}{n_1\phi_{11}}\left(\frac{\partial \mathcal{S}}{\partial \zeta}\right)^2 + \frac{1}{n_2\phi_{22}}\left(\frac{\partial \mathcal{S}}{\partial \vartheta}\right)^2$$

where,

$$\left(\frac{\partial \mathcal{S}}{\partial \zeta}\right)^2 = \left[\left(\frac{\partial S(t)}{\partial \zeta}\right)^2, \quad \left(\frac{\partial h(t)}{\partial \zeta}\right)^2\right], \qquad \left(\frac{\partial \mathcal{S}}{\partial \vartheta}\right)^2 = \left[\left(\frac{\partial h(t)}{\partial \vartheta}\right)^2, \quad \left(\frac{\partial h(t)}{\partial \vartheta}\right)^2\right]$$

and

$$\frac{\partial S(t)}{\partial \zeta} = \left(1 + \frac{\vartheta}{t}\right)^{-\zeta} \log\left(1 + \tfrac{\vartheta}{t}\right), \qquad \frac{\partial S(t)}{\partial \vartheta} = \frac{\zeta}{t}\left(1 + \tfrac{\vartheta}{t}\right)^{-(\zeta+1)},$$

$$\frac{\partial h(t)}{\partial \zeta} = \frac{\frac{\zeta\vartheta}{t^2}(1 + \frac{\vartheta}{t})^{-(\zeta+1)}}{1 - (1 + \frac{\vartheta}{t})^{-\zeta}}\left[\frac{1}{\zeta} - \log\left(1 + \tfrac{\vartheta}{t}\right) - \frac{\left(1+\frac{\vartheta}{t}\right)^{-\zeta}\log\left(1+\frac{\vartheta}{t}\right)}{1-\left(1+\frac{\vartheta}{t}\right)^{-\zeta}}\right],$$

$$\frac{\partial h(t)}{\partial \vartheta} = \frac{\frac{\zeta\vartheta}{t^2}(1 + \frac{\vartheta}{t})^{-(\zeta+1)}}{1 - (1 + \frac{\vartheta}{t})^{-\zeta}}\left[\frac{1}{\vartheta} - \frac{1+\zeta}{(t+\vartheta)} - \frac{\zeta\left(1+\frac{\vartheta}{t}\right)^{-\zeta-1}}{t\left(1-\left(1+\frac{\vartheta}{t}\right)^{-\zeta}\right)}\right],$$

respectively. Hence, using the asymptotic distribution of $\mathcal{S}$ given in Equation (4.4) the $100(1-\tau)\%$ ACI for $\mathcal{S}$ is obtained as

$$[\hat{\mathcal{S}}_L, \hat{\mathcal{S}}_U] \in [\hat{\mathcal{S}} \mp Z_{\frac{\tau}{2}}\sqrt{Var(\hat{\mathcal{S}})}].$$

5. BOOTSTRAP CONFIDENCE INTERVAL

The number of observations obtained through any life testing experiments are often not large enough, therefore the ACI may not be an appropriate choice. Thus, in this section, an alternative confidence interval construction procedure, known as the bootstrap method, suggested by Efron and Tibshirani (1993), is considered. This method of finding confidence interval is the most efficient sampling and re-sampling procedures without the need of pivotal quantity. Here, we discuss the different types of BCIs, namely $(s-\text{boot})$, $(p-\text{boot})$, and $(t-\text{boot})$. The following steps may be used to construct different 95% BCIs.

- Specify the values of censoring parameters n, m, R_i and model parameters ζ, ϑ.
- Generate $X_{1:m:n}, X_{2:m:n}, ..., X_{m:m:n}$ ordered PCS-II from ILD such that $m \leq n$.
- Compute MLE $(\hat{\zeta}, \hat{\vartheta})$ of the parameters (ζ, ϑ) and obtain $\hat{S}(t), \hat{h}(t)$ of $\mathcal{S}$ using $X_{1:m:n}, ..., X_{m:m:n}$.
- Again generate PCS-II bootstrap samples $X^*_{1:m:n}, X^*_{2:m:n}, ..., X^*_{m:m:n}$ from ILD using $\hat{\Theta} = (\hat{\zeta}, \hat{\vartheta})$ as population parameters and calculate the MLE $\hat{\Theta}^*$.
- Compute MLE $\hat{\mathcal{S}}^*_1 = [\hat{S^*_1}(t), \hat{h^*_1}(t)]$ of $\mathcal{S}$ using $\hat{\Theta}^*$.
- Repeat above steps B times to generate $\hat{\mathcal{S}}^*_i$; $i = 2, ..., B$.

5.1. s-boot

Let $\bar{\mathcal{S}}^*$ and V^* be the sample mean and sample standard deviation of $\hat{S}^*, i = 1, 2, ..., B$.

$$\bar{\mathcal{S}}^* = \frac{1}{B}\sum_{i=1}^{B} \hat{\mathcal{S}}^*_i \quad \textit{and} \quad V^* = \sqrt{\frac{1}{B}\sum_{i=1}^{B}(\hat{\mathcal{S}}^*_i - \bar{\mathcal{S}}^*)^2},$$

respectively. Thus, $100(1-\tau)\%$ s-boot confidence interval for $\mathcal{S}$ is given by

$$\left(\hat{\mathcal{S}}^s_L \qquad \hat{\mathcal{S}}^s_U\right) \quad \in \quad \left(\hat{\mathcal{S}}^* - Z_{\tau/2}\, V^* \qquad \hat{\mathcal{S}}^* + Z_{\tau/2}\, V^*\right).$$

5.2. p-boot

Let $\hat{\mathcal{S}}^{*(\delta)}$ be the δ-percentile of $(\hat{\mathcal{S}}^*_{(i)};\ i = 1, 2, ..., B)$ and $\hat{\mathcal{S}}^{*(\delta)}$ is such that

$$\frac{1}{B}\sum_{i=1}^{B} I(\mathcal{S}^*_{(i)} \leq \hat{\mathcal{S}}^{*(\delta)}) = \delta \quad : 0 \leq \delta \leq 1.$$

where, $I(.)$ is the indicator function. Then $100(1-\tau)\%$ p-boot confidence interval is given by

$$\left(\hat{\mathcal{S}}^p_L \qquad \hat{\mathcal{S}}^p_U\right) \quad \in \quad \left(\hat{\mathcal{S}}^{*[B\frac{\tau}{2}]} \qquad \hat{\mathcal{S}}^{*[B\frac{1-\tau}{2}]}\right).$$

5.3. t-boot

The student's t-bootstrap confidence interval is obtained by the following additional steps:

- Generate again bootstrap sample $X^{**}_{1:m:n}, X^{**}_{2:m:n}, ..., X^{**}_{m:m:n}$ of size $m \leq n$ using $\hat{\Theta}^*$.
- Compute MLE of Θ say $\hat{\Theta}_i{}^{**}$ and obtain $\hat{\mathcal{S}}^{**}_i$, MLE of $\mathcal{S}$, $\forall i = 1, ..., B$.
- Calculate $V^{**} = \sqrt{\frac{1}{B}\sum_{i=1}^{B}(\hat{\mathcal{S}}^{**}_i - \bar{\mathcal{S}}^{**})^2}$ where $\bar{\mathcal{S}}^{**} = \frac{1}{B}\sum_{i=1}^{B}\hat{\mathcal{S}}^{**}_i$.

- Compute the statistic $T = \dfrac{\hat{\mathcal{S}}^{**} - \bar{\mathcal{S}}^{**}}{V^{**}}$. The $100(1-\tau)\%$ t-boot confidence interval for $\mathcal{S}$ is given by

$$\left(\hat{\mathcal{S}}_L^p \qquad \hat{\mathcal{S}}_U^p\right) \quad \in \quad \left(\bar{\mathcal{S}}^{**} - t^{\tau/2}\, V^{**} \qquad \bar{\mathcal{S}}^{**} + t^{\tau/2}\, V^{**}\right).$$

To study the different CIs, we consider their estimated average widths ($\mathcal{W}$) and coverage probability ($\mathcal{P}$). For each of the considered methods, the average width of the BCIs are computed based on B different trials. The average width and coverage probability are given by

$$\mathcal{W} = \frac{\sum_{i=1}^{B}(U_i - L_i)}{B},$$
$$\mathcal{P} = \frac{\#\,(L \le \Theta \le U)}{B},$$

where L and U are the lower and upper limit of the corresponding $100(1-\tau)\%$ CI based on B replicates, respectively.

6. BAYESIAN ESTIMATION

In this section, Bayes estimators for $\mathcal{S}$ are derived using the posterior distribution of ζ and ϑ. It is well known that in the Bayesian paradigm, model parameter(s) are treated as a random variable that follows some standard probability distribution, quantified as the prior distribution. The accuracy of the Bayes estimators is measured by appropriately chosen loss functions. Here, we have chosen gamma prior and general class of asymmetric loss function (ASLF) in our study. The considered prior is more flexible in terms of accommodating variety of shapes of other standard distributions. The prior for ζ and ϑ are given as

$$g_1(\zeta;\mu,\omega) = \frac{\omega^{\mu}}{\Gamma\mu} e^{-\zeta\omega}\zeta^{\mu-1} \tag{6.1a}$$

and

$$g_2(\vartheta;\kappa,\nu) = \frac{\nu^{\kappa}}{\Gamma\kappa} e^{-\vartheta\nu}\vartheta^{\kappa-1}, \tag{6.1b}$$

respectively, where, μ, ω, κ & ν are the hyper-parameters which are assumed to be known and positive. Since, the considered priors are independent in nature, the joint posterior given data using the Equations (3.1) and (6.1a-6.1b) is obtained as follows:

$$P(\zeta,\vartheta|data) = \Lambda\, e^{-(\omega\zeta+\nu\vartheta)}\zeta^{\mu+m-1}\vartheta^{\omega+m-1}\prod_{i=1}^{m}\left(1+\frac{\vartheta}{x_i}\right)^{-(\zeta+1)}\left[1-\left(1+\frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i}, \tag{6.2}$$

where Λ is the normalizing constant, given as

$$\Lambda = \int_0^{\infty}\int_0^{\infty} e^{-(\omega\zeta+\nu\vartheta)}\zeta^{\mu+m-1}\vartheta^{\kappa+m-1}\prod_{i=1}^{m}\left(1+\frac{\vartheta}{x_i}\right)^{-(\zeta+1)}\left[1-\left(1+\frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i} d\zeta\, d\vartheta.$$

The general form of the generalized class of ASLF was proposed by Calabria and Pulcini (1996) as the modified version of the Linex loss function. The general mathematical form of generalized ASLF is given as

$$L_{as}(\delta,\hat{\delta}) \propto \left(\frac{\hat{\delta}}{\delta}\right)^{\varepsilon} - \varepsilon\,\ln\left(\frac{\hat{\delta}}{\delta}\right) - 1.$$

where ε is the loss parameter that reflects the departure from symmetry. The Bayes estimates of the parameter δ concerning the considered loss function is obtained by using the following expression:

$$\hat{\delta}_{as} = \left[E(\delta^{-\varepsilon})\right]^{-\frac{1}{\varepsilon}}, \tag{6.3}$$

provided the expectation exists and is finite. It is to be noted that for different values of ε as -2, -1, 1, 2 the ASLF is reduced to precautionary loss function (PLF), squared error loss function (SELF), entropy loss function (ELF) and general entropy loss function (GELF), respectively. Hence, using the Equation (6.3), the Bayes estimator of $\mathcal{S}$ under ASLF are obtained as

(6.4)
$$\hat{S}(t)_{as} = \left[\Lambda \int_0^\infty \int_0^\infty \left[1-\left(1+\frac{\vartheta}{t}\right)^{-\zeta}\right]^{-\varepsilon} e^{-(\omega\zeta+\nu\vartheta)}\zeta^{\mu+m-1}\vartheta^{\kappa+m-1}\prod_{i=1}^{m}\left(1+\frac{\vartheta}{x_i}\right)^{-(\zeta+1)}\left[1-\left(1+\frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i} d\zeta\, d\vartheta\right]^{-1/\varepsilon},$$

(6.5)
$$\hat{h}(t)_{as} = \left[\Lambda \int_0^\infty \int_0^\infty \left[\frac{\zeta\vartheta}{t^2+\vartheta t}\right]^{-\varepsilon} e^{-(\omega\zeta+\nu\vartheta)}\zeta^{\mu+m-1}\vartheta^{\kappa+m-1}\prod_{i=1}^{m}\left(1+\frac{\vartheta}{x_i}\right)^{-(\zeta+1)}\left[1-\left(1+\frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i} d\zeta\, d\vartheta\right]^{-1/\varepsilon}.$$

Since all the Bayesian estimates obtained above involve a ratio of two integrals, they are not easy to compute analytically. Therefore, resorting to approximation techniques or sample generation techniques from high dimensional posterior setting is an immediate option. Here, we used the Markov Chain Monte Carlo method to obtain the approximated Bayes estimates of the survival characteristics using the generated posterior sample.

6.1. Markov Chain Monte Carlo method

Markov Chain Monte Carlo (MCMC) method is one of the most appropriate and efficient computational procedures to approximate the Equations (6.4-6.5) based on generated sequences of ζ and ϑ. MCMC comprises various Bayes computational techniques. Among all of the accept and reject MCMC methods, the M-H algorithm suggested by Hastings (1970) is used extensively by the researchers. This method efficiently works when the marginal posterior distribution does not assume any standard form. The detailed description of this method may be found in Smith and Roberts (1993), Gelfand and Smith (1990), Upadhyay et al. (2001), Yadav et al. (2022). The marginal posterior distribution of ζ and ϑ are given as

$$\pi_\zeta(\zeta,|data,\vartheta) \propto e^{-\omega\zeta}\,\zeta^{\mu+m-1}\prod_{i=1}^{m}\left(1+\frac{\vartheta}{x_i}\right)^{-(\zeta+1)}\left[1-\left(1+\frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i} \tag{6.6}$$

and

$$\pi_\vartheta(\vartheta,|data,\zeta) \propto e^{-\nu\vartheta}\,\vartheta^{\kappa+m-1}\prod_{i=1}^{m}\left(1+\frac{\vartheta}{x_i}\right)^{-(\zeta+1)}\left[1-\left(1+\frac{\vartheta}{x_i}\right)^{-\zeta}\right]^{R_i}, \tag{6.7}$$

respectively. The following steps are used to extract the posterior sample from the above marginal densities, given in the Equation (6.6-6.7).

- start with initial values (ζ^0, ϑ^0).

- generate N sequences of ζ, ϑ as $\zeta_1, \zeta_2, ..., \zeta_N$, $\vartheta_1, \vartheta_2, ..., \vartheta_N$ using normal distribution as proposal density.
- generate the sequence for $\mathcal{S}$ for any mission time $t > 0$ as

$$S_1, S_2, ..., S_N \quad \text{and} \quad h_1, h_2, ..., h_N.$$

- The Bayes estimates of $\mathcal{S}$ under ASLF are obtained as

$$\hat{S}(t)_{as} = \left(\frac{1}{N - N_b} \sum_{j=1}^{N-N_b} S_j^{-\varepsilon} \right)^{-1/\varepsilon} \qquad \hat{h}(t)_{as} = \left(\frac{1}{N - N_b} \sum_{j=1}^{N-N_b} h_j^{-\varepsilon} \right)^{-1/\varepsilon},$$

Where, N_b denotes the burn-in period. The Bayes credible/highest posterior density interval (HPDI) for survival characteristics may be constructed by employing the algorithm by Chen and Shao (1999).

7. MONTE CARLO SIMULATIONS

From the previous section, it is clear that the exact expressions of MLE and Bayes estimators are not available in explicit form, thus the theoretical comparison among the obtained estimators is not possible. Hence, in this section, Monte Carlo simulations have been performed to compare the performances of the proposed estimators. The simulation study has been conducted for different variations of the censoring parameters (n, m, R_i) [see, Table 1] along with the parameters' variation as $[(0.5, 0.75), (0.85, 1), (1.5, 0.85)\ \&\ (2, 2)]$. The

Table 1: Censoring schemes for simulation.

n, m	Schemes	Place of removals R_i
20, 10	$R_1 : (20, 10, 10, 0^9)$	Removals occurs at first stage
	$R_2 : (20, 10, 0^9, 10)$	Removals occurs at last stage
	$R_3 : (20, 10, 1^{10})$	Removals occurs at each stage
	$R_4 : (20, 10, 5, 0^8, 5)$	Removals occurs at first and last stages
	$R_5 : (20, 10, 0^4, 5^2, 0^4)$	Removals occurs at some intermediate stages
40, 20	$R_6 : (40, 20, 5^4, 0^{16})$	Removals occurs at some beginning stages
	$R_7 : (40, 20, 0^{16}, 5^4)$	Removals occurs at some last stage
	$R_8 : (40, 20, 1^{20})$	Removals occurs at each stage
	$R_9 : (40, 20, 2^5, 0^{10}, 2^5)$	Removals occurs at first and last stages
	$R_{10} : (40, 20, 0^8, 5^4, 0^8)$	Removals occurs at some intermediate stages

sequence of PCS-II samples is generated using the algorithm suggested by Balakrishnan and Sandhu (1995). The value of effective sample size m is chosen such that the sample information is 50% censored. The MLEs of the parameters are calculated using the N-R method and invariance property has been used to construct MLEs of survival characteristics for known mission time $t = 4$. The point estimates of the $\mathcal{S}$ are compared in terms of average mean square (MSE) and interval estimates are compared in terms of the average width of the interval. Bayes estimators are obtained with gamma informative priors under ASLF. The values of loss parameter ε are chosen as $(-2, -1, 1, 2)$. A positive value of ε indicates that over-estimation is more serious than under-estimation and vice versa. The Bayes estimators for the considered values of loss parameter correspond to the one obtained under PLF, SELF,

ELF, and GELF respectively. Since the Bayes estimator assumes the ratio of two integrals, hence MCMC technique has been used to obtain the Bayes point and interval estimates. The values of the hyper-parameters are obtained in such a way that the prior mean is informative (less prior variance). The ACIs of $\mathcal{S}$ are constructed by following the idea of the delta method. Different parametric BCIs (s-boot, p-boot, t-boot) are computed. HPDIs are also computed for the same setup and compared with ACIs and BCIs. All simulation procedures has been performed using *R*-software and results are reported based on 3000 replications [see Tables 2-6]. In the reported tables, $\hat{S}(t)_M$, $\hat{S}(t)_P$, $\hat{S}(t)_S$, $\hat{S}(t)_E$, $\hat{S}(t)_G$ and $\hat{h}(t)_M$, $\hat{h}(t)_P$, $\hat{h}(t)_S$, $\hat{h}(t)_E$, $\hat{h}(t)_G$ denote the MLEs and Bayes estimates of $\mathcal{S}$ and R_M, R_P, R_S, R_E, R_G denotes the corresponding MSEs of the same obtained under PLF, SELF, ELF and GELF respectively. The following conclusions are made based on Monte Carlo simulations.

- The MSEs of the MLEs and Bayes estimators are decreasing by increasing sample sizes and effective sample sizes.
- For $S(t)$, the Bayes estimators under ELF are less penalized as compared to the MLE and other Bayes estimators.
- For $h(t)$, the Bayes estimators under PLF outperform MLE and other Bayes estimators. However, for parametric combination $\zeta = 2$ and $\vartheta = 2$ for some censoring schemes Bayes estimator under SELF and PLF perform equally well.
- The width of the BCIs is lesser than the ACIs for $S(t)$ as well as for $h(t)$. Although boot-p has the smallest width among BCIs for $S(t)$, some variation is reported when $\zeta = 2$ and $\vartheta = 2$ where boot-t bags the position.
- Similarly boot-t has smallest width among BCIs for $h(t)$ but for $\zeta = 2$ and $\vartheta = 2$ for some censoring schemes boot-p hails.
- The width of HPDIs is smaller than that of ACIs and BCIs for both $S(t)$ and $h(t)$.

8. REAL DATA APPLICATIONS

This section explored the application of the proposed procedures through two medical data sets. The data sets are taken from Efron (1988) and Collett (2023). Data set-I represents the survival times of 58 head and neck cancer patients treated with radiotherapy and data set-II represents the survival times of 38 prostate cancer patients. The fitting of the ILD model to the considered data sets is appropriately explained by Yadav et al. (2019) in the stress-strength reliability estimation study. They have fitted ILD for these two data sets and also shown the model compatibility among the most popular family of inverted distributions such as inverted exponential distribution, generalized inverted exponential distribution, inverse Weibull distribution, and remarked that ILD might be an alternative choice to explain the real phenomenon with hump type hazard rate. Hence, here we have taken the same data sets to estimate the survival characteristics with different censoring schemes, given in Table 7.

In Table 7, the quantity $a_1^{b_1}$ represents that the number a_1 is repeated b_1 times. For the considered data sets, MLEs and Bayes estimates of the survival characteristics are obtained using censoring schemes reported in Table 7 for arbitrarily chosen mission time $t = 50, 200$ for data set-I and $t = 50, 65$ for data set-II. The boxplot summary of the generated PCS-II samples is given in Figure 1 & 2 respectively. The Bayes estimates are obtained with non-informative prior $(\mu, \omega, \kappa, \nu \to 0.0001)$ under different loss functions using the MCMC technique. To implement MCMC technique in the real data set, the stationarity of Markov

Table 2: Average estimates and MSEs of the estimator $S(t)|_{t=4}$.

Schemes	Parameter	$\hat{S}(t)_M$	$\hat{S}(t)_P$	$\hat{S}(t)_S$	$\hat{S}(t)_E$	$\hat{S}(t)_G$	R_M	R_P	R_S	R_E	R_G
R_1	0.5, 0.75	0.14646	0.13854	0.13050	0.11354	0.10440	0.00565	0.00310	0.00287	0.00282	0.00305
R_2		0.13810	0.13605	0.13060	0.11942	0.11358	0.00331	0.00232	0.00220	0.00215	0.00224
R_3		0.13968	0.13657	0.13040	0.11770	0.11105	0.00395	0.00260	0.00246	0.00242	0.00253
R_4		0.13707	0.13369	0.12759	0.11490	0.10815	0.00410	0.00273	0.00262	0.00262	0.00276
R_5		0.13774	0.13378	0.12750	0.11468	0.10806	0.00443	0.00271	0.00259	0.00260	0.00274
R_6		0.13749	0.13667	0.13233	0.12333	0.11861	0.00282	0.00214	0.00204	0.00197	0.00200
R_7		0.13382	0.13455	0.13155	0.12546	0.12232	0.00168	0.00143	0.00139	0.00137	0.00139
R_8		0.13629	0.13609	0.13265	0.12561	0.12196	0.00226	0.00175	0.00169	0.00164	0.00165
R_9		0.13716	0.13713	0.13374	0.12679	0.12318	0.00214	0.00172	0.00166	0.00160	0.00161
R_{10}		0.13608	0.13610	0.13250	0.12516	0.12139	0.00243	0.00190	0.00182	0.00175	0.00175
R_1	0.85, 1.0	0.17978	0.16737	0.15896	0.14114	0.13144	0.00848	0.00411	0.00402	0.00430	0.00474
R_2		0.17339	0.16918	0.16322	0.15091	0.14443	0.00481	0.00302	0.00296	0.00307	0.00325
R_3		0.17522	0.16884	0.16219	0.14839	0.14108	0.00564	0.00329	0.00322	0.00337	0.00361
R_4		0.17370	0.16797	0.16127	0.14725	0.13972	0.00534	0.00322	0.00316	0.00335	0.00361
R_5		0.17012	0.16398	0.15730	0.14354	0.13635	0.00599	0.00335	0.00333	0.00359	0.00388
R_6		0.16965	0.16761	0.16291	0.15315	0.14801	0.00407	0.00273	0.00270	0.00277	0.00288
R_7		0.17250	0.17186	0.16848	0.16157	0.15800	0.00267	0.00199	0.00195	0.00194	0.00197
R_8		0.16804	0.16711	0.16337	0.15567	0.15168	0.00306	0.00223	0.00221	0.00227	0.00235
R_9		0.17348	0.17199	0.16818	0.16033	0.15625	0.00299	0.00215	0.00210	0.00210	0.00215
R_{10}		0.16869	0.16797	0.16407	0.15609	0.15196	0.00332	0.00237	0.00234	0.00238	0.00246
R_1	1.5, 0.85	0.25891	0.24748	0.23779	0.21736	0.20637	0.01033	0.00582	0.00577	0.00624	0.00682
R_2		0.25491	0.25008	0.24346	0.22973	0.22251	0.00598	0.00408	0.00402	0.00418	0.00440
R_3		0.25781	0.25156	0.24416	0.22878	0.22068	0.00752	0.00486	0.00476	0.00490	0.00515
R_4		0.26042	0.25358	0.24591	0.22986	0.22132	0.00762	0.00504	0.00491	0.00501	0.00526
R_5		0.25184	0.24650	0.23915	0.22398	0.21605	0.00685	0.00437	0.00435	0.00464	0.00497
R_6		0.25056	0.25027	0.24523	0.23486	0.22948	0.00512	0.00373	0.00367	0.00371	0.00381
R_7		0.25196	0.25217	0.24860	0.24132	0.23758	0.00315	0.00262	0.00258	0.00258	0.00262
R_8		0.25109	0.25125	0.24730	0.23923	0.23507	0.00401	0.00316	0.00311	0.00311	0.00316
R_9		0.25204	0.25177	0.24772	0.23943	0.23514	0.00381	0.00299	0.00295	0.00295	0.00300
R_{10}		0.24844	0.24912	0.24513	0.23701	0.23285	0.00397	0.00308	0.00305	0.00308	0.00315
R_1	2,2	0.55021	0.52093	0.51302	0.49598	0.48668	0.01354	0.00874	0.00948	0.01148	0.01282
R_2		0.55507	0.53545	0.52979	0.51779	0.51136	0.00962	0.00645	0.00679	0.00771	0.00832
R_3		0.55617	0.53402	0.52787	0.51479	0.50777	0.01049	0.00652	0.00691	0.00799	0.00871
R_4		0.55671	0.53304	0.52657	0.51275	0.50529	0.01146	0.00730	0.00773	0.00892	0.00972
R_5		0.55302	0.53271	0.52666	0.51380	0.50692	0.01011	0.00661	0.00700	0.00808	0.00878
R_6		0.55058	0.53913	0.53490	0.52610	0.52150	0.00750	0.00459	0.00479	0.00534	0.00568
R_7		0.55396	0.54659	0.54351	0.53713	0.53382	0.00516	0.00363	0.00373	0.00398	0.00415
R_8		0.54650	0.54061	0.53719	0.53011	0.52643	0.00598	0.00418	0.00433	0.00471	0.00495
R_9		0.55276	0.54358	0.54009	0.53287	0.52911	0.00603	0.00405	0.00418	0.00453	0.00475
R_{10}		0.55091	0.54381	0.54043	0.53345	0.52983	0.00580	0.00386	0.00398	0.00432	0.00453

Chain has been investigated via tuning of the variance, trace plot, auto-correlation function plot, etc. About 500 generated posterior samples from full conditional densities in the very beginning of the process (burn-in period, N_b) are excluded from the total generated sequence of samples and it is observed that the generated sequence of posterior samples is well mixed and follows the Markov property. The MCMC convergence and estimated density plots of the survival and hazard function based on generated MCMC samples for $t = 50$ are given in Figure 3 & 4, respectively. Different interval estimates (ACIs, BCIs, and HPDIs) for both survival characteristics have been obtained for the same set up of design and reported in Table 6. Table 6 indicates that the width of the BCIs is less as compared to the ACIs for all set up of designs. However, the width of HPD intervals are better throughout the analysis than these two in terms of the width of the interval. Further, it is noticed that for $S(t)$, t-boot has lesser width among all BCIs for data set-I and p-boot for data set-II. Similarly for $h(t)$, p-boot is preferable for the first data set whereas t-boot is for the second set.

Table 3: Average estimates and MSEs of the estimators $h(t)|_{t=4}$.

Schemes	Parameter	$\hat{h}(t)_M$	$\hat{h}(t)_P$	$\hat{h}(t)_S$	$\hat{h}(t)_E$	$\hat{h}(t)_G$	R_M	R_P	R_S	R_E	R_G
R_1		0.20999	0.21500	0.21454	0.21358	0.21307	0.00063	0.00027	0.00028	0.00031	0.00032
R_2		0.21330	0.21545	0.21515	0.21451	0.21419	0.00040	0.00023	0.00024	0.00025	0.00026
R_3		0.21282	0.21557	0.21523	0.21450	0.21413	0.00044	0.00024	0.00025	0.00027	0.00028
R_4		0.21309	0.21589	0.21556	0.21486	0.21449	0.00047	0.00026	0.00027	0.00029	0.00030
R_5	0.5, 0.75	0.21309	0.21628	0.21592	0.21517	0.21477	0.00053	0.00026	0.00026	0.00028	0.00029
R_6		0.21308	0.21466	0.21442	0.21391	0.21365	0.00032	0.00022	0.00023	0.00024	0.00024
R_7		0.21456	0.21513	0.21497	0.21463	0.21446	0.00021	0.00016	0.00017	0.00017	0.00018
R_8		0.21323	0.21431	0.21411	0.21370	0.21350	0.00028	0.00020	0.00020	0.00021	0.00022
R_9		0.21332	0.21428	0.21409	0.21371	0.21351	0.00026	0.00019	0.00020	0.00020	0.00021
R_{10}		0.21355	0.21462	0.21441	0.21397	0.21375	0.00029	0.00021	0.00021	0.00022	0.00023
R_1		0.20135	0.20780	0.20725	0.20609	0.20548	0.00088	0.00034	0.00035	0.00038	0.00039
R_2		0.20428	0.20724	0.20684	0.20602	0.20559	0.00054	0.00028	0.00029	0.00030	0.00031
R_3		0.20338	0.20732	0.20688	0.20595	0.20547	0.00062	0.00030	0.00031	0.00032	0.00033
R_4		0.20401	0.20768	0.20724	0.20634	0.20588	0.00060	0.00030	0.00030	0.00032	0.00033
R_5	0.85, 1.0	0.20515	0.20909	0.20865	0.20772	0.20724	0.00070	0.00033	0.00033	0.00035	0.00036
R_6		0.20484	0.20693	0.20662	0.20596	0.20563	0.00046	0.00028	0.00029	0.00030	0.00031
R_7		0.20434	0.20550	0.20526	0.20478	0.20453	0.00033	0.00023	0.00024	0.00024	0.00025
R_8		0.20508	0.20650	0.20624	0.20572	0.20545	0.00037	0.00025	0.00026	0.00026	0.00027
R_9		0.20352	0.20512	0.20486	0.20432	0.20404	0.00036	0.00024	0.00024	0.00025	0.00026
R_{10}		0.20542	0.20683	0.20656	0.20600	0.20571	0.00039	0.00026	0.00026	0.00027	0.00028
R_1		0.19262	0.19784	0.19712	0.19559	0.19477	0.00086	0.00035	0.00037	0.00040	0.00043
R_2		0.19478	0.19704	0.19654	0.19551	0.19497	0.00054	0.00028	0.00029	0.00031	0.00033
R_3		0.19380	0.19678	0.19621	0.19502	0.19440	0.00065	0.00032	0.00033	0.00036	0.00038
R_4		0.19371	0.19672	0.19615	0.19494	0.19431	0.00063	0.00031	0.00032	0.00035	0.00037
R_5	1.5, 0.85	0.19519	0.19798	0.19742	0.19626	0.19565	0.00061	0.00029	0.00030	0.00033	0.00034
R_6		0.19659	0.19753	0.19715	0.19636	0.19596	0.00044	0.00028	0.00028	0.00030	0.00031
R_7		0.19616	0.19653	0.19625	0.19570	0.19542	0.00028	0.00020	0.00021	0.00021	0.00022
R_8		0.19660	0.19708	0.19678	0.19616	0.19584	0.00037	0.00025	0.00025	0.00026	0.00027
R_9		0.19603	0.19670	0.19639	0.19576	0.19544	0.00035	0.00024	0.00024	0.00025	0.00026
R_{10}		0.19713	0.19752	0.19721	0.19658	0.19626	0.00037	0.00024	0.00025	0.00026	0.00026
R_1		0.13215	0.14127	0.13976	0.13652	0.13474	0.00104	0.00038	0.00037	0.00037	0.00038
R_2		0.13336	0.13834	0.13720	0.13480	0.13354	0.00077	0.00031	0.00031	0.00032	0.00033
R_3		0.13236	0.13862	0.13739	0.13478	0.13339	0.00087	0.00032	0.00031	0.00032	0.00033
R_4		0.13198	0.13878	0.13749	0.13475	0.13329	0.00090	0.00035	0.00034	0.00036	0.00037
R_5	2,2	0.13333	0.13898	0.13777	0.13520	0.13383	0.00083	0.00031	0.00031	0.00031	0.00032
R_6		0.13454	0.13765	0.13674	0.13484	0.13384	0.00072	0.00027	0.00027	0.00027	0.00028
R_7		0.13501	0.13630	0.13561	0.13421	0.13349	0.00051	0.00024	0.00024	0.00024	0.00024
R_8		0.13720	0.13800	0.13726	0.13573	0.13495	0.00057	0.00026	0.00026	0.00026	0.00026
R_9		0.13442	0.13661	0.13584	0.13426	0.13345	0.00058	0.00025	0.00025	0.00026	0.00026
R_{10}		0.13539	0.13698	0.13623	0.13468	0.13388	0.00056	0.00024	0.00024	0.00024	0.00025

9. CONCLUDING REMARKS

In this paper, some distributional properties including some survival characteristics of ILD have been studied. ILD can be used as a suitable lifetime model in several cancer or tumor-related studies where hazard initially increases and decreases after a certain point. Analyzing such data under PCS-II gives an edge to practitioners where subjects often leave the study during the experiment for several reasons. Further, MLE and Bayes estimators for $\mathcal{S}$ are discussed for specified t when the observed sample information is obtained under PCS-II. The MLEs for the same are computed by using the invariance property. The Bayes estimators are derived with two independent gamma prior under ASLF, and computed via MCMC method also. Different confidence interval estimation procedures namely ACIs, BCIs (boot-s, boot-p, and boot-t) & HPDIs are computed numerically for $\mathcal{S}$. ACIs for the same are computed by using the concept of delta method. A Monte Carlo simulation study has been performed to investigate the performances of the proposed estimators for the different censoring parameters. From this extensive simulation study, it is noted that Bayes estimators

Table 4: Average width ($\mathcal{W}$) and coverage probability ($\mathcal{P}$) of ACIs and HPDIs for $\mathcal{S}$.

Schemes	Parameter	ACI				HPDI			
		S(t)		h(t)		S(t)		h(t)	
		$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$
R_1	0.5, 0.75	0.59439	0.982	0.13671	0.961	0.17478	0.838	0.04704	0.879
R_2		0.49384	0.988	0.11239	0.967	0.14513	0.825	0.03917	0.842
R_3		0.53742	0.992	0.12221	0.972	0.15341	0.919	0.04138	0.845
R_4		0.51180	0.993	0.11558	0.970	0.15156	0.889	0.04077	0.871
R_5		0.55872	0.991	0.12837	0.963	0.15228	0.895	0.04149	0.826
R_6		0.38232	0.990	0.09041	0.965	0.13115	0.896	0.03574	0.913
R_7		0.33347	0.978	0.07757	0.970	0.10915	0.879	0.02996	0.922
R_8		0.34861	0.983	0.08254	0.960	0.11726	0.861	0.03243	0.902
R_9		0.34067	0.985	0.07997	0.967	0.11713	0.897	0.03195	0.723
R_{10}		0.36988	0.993	0.08759	0.959	0.11939	0.908	0.03315	0.814
R_1	0.85, 1.0	0.68697	0.980	0.15588	0.949	0.19845	0.893	0.05225	0.849
R_2		0.61617	0.987	0.13673	0.973	0.17033	0.886	0.04513	0.856
R_3		0.62798	0.991	0.14202	0.969	0.17880	0.911	0.04740	0.851
R_4		0.61815	0.994	0.13753	0.970	0.17961	0.914	0.04707	0.841
R_5		0.68605	0.993	0.15267	0.963	0.17562	0.898	0.04663	0.905
R_6		0.45690	0.987	0.10584	0.952	0.15196	0.916	0.04064	0.892
R_7		0.41171	0.988	0.09450	0.973	0.13162	0.931	0.03560	0.897
R_8		0.41675	0.994	0.09675	0.960	0.13625	0.902	0.03699	0.903
R_9		0.42230	0.979	0.09774	0.968	0.13959	0.889	0.03773	0.958
R_{10}		0.44285	0.992	0.10264	0.965	0.13870	0.879	0.03771	0.918
R_1	1.5, 0.85	1.23244	0.973	0.23652	0.978	0.26140	0.907	0.05967	0.941
R_2		1.09173	0.940	0.20572	0.977	0.22025	0.881	0.05049	0.893
R_3		1.18641	0.969	0.22391	0.977	0.23249	0.871	0.05340	0.913
R_4		1.16277	0.943	0.21935	0.979	0.23800	0.881	0.05404	0.913
R_5		1.17021	0.933	0.22264	0.981	0.22884	0.865	0.05280	0.882
R_6		0.85517	0.923	0.16247	0.983	0.19315	0.857	0.04438	0.873
R_7		0.72049	0.972	0.13767	0.985	0.16438	0.868	0.03819	0.907
R_8		0.76727	0.952	0.14556	0.987	0.17222	0.853	0.03978	0.926
R_9		0.74240	0.949	0.14193	0.989	0.17471	0.860	0.04030	0.888
R_{10}		0.79104	0.982	0.15036	0.991	0.17188	0.871	0.04006	0.948
R_1	2,2	2.20791	0.972	0.47009	0.993	0.34879	0.909	0.07879	0.934
R_2		1.98079	0.961	0.42167	0.996	0.30033	0.914	0.06798	0.924
R_3		2.05598	0.951	0.43625	0.988	0.31230	0.924	0.07082	0.933
R_4		2.03495	0.941	0.43308	0.990	0.32008	0.915	0.07240	0.919
R_5		2.10546	0.931	0.44696	0.989	0.30987	0.928	0.07038	0.930
R_6		1.57591	0.921	0.33118	0.968	0.26156	0.909	0.06095	0.915
R_7		1.38112	0.940	0.29048	0.943	0.22560	0.934	0.05259	0.892
R_8		1.47554	0.980	0.30985	0.973	0.23614	0.928	0.05498	0.904
R_9		1.38589	0.949	0.29213	0.949	0.23907	0.927	0.05571	0.917
R_{10}		1.44112	0.980	0.30372	0.979	0.23528	0.927	0.05521	0.908

for the designed set up of simulation provides more accurate and precise results in terms of MSE as compared to the classical (MLE) estimators. Subsequently, the HPDIs show a shorter length of the interval in comparison to ACIs and BCIs. However, the interval obtained through the bootstrap method is better in terms of the length of the interval than

Table 5: Average width and corresponding coverage probability of different BCIs for the $\mathcal{S}$.

Schemes	Parameter	s-boot				p-boot				t-boot			
		S(t)		h(t)		S(t)		h(t)		S(t)		h(t)	
		$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$	$\mathcal{W}$	$\mathcal{P}$
R_1	0.5, 0.75	0.29544	0.873	0.07523	0.891	0.28378	0.886	0.07324	0.835	0.34628	0.895	0.05505	0.895
R_2		0.21038	0.859	0.07314	0.842	0.20573	0.936	0.07104	0.896	0.24435	0.942	0.05714	0.888
R_3		0.24127	0.889	0.07045	0.908	0.23423	0.889	0.07528	0.897	0.28671	0.879	0.05471	0.897
R_4		0.23853	0.869	0.07411	0.893	0.23301	0.966	0.07069	0.914	0.28454	0.946	0.05193	0.897
R_5		0.23643	0.918	0.08954	0.901	0.22909	0.942	0.08528	0.913	0.28475	0.952	0.05879	0.895
R_6		0.20426	0.910	0.07011	0.895	0.20010	0.928	0.06828	0.926	0.24157	0.936	0.05224	0.886
R_7		0.15742	0.892	0.05649	0.869	0.15501	0.920	0.05525	0.908	0.17670	0.923	0.04606	0.871
R_8		0.17541	0.902	0.06093	0.877	0.17263	0.934	0.05945	0.907	0.20083	0.936	0.04800	0.823
R_9		0.17308	0.932	0.05935	0.877	0.17041	0.922	0.05797	0.917	0.19856	0.925	0.04699	0.853
R_{10}		0.18643	0.894	0.06420	0.843	0.18310	0.922	0.06233	0.886	0.21522	0.942	0.04946	0.859
R_1	0.85, 1.0	0.31824	0.843	0.08922	0.879	0.30892	0.848	0.08057	0.867	0.38068	0.898	0.06723	0.867
R_2		0.23656	0.861	0.08339	0.868	0.23125	0.870	0.07946	0.880	0.26713	0.888	0.06639	0.887
R_3		0.27553	0.870	0.08568	0.875	0.26815	0.956	0.08173	0.878	0.31452	0.943	0.06277	0.902
R_4		0.25835	0.900	0.09012	0.887	0.25231	0.940	0.08617	0.900	0.30549	0.950	0.06978	0.885
R_5		0.30012	0.850	0.09317	0.872	0.29171	0.930	0.08824	0.898	0.33972	0.920	0.06900	0.867
R_6		0.23756	0.890	0.07857	0.868	0.23329	0.917	0.07655	0.903	0.27315	0.923	0.06178	0.881
R_7		0.18867	0.896	0.06437	0.854	0.18568	0.929	0.06299	0.898	0.20540	0.923	0.05499	0.890
R_8		0.20558	0.897	0.06899	0.862	0.20233	0.921	0.06737	0.894	0.22904	0.917	0.05674	0.898
R_9		0.20092	0.896	0.06785	0.871	0.19771	0.932	0.06637	0.909	0.22418	0.931	0.05625	0.884
R_{10}		0.21548	0.891	0.07447	0.863	0.21147	0.915	0.07256	0.897	0.23939	0.911	0.06030	0.898
R_1	1.5, 0.85	0.36231	0.882	0.10451	0.860	0.35456	0.930	0.09965	0.927	0.43190	0.934	0.07256	0.886
R_2		0.27880	0.891	0.08134	0.847	0.27348	0.926	0.07812	0.909	0.31423	0.932	0.06353	0.878
R_3		0.30380	0.875	0.08848	0.837	0.29800	0.924	0.08490	0.904	0.34757	0.930	0.06529	0.921
R_4		0.30945	0.899	0.08723	0.848	0.30335	0.932	0.08379	0.911	0.35705	0.944	0.06391	0.900
R_5		0.30762	0.886	0.08911	0.889	0.30105	0.919	0.08494	0.879	0.34854	0.923	0.06491	0.898
R_6		0.26792	0.884	0.07936	0.850	0.26395	0.923	0.07704	0.901	0.29845	0.924	0.06365	0.897
R_7		0.21408	0.912	0.06396	0.838	0.21132	0.928	0.06247	0.884	0.22868	0.932	0.05570	0.898
R_8		0.22800	0.895	0.06930	0.844	0.22495	0.916	0.06750	0.887	0.24731	0.921	0.05867	0.910
R_9		0.23282	0.898	0.07046	0.873	0.22967	0.928	0.06882	0.919	0.25259	0.930	0.06003	0.908
R_{10}		0.23319	0.874	0.07193	0.842	0.22967	0.914	0.06981	0.891	0.25031	0.907	0.06072	0.886
R_1	2,2	0.44109	0.898	0.13105	0.873	0.43231	0.931	0.12650	0.948	0.44720	0.897	0.11339	0.858
R_2		0.36824	0.910	0.10699	0.863	0.36314	0.932	0.10361	0.930	0.36708	0.927	0.09854	0.874
R_3		0.38664	0.899	0.11435	0.861	0.38031	0.926	0.11052	0.929	0.38435	0.916	0.10481	0.902
R_4		0.39495	0.898	0.11599	0.873	0.38847	0.917	0.11222	0.937	0.39559	0.895	0.10499	0.868
R_5		0.38736	0.891	0.11455	0.851	0.38084	0.918	0.11065	0.921	0.38019	0.902	0.10416	0.826
R_6		0.33297	0.896	0.10286	0.842	0.32817	0.922	0.09988	0.903	0.31694	0.901	0.09924	0.887
R_7		0.27225	0.911	0.08629	0.879	0.26956	0.927	0.08418	0.911	0.26040	0.909	0.08732	0.834
R_8		0.29191	0.935	0.09239	0.872	0.28866	0.945	0.09001	0.913	0.27892	0.934	0.09280	0.882
R_9		0.29427	0.928	0.09188	0.869	0.29071	0.950	0.08955	0.912	0.28156	0.930	0.09108	0.892
R_{10}		0.29404	0.918	0.09288	0.853	0.29057	0.934	0.09030	0.916	0.27453	0.914	0.09292	0.900

that of ACIs. Consequently, the coverage probability goes down for the shorter length of the confidence interval. Hence, the proposed estimation procedures under PCS-II might be helpful for the lifetime practitioner/researchers to develop the guidelines for choosing the best estimation techniques.

ACKNOWLEDGMENTS

We sincerely thank the Editor, Associate Editor, and anonymous referees for their insightful comments and suggestions, which significantly improved the quality and clarity of this work. Their valuable input was instrumental in refining the manuscript, and we deeply appreciate their efforts.

Table 6: Estimates of the survival characteristics $\mathcal{S}$ for different t in case of real data.

Scheme	time	$\hat{S}(t)_M$	$\hat{S}(t)_P$	$\hat{S}(t)_S$	$\hat{S}(t)_E$	$\hat{S}(t)_G$	$\hat{h}(t)_M$	$\hat{h}(t)_P$	$\hat{h}(t)_S$	$\hat{h}(t)_E$	$\hat{h}(t)_G$
S_c^1		0.78344	0.77581	0.77476	0.77261	0.77151	0.00602	0.00621	0.00615	0.00603	0.00596
S_1^1		0.93249	0.92325	0.92275	0.92172	0.92118	0.00247	0.00280	0.00268	0.00244	0.00231
S_2^1	t=50	0.81057	0.80386	0.80295	0.80109	0.80014	0.00477	0.00496	0.00490	0.00477	0.00470
S_3^1		0.92374	0.91639	0.91596	0.91508	0.91463	0.00307	0.00334	0.00324	0.00304	0.00293
S_4^1		0.84957	0.84302	0.84235	0.84098	0.84029	0.00307	0.00324	0.00318	0.00307	0.00302
S_c^1		0.39219	0.38911	0.38679	0.38207	0.37967	0.00344	0.00346	0.00345	0.00344	0.00344
S_1^1		0.61751	0.60872	0.60527	0.59804	0.59421	0.00251	0.00256	0.00255	0.00251	0.00249
S_2^1	t=200	0.45579	0.45241	0.44974	0.44428	0.44147	0.00301	0.00303	0.00302	0.00301	0.00300
S_3^1		0.56303	0.55908	0.55649	0.55115	0.54839	0.00288	0.00290	0.00289	0.00287	0.00286
S_4^1		0.58007	0.57479	0.57229	0.56709	0.56439	0.00212	0.00217	0.00215	0.00212	0.00210
S_c^2		0.50168	0.49672	0.49353	0.48698	0.48360	0.01252	0.01262	0.01258	0.01250	0.01246
S_1^2		0.59031	0.58177	0.57829	0.57100	0.56714	0.01093	0.01113	0.01107	0.01094	0.01088
S_2^2	t=50	0.54012	0.53310	0.52976	0.52282	0.51921	0.01156	0.01172	0.01167	0.01157	0.01152
S_3^2		0.58363	0.57580	0.57260	0.56595	0.56248	0.01110	0.01127	0.01121	0.01110	0.01104
S_4^2		0.54913	0.54376	0.54058	0.53399	0.53056	0.01135	0.01148	0.01143	0.01133	0.01128
S_c^2		0.42186	0.41721	0.41407	0.40761	0.40427	0.01067	0.01075	0.01073	0.01069	0.01067
S_1^2		0.50626	0.49833	0.49448	0.48652	0.48237	0.00959	0.00972	0.00969	0.00962	0.00958
S_2^2	t=65	0.45971	0.45490	0.45153	0.44460	0.44102	0.01000	0.01009	0.01006	0.01001	0.00998
S_3^2		0.49947	0.49607	0.49243	0.48486	0.48089	0.00971	0.00977	0.00974	0.00967	0.00964
S_4^2		0.46863	0.46541	0.46195	0.45480	0.45109	0.00985	0.00991	0.00988	0.00982	0.00979

Scheme	time	ACIs		BCIs						HPDIs	
		$S(t)$	$h(t)$		$S(t)$ $[\mathcal{W}]$			$h(t)$ $[\mathcal{W}]$		$S(t)$	$h(t)$
		$\mathcal{W}$	$\mathcal{W}$	s-boot	p-boot	t-boot	s-boot	p-boot	t-boot	$\mathcal{W}$	$\mathcal{W}$
S_c^1		0.63162	0.01284	0.17508	0.17762	0.17712	0.00489	0.00481	0.00565	0.15711	0.00339
S_1^1		0.39312	0.01103	0.12814	0.12863	0.10409	0.00370	0.00385	0.00468	0.11203	0.00311
S_2^1	t=50	0.57287	0.01112	0.16770	0.16900	0.16619	0.00458	0.00440	0.00561	0.14882	0.00315
S_3^1		0.46221	0.01390	0.11355	0.11089	0.10422	0.00361	0.00345	0.00387	0.10520	0.00308
S_4^1		0.41529	0.00689	0.16008	0.15436	0.15314	0.00331	0.00336	0.00445	0.12863	0.00231
S_c^1		0.65159	0.00253	0.20771	0.20474	0.20857	0.00127	0.00124	0.00125	0.16599	0.00074
S_1^1		0.90544	0.00402	0.29154	0.29052	0.26534	0.00194	0.00184	0.00197	0.25247	0.00121
S_2^1	t=200	0.70909	0.00294	0.23085	0.22830	0.22120	0.00160	0.00156	0.00162	0.18984	0.00093
S_3^1		0.94641	0.00400	0.24873	0.24948	0.25422	0.00163	0.00161	0.00165	0.21125	0.00095
S_4^1		0.65089	0.00281	0.26672	0.26874	0.21355	0.00181	0.00178	0.00215	0.21436	0.00107
S_c^2		0.97569	0.01566	0.27779	0.27633	0.28298	0.00645	0.00636	0.00586	0.22187	0.00385
S_1^2		0.98259	0.01701	0.29669	0.28999	0.28198	0.00766	0.00732	0.00732	0.25065	0.00453
S_2^2	t=50	0.95709	0.01595	0.27914	0.27271	0.27563	0.00696	0.00669	0.00701	0.23077	0.00417
S_3^2		0.99988	0.01722	0.27835	0.27135	0.27176	0.00722	0.00702	0.00710	0.23845	0.00432
S_4^2		0.94550	0.01588	0.29541	0.30014	0.28853	0.00700	0.00688	0.00686	0.22929	0.00426
S_c^2		0.90097	0.01046	0.26271	0.26208	0.28382	0.00074	0.00071	0.00055	0.19974	0.00252
S_1^2		0.94943	0.01174	0.30240	0.29761	0.30065	0.00104	0.00095	0.00072	0.24200	0.00316
S_2^2	t=65	0.90499	0.01094	0.26274	0.26014	0.26846	0.00094	0.00092	0.00069	0.21692	0.00287
S_3^2		0.96209	0.01182	0.28030	0.27269	0.27761	0.00091	0.00088	0.00068	0.23597	0.00309
S_4^2		0.89885	0.01096	0.27326	0.26975	0.27870	0.00101	0.00098	0.00073	0.22157	0.00295

Table 7: Censoring schemes for the considered real data sets.

n, m	Schemes	Place of removals R_i
58, 30	$S_c^1 : (58, 58, 0^{58})$	Complete sample: no removals
	$S_1^1 : (58, 30, 28, 0^{29})$	Removals occurs at first stage
	$S_2^1 : (58, 30, 0^{29}, 28)$	Removals occurs at last stage
	$S_3^1 : (58, 30, 14, 0^{28}, 14)$	Removals occurs at first and last stage
	$S_4^1 : (58, 30, 0^{14}, 14^2, 0^{14})$	Removals occurs at some intermediate stages
38, 30	$S_c^2 : (38, 38, 0^{38})$	Complete sample: no removals
	$S_1^2 : (38, 30, 8, 0^{29})$	Removals occurs at first stage
	$S_2^2 : (38, 30, 0^{29}, 8)$	Removals occurs at last stage
	$S_3^2 : (38, 30, 4, 0^{28}, 4)$	Removals occurs at first and last stage
	$S_4^2 : (38, 30, 0^{14}, 4^2, 0^{14})$	Removals occurs at some intermediate stages

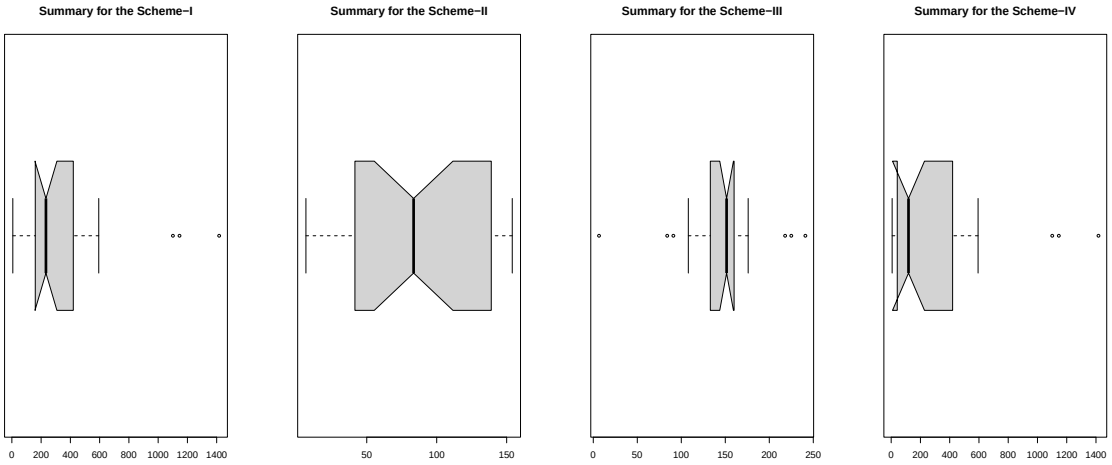


Figure 1: Box plot summary for the data set-I based on generated PCS-I samples.

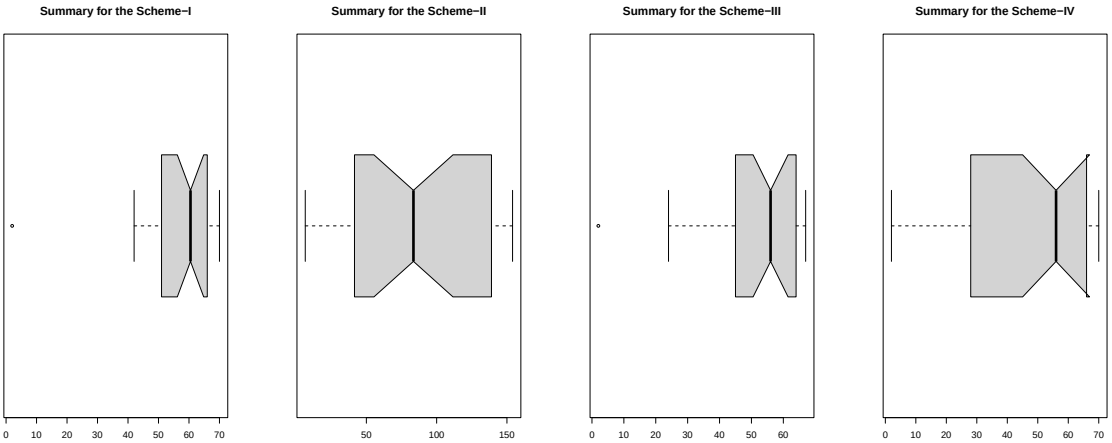


Figure 2: Box plot summary for the data set-II based on generated PCS-II samples.

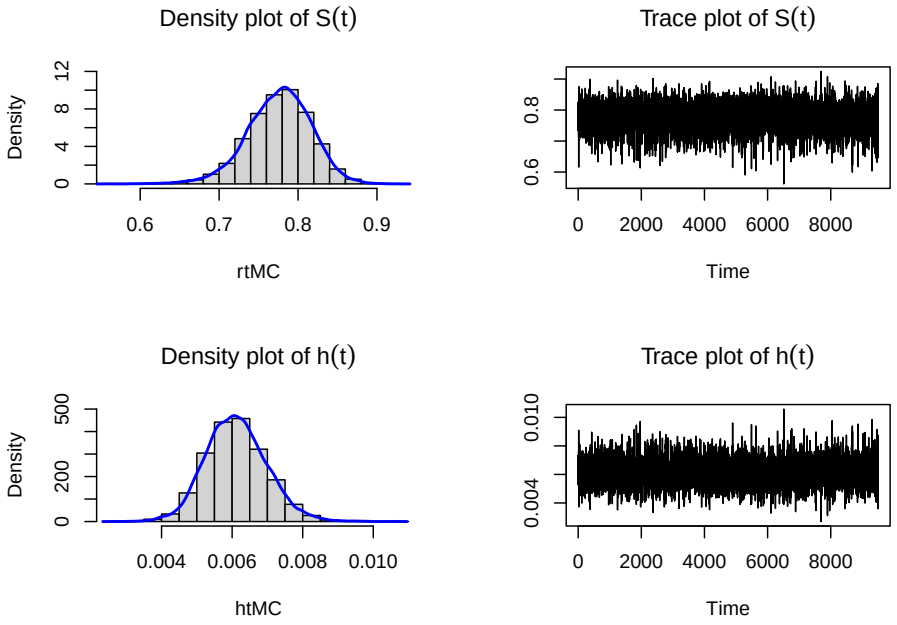


Figure 3: Posterior density and trace plot for the survival function for data set-I when $t = 50$ based on 10000 MCMC samples.

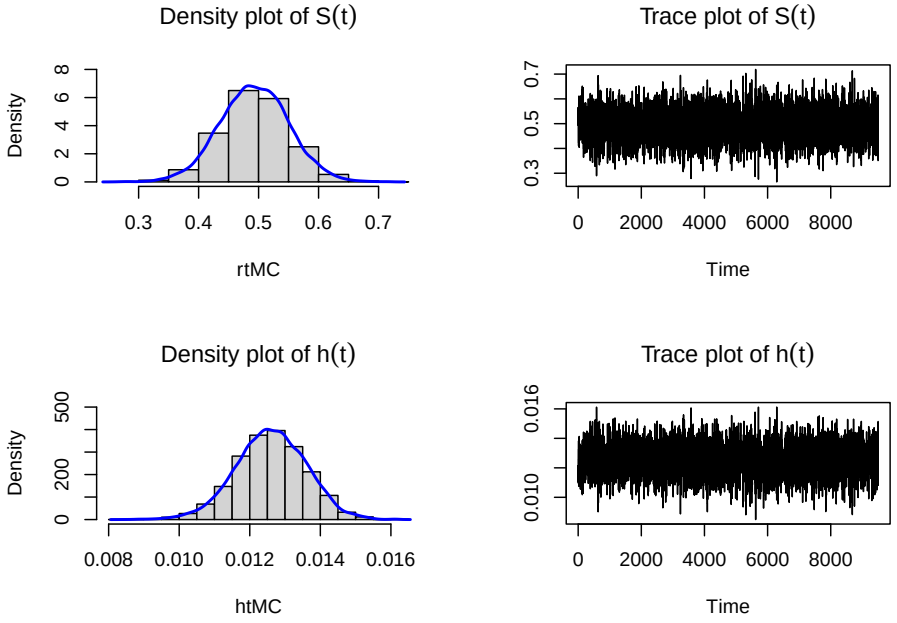


Figure 4: Posterior density and trace plot for the survival function for data set-II when $t = 50$ based on 10000 MCMC samples.

REFERENCES

Balakrishnan, N. (2007). Progressive censoring methodology: an appraisal. *TEST*, 16(2):211–259.

Balakrishnan, N. and Aggarwala, R. (2000). *Progressive censoring: theory, methods, and applications*. Statistics for Industry and Technology. Birkhäuser, Boston Basel.

Balakrishnan, N. and Sandhu, R. A. (1995). A Simple Simulational Algorithm for Generating Progressive Type-II Censored Samples. *The American Statistician*, 49(2):229–230.

Boucher, K. and Kerber, R. (2001). The shape of the hazard function for cancer incidence. *Mathematical and Computer Modelling*, 33(12-13):1361–1375.

Calabria, R. and Pulcini, G. (1996). Point estimation under asymmetric loss functions for left-truncated exponential samples. *Communications in Statistics - Theory and Methods*, 25(3):585–600.

Chen, M.-H. and Shao, Q.-M. (1999). Monte Carlo Estimation of Bayesian Credible and HPD Intervals. *Journal of Computational and Graphical Statistics*, 8(1):69–92.

Cohen, A. C. (1963). Progressively Censored Samples in Life Testing. *Technometrics*, 5(3):327–339.

Collett, D. (2023). *Modelling survival data in medical research*. Chapman and Hall/CRC.

Dey, S., Singh, S., Tripathi, Y. M., and Asgharzadeh, A. (2016). Estimation and prediction for a progressively censored generalized inverted exponential distribution. *Statistical Methodology*, 32:185–202.

Efron, B. (1988). Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve. *Journal of the American Statistical Association*, 83(402):414–425.

Efron, B. and Tibshirani, R. (1993). *An introduction to the bootstrap*. Number 57 in Monographs on statistics and applied probability. Chapman & Hall, New York.

Gelfand, A. E. and Smith, A. F. M. (1990). Sampling-Based Approaches to Calculating Marginal Densities. *Journal of the American Statistical Association*, 85(410):398–409.

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

Jiang, R., Ji, P., and Xiao, X. (2003). Aging property of unimodal failure rate models. *Reliability Engineering & System Safety*, 79(1):113–116.

Kayal, T., Tripathi, Y. M., Singh, D. P., and Rastogi, M. K. (2017). Estimation and prediction for Chen distribution with bathtub shape under progressive censoring. *Journal of Statistical Computation and Simulation*, 87(2):348–366.

Kleiber, C. and Kotz, S. (2003). *Statistical size distributions in economics and actuarial sciences*. Wiley Series in Probability and Statistics. Wiley, 1 edition.

Kumar, D., Yadav, A. S., Kumar, P., Kumar, P., Singh, S., and Singh, U. (2021). Transmuted inverse lomax distribution and its properties. *International Journal of Agricultural and Statistical Sciences*, 17(1):1–8.

Kundu, D. (2008). Bayesian Inference and Life Testing Plan for the Weibull Distribution in Presence of Progressive Censoring. *Technometrics*, 50(2):144–154.

Lodhi, C., Mani Tripathi, Y., and Kumar Rastogi, M. (2021). Estimating the parameters of a truncated normal distribution under progressive type II censoring. *Communications in Statistics - Simulation and Computation*, 50(9):2757–2781.

Maurya, R. K., Tripathi, Y. M., Rastogi, M. K., and Asgharzadeh, A. (2017). Parameter Estimation for a Burr XII Distribution under Progressive Censoring. *American Journal of Mathematical and Management Sciences*, 36(3):259–276.

Maurya, R. K., Tripathi, Y. M., Sen, T., and Rastogi, M. K. (2019). Inference for an Inverted Exponentiated Pareto Distribution Under Progressive Censoring. *Journal of Statistical Theory and Practice*, 13(1):2.

Nanda, A. K., Bhattacharjee, S., and Alam, S. (2007). Properties of aging intensity function. *Statistics & Probability Letters*, 77(4):365–373.

Pradhan, B. and Kundu, D. (2009). On progressively censored generalized exponential distribution. *TEST*, 18(3):497–515.

Rahman, J., Aslam, M., and Ali, S. (2013). Estimation and prediction of inverse lomax model via bayesian approach. *Caspian Journal of Applied Sciences Research*, 2(3):43–56.

Rastogi, M. K. and Tripathi, Y. M. (2014). Estimation for an inverted exponentiated Rayleigh distribution under type II progressive censoring. *Journal of Applied Statistics*, 41(11):2375–2405.

Shaked, M. and Shanthikumar, J. G., editors (1994). *Stochastic orders and their applications*. Probability and mathematical statistics. Academic Press, Boston.

Singh, S., Singh, U., and Yadav, A. (2016). Reliability estimation for inverse lomax distribution under type ii censored data using markov chain monte carlo method. *International Journal of Mathematics and Statistics*, 17(1):128–146.

Singh, S., Tripathi, Y. M., and Wu, S.-J. (2015a). On estimating parameters of a progressively censored lognormal distribution. *Journal of Statistical Computation and Simulation*, 85(6):1071–1089.

Singh, U., Singh, S. K., and Yadav, A. S. (2015b). Bayesian estimation for extension of exponential distribution under progressive type-ii censored data using markov chain monte carlo method. *Journal of Statistics Applications & Probability*, 4(2):275.

Smith, A. F. M. and Roberts, G. O. (1993). Bayesian Computation Via the Gibbs Sampler and Related Markov Chain Monte Carlo Methods. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 55(1):3–23.

Upadhyay, S., Vasishta, N., and Smith, A. (2001). Bayes inference in life testing and reliability via markov chain monte carlo simulation. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, 63(1):15–40.

Yadav, A. S., Bakouch, H. S., and Chesneau, C. (2022). Bayesian estimation of the survival characteristics for Hjorth distribution under progressive type-II censoring. *Communications in Statistics - Simulation and Computation*, 51(3):882–900.

Yadav, A. S., Singh, S. K., and Singh, U. (2016). On hybrid censored inverse lomax distribution: Application to the survival data. *Statistica*, 76(2):185–203.

Yadav, A. S., Singh, S. K., and Singh, U. (2019). Bayesian estimation of $R = P[Y<X]$ for inverse Lomax distribution under progressive type-II censoring scheme. *International Journal of System Assurance Engineering and Management*, 10(5):905–917.

REVSTAT – Statistical Journal
Volume 23, Number 5, Special Edition 2025, 569–583
https://doi.org/10.57805/revstat.v23i5.871

Prediction Methods for Upper and Lower Record Values and their Interrelationships

Authors: Christina Empacher ✉
– Institute of Statistics, RWTH Aachen University, Germany
christina.empacher@rwth-aachen.de

Udo Kamps
– Institute of Statistics, RWTH Aachen University, Germany
udo.kamps@rwth-aachen.de

Received: November 2024 Revised: June 2025 Accepted: June 2025

Abstract:

- Point and interval prediction methods for upper record values from some continuous distributions are adapted to respective statistical methods for lower record values from corresponding underlying distributions, namely from inverted distributions as their related counterparts. Interrelationships between corresponding point predictors for future upper and lower record values, these are maximum observed likelihood and maximum-product-of-spacings predictors, as well as corresponding prediction intervals for record values based on pivot statistics are shown. By this, known and new point predictors and prediction intervals for upper or lower record values from particular distributions can directly be transferred to analogous methods for lower or upper record values from the respective inverted distributions. The findings are exemplified for upper record values from an exponential and a Pareto distribution related to lower record values from an inverted exponential and a power function distribution.

Keywords:

- *upper record values; lower record values; inverted distribution; point prediction; maximum likelihood prediction; interval prediction.*

AMS Subject Classification:

- 62F99.

✉ Corresponding author

1. INTRODUCTION

There is a rich literature on statistical prediction methods for future record values based on various continuous distributions; we refer to, e.g., Ahsanullah (1980), Awad and Raqab (2000), Raqab (2007), Volovskiy and Kamps (2020a,b), Navarro (2022), Empacher et al. (2025), and the references, therein. Results in the literature mainly address upper record values rather than lower record values. However, as examined in this article, there is a close connection between point prediction as well as interval prediction for upper record values from a certain distribution and related inferential methods for lower record values from so-called inverted distributions as their counterparts. We consider maximum observed likelihood and maximum-product-of-spacings predictors and prediction intervals based on pivot statistics. These pivotal quantities make use of the cumulative hazard rate in case of upper records and the cumulative reversed hazard rate in case of lower records. The findings may be used to directly derive related results on point and interval prediction for lower and upper record values from known or new prediction methods for future upper or lower record values, respectively. The results are exemplified for upper record values from an exponential and a Pareto distribution related to lower record values from an inverted exponential and a power function distribution, respectively.

Let $(X_j)_{j\in\mathbb{N}}$ be a sequence of independent and identically distributed random variables with absolutely continuous cumulative distribution function (cdf) F_θ and probability density function (pdf) f_θ, where $\theta \in \Theta \subset \mathbb{R}^d$, $d \in \mathbb{N}$, is a vector of distributional parameters. Then, the random variables

$$T_u(1) = 1, \quad T_u(n+1) = \min\{j > T_u(n) \colon X_j > X_{T_u(n)}\},$$

$$T_l(1) = 1, \quad T_l(n+1) = \min\{j > T_l(n) \colon X_j < X_{T_l(n)}\}, \quad n \in \mathbb{N},$$

are called upper and lower record times, respectively. Then, the corresponding random variables $R_n = X_{T_u(n)}$ and $L_n = X_{T_l(n)}$, $n \in \mathbb{N}$, are called upper and lower record values, respectively.

The joint density functions are given by

$$f_\theta^{R_1,\dots,R_r}(x_1,\dots,x_r) = f_\theta(x_r) \prod_{j=1}^{r-1} \frac{f_\theta(x_j)}{1 - F_\theta(x_j)}$$

for upper record values $R_1, \dots, R_r$ and $-\infty < x_1 < x_2 < \cdots < x_r < \infty$ and by

$$f_\theta^{L_1,\dots,L_r}(y_1,\dots,y_r) = f_\theta(y_r) \prod_{j=1}^{r-1} \frac{f_\theta(y_j)}{F_\theta(y_j)}$$

for lower record values $L_1, \dots, L_r$ and $-\infty < y_r < \cdots < y_2 < y_1 < \infty$, respectively.

Let further H_θ with $H_\theta(x) = -\ln(1 - F_\theta(x))$ denote the cumulative hazard rate and G_θ with $G_\theta(x) = -\ln(F_\theta(x))$ denote the cumulative reversed hazard rate of F_θ. Then, for upper and lower record values, we find

$$H_\theta(R_r) \overset{d}{=} \sum_{i=1}^{r} Z_i \quad \text{and} \quad G_\theta(L_r) \overset{d}{=} \sum_{i=1}^{r} Z_i, \tag{1.1}$$

where $Z_1, \ldots, Z_r$ are independent and standard exponentially distributed random variables for all $r \in \mathbb{N}$ (see Arnold et al., 1998, p. 9). Those properties are helpful in order to construct pivotal quantities.

In this article we first recap three point prediction methods, namely maximum likelihood, maximum observed likelihood and maximum-product-of-spacings prediction for upper record values, which are known from literature. By analogy, corresponding methods for lower record values are obtained. In the next sections, interrelationships between point and interval prediction methods for upper and lower future record values are studied. The findings are illustrated for upper record values from exponential and Pareto distributions leading to related (new) statistical methods for lower record values from inverted exponential and power function distributions.

2. PREDICTION OF RECORD VALUES

We consider a setting where the first $r \in \mathbb{N}$ record values are observed and a s-th future record value with $s \in \mathbb{N}$, $s > r$ is to be predicted.

A well known approach is maximum likelihood prediction (see Kaminsky and Rhodin, 1985). In the case of upper record values the predictive likelihood function is given by the joint density function of $R_1, \ldots, R_r$ and R_s given by

(2.1)
$$f_\theta^{R_1,\ldots,R_r,R_s}(x_1,\ldots,x_r,x_s) = \frac{1}{(s-r-1)!}\left(\prod_{j=1}^{r}\frac{f_\theta(x_j)}{1-F_\theta(x_j)}\right) f_\theta(x_s)\left(-\ln\left(\frac{1-F_\theta(x_s)}{1-F_\theta(x_r)}\right)\right)^{s-r-1}$$

with $x_1 < \cdots < x_r < x_s$, and which, in order to derive a maximum likelihood predictor, is maximized regarding x_s and the unknown parameters θ. For lower record values based on the same cdf F_θ, the predictive likelihood function is given by

(2.2)
$$f_\theta^{L_1,\ldots,L_r,L_s}(y_1,\ldots,y_r,y_s) = \frac{1}{(s-r-1)!}\left(\prod_{j=1}^{r}\frac{f_\theta(y_j)}{F_\theta(y_j)}\right) f_\theta(y_s)\left(-\ln\left(\frac{F_\theta(y_s)}{F_\theta(y_r)}\right)\right)^{s-r-1}$$

with $y_1 > \cdots > y_r > y_s$. Again, the function has to be maximized regarding x_s and θ. For the joint densities we refer to Arnold et al. (1998, p. 10) and to Volovskiy and Kamps (2020a).

Let, in the following, $x_\star = (x_1, \ldots, x_r), y_\star = (y_1, \ldots, y_r)$, $R_\star = (R_1, \ldots, R_r)$ and $L_\star = (L_1, \ldots, L_r)$.

Another prediction method called maximum observed likelihood prediction (MOLP) is introduced in Volovskiy and Kamps (2020a) for upper record values. In the maximum observed likelihood procedure a so called observed predictive likelihood function $f_\theta^{R_1,\ldots,R_r|R_s}$, i.e., the conditional density function of $R_1, \ldots, R_r$ given R_s, is maximized regarding θ and x_s. It has been shown that the MOLP can be obtained by maximizing the function Ψ_{F_θ} given by

(2.3)
$$\Psi_{F_\theta}(\theta \mid x_\star) = \prod_{j=1}^{r}\frac{f_\theta(x_j)}{(1-F_\theta(x_j))\,(-\ln(1-F_\theta(x_r)))}, \quad x_1 < \cdots < x_r$$

regarding the distributional parameter θ with solution

$$\underset{\theta\in\Theta}{\operatorname{argmax}}\ \Psi_{F_\theta}(\theta \mid x_\star) = \theta_u^{(1)}(x_\star),\ \text{say},$$

and plugging in $\theta_u^{(1)}(R_\star)$ into the form of the MOLP which is given by

$$\pi_{MOLP}^{(s)upper} = \pi_{MOLP}^{(s)upper}(R_\star) = F^{-1}_{\theta_u^{(1)}(R_\star)}\left(1-\left(1-F_{\theta_u^{(1)}(R_\star)}(R_r)\right)^{\frac{s-1}{r}}\right). \tag{2.4}$$

By analogy with the derivation in Volovskiy and Kamps (2020a), a corresponding result for the prediction of lower records can be obtained by maximizing the function Φ_{F_θ} given by

$$\Phi_{F_\theta}\ (\theta \mid y_\star) = \prod_{j=1}^{r} \frac{f_\theta(y_j)}{F_\theta(y_j)\,(-\ln(F_\theta(y_r)))}, \quad y_1 > \cdots > y_r \tag{2.5}$$

regarding θ with solution

$$\underset{\theta\in\Theta}{\operatorname{argmax}}\ \Phi_{F_\theta}(\theta \mid y_\star) = \theta_l^{(1)}(y_\star),\ \text{say},$$

and plugging in $\theta_l^{(1)}(L_\star)$ into the form of the MOLP for lower records, which is given by

$$\pi_{MOLP}^{(s)lower} = \pi_{MOLP}^{(s)lower}(L_\star) = F^{-1}_{\theta_l^{(1)}(L_\star)}\left(\left(F_{\theta_l^{(1)}(L_\star)}(L_r)\right)^{\frac{s-1}{r}}\right). \tag{2.6}$$

In both cases of upper and lower record values the maximum observed likelihood procedure leads to predictors of the next record values R_{r+1} and L_{r+1} given by the last observed record value R_r and L_r, respectively. Since such a prediction is not useful in applications, another approach, which gives a non-trivial prediction of the very next record value, was introduced in Volovskiy and Kamps (2020b).

In the maximum-product-of-spacings prediction (MPSP) procedure for upper records, a function

$$\mathcal{P}_s(\theta \mid x_\star) = \prod_{i=1}^{r}\left(\frac{H_\theta(x_i)}{H_\theta(x_s)} - \frac{H_\theta(x_{i-1})}{H_\theta(x_s)}\right), \quad -\infty = x_0 < x_1 < \cdots < x_r$$

is maximized regarding the parameter θ and x_s. In Volovskiy and Kamps (2020b) it is shown that the procedure can be simplified by maximizing

$$\mathcal{P}_r(\theta \mid x_\star) = \mathcal{P}_{r,F_\theta}(\theta \mid x_\star) = \prod_{i=1}^{r}\left(\frac{H_\theta(x_i)}{H_\theta(x_r)} - \frac{H_\theta(x_{i-1})}{H_\theta(x_r)}\right), \quad -\infty = x_0 < x_1 < \cdots < x_s \tag{2.7}$$

regarding the parameter θ. If

$$\underset{\theta\in\Theta}{\operatorname{argmax}}\ \mathcal{P}_r(\theta \mid x_\star) = \theta_u^{(2)}(x_\star),\ \text{say},$$

is the respective solution, then $\theta_u^{(2)}(R_\star)$ is plugged in into the form of the MPSP, which is given by

$$\pi_{MPSP}^{(s)upper} = \pi_{MPSP}^{(s)upper}(R_\star) = F^{-1}_{\theta_u^{(2)}(R_\star)}\left(1-\left(1-F_{\theta_u^{(2)}(R_\star)}(R_r)\right)^{\frac{s}{r}}\right). \tag{2.8}$$

Again, by analogy with the derivation for upper record values it is seen that in a setting with lower record values with underlying cdf F_θ, a function

$$\mathcal{S}_r(\theta \mid y_\star) = \mathcal{S}_{r,F_\theta}(\theta \mid y_\star) = \prod_{i=1}^{r} \left(\frac{G_\theta(y_i)}{G_\theta(y_r)} - \frac{G_\theta(y_{i-1})}{G_\theta(y_r)} \right), \quad y_1 > \cdots > y_r,\ y_0 = \infty \tag{2.9}$$

has to be maximized regarding θ. If

$$\operatorname*{argmax}_{\theta \in \Theta} \mathcal{S}_r(\theta \mid y_\star) = \theta_l^{(2)}, \text{ say,}$$

is the resulting solution, then $\theta_l^{(2)}(L_\star)$ is plugged in into the predictor with the form

$$\pi_{MPSP}^{(s)lower} = \pi_{MPSP}^{(s)lower}(L_\star) = F^{-1}_{\theta_l^{(2)}(L_\star)} \left(\left(F_{\theta_l^{(2)}(L_\star)}(L_r) \right)^{\frac{s}{r}} \right). \tag{2.10}$$

Detailed proofs for the lower records versions of MOLP and MPSP can be found in Empacher (2024).

3. RELATIONSHIP BETWEEN PREDICTORS OF UPPER AND LOWER RECORDS

A strong relationship between predictors for upper and lower record values correlates with a relationship between the underlying distributions. According to Lehmann and Shaffer (1988), two positive random variables X and Y are said to have inverted distributions to each other if $X \overset{d}{=} 1/Y$. If X follows a strictly increasing and absolutely continuous cdf F_θ with pdf f_θ and quantile function F_θ^{-1}, the cdf $\widetilde{F}_\theta$, pdf $\tilde{f}_\theta$ and quantile function $\widetilde{F}_\theta^{-1}$ of Y can be written as

$$\begin{aligned}
\widetilde{F}_\theta(y) &= 1 - F_\theta\left(\frac{1}{y}\right), \quad y \in \left(\frac{1}{\omega(F_\theta)}, \frac{1}{\alpha(F_\theta)} \right), \\
\tilde{f}_\theta(y) &= f_\theta\left(\frac{1}{y}\right) \frac{1}{y^2}, \quad y \in \left(\frac{1}{\omega(F_\theta)}, \frac{1}{\alpha(F_\theta)} \right), \\
\widetilde{F}_\theta^{-1}(y) &= \frac{1}{F_\theta^{-1}(1-y)}, \quad y \in (0,1),
\end{aligned}$$

where $\alpha(F_\theta)$ and $\omega(F_\theta)$ denote the left and right endpoints of the support of the distribution defined by F_θ and, by convention $\frac{1}{0} = \infty$, $\frac{1}{\infty} = 0$. Such a distributional relationship between upper and lower record values yields the following relationship between MOLPs and MPSPs of upper and lower record values. Thus, the MOLP of an upper record value based on F_θ can directly be obtained from the MOLP of the respective lower record value based on its counterpart $\widetilde{F}_\theta$ and vice versa.

Theorem 3.1. *Let $\pi_{MOLP}^{(s)lower}$ denote the MOLP of the s-th lower record value based on a strictly increasing and absolutely continuous cdf F_θ with positive support and $\pi_{MOLP}^{(s)upper}$*

the MOLP of the s-th upper record value based on the inverted cdf $\widetilde{F}_\theta$*. Then, the MOLPs for lower and upper record values satisfy*

$$\pi^{(s)lower}_{MOLP}(L_\star) = \frac{1}{\pi^{(s)upper}_{MOLP}\left(\frac{1}{L_\star}\right)} \quad \text{and} \quad \pi^{(s)upper}_{MOLP}(R_\star) = \frac{1}{\pi^{(s)lower}_{MOLP}\left(\frac{1}{R_\star}\right)},$$

where, for short, $\frac{1}{x_\star} = \left(\frac{1}{x_1}, \ldots, \frac{1}{x_r}\right)$ *for* $x_\star = (x_1, \ldots, x_r) \in \mathbb{R}^r$.

Proof: For $s = r+1$, the result is obvious since $\pi^{(r+1)upper}_{MOLP}(R_\star) = R_r$ and $\pi^{(r+1)lower}_{MOLP}(L_\star) = L_r$. So let $s > r+1$.

For $x_1, \ldots, x_r \in (\alpha(F_\theta), \omega(F_\theta))$ with $0 < x_1 < \cdots < x_r$ we have $\frac{1}{x_1}, \ldots, \frac{1}{x_r} \in (\alpha(\widetilde{F}_\theta), \omega(\widetilde{F}_\theta))$ with $\frac{1}{x_1} > \cdots > \frac{1}{x_r} > 0$. Then, by using (2.3) and (2.5) we find

$$\begin{aligned}
\operatorname*{argmax}_{\theta\in\Theta} \Psi_{\widetilde{F}_\theta}(\theta \mid x_\star) &= \operatorname*{argmax}_{\theta\in\Theta} \left(\prod_{i=1}^{r} \frac{\tilde{f}_\theta(x_i)}{\left(1-\widetilde{F}(x_i)\right)\left(-\ln\left(1-\widetilde{F}(x_r)\right)\right)} \right) \\
&= \operatorname*{argmax}_{\theta\in\Theta} \left(\prod_{i=1}^{r} \frac{f_\theta\left(\frac{1}{x_i}\right)\frac{1}{x_i^2}}{F_\theta\left(\frac{1}{x_i}\right)\left(-\ln F_\theta\left(\frac{1}{x_r}\right)\right)} \right) \\
&= \operatorname*{argmax}_{\theta\in\Theta} \Phi_{F_\theta}\left(\theta \,\middle|\, \frac{1}{x^\star}\right) \cdot \prod_{i=1}^{r} \frac{1}{x_i^2} = \operatorname*{argmax}_{\theta\in\Theta} \Phi_{F_\theta}\left(\theta \,\middle|\, \frac{1}{x_\star}\right) \qquad (3.1)
\end{aligned}$$

which means that both maximizations lead to the same solution regarding $\theta \in \Theta$ given by $\hat{\theta} = \hat{\theta}(x_\star)$, say.

Moreover, using (2.4) and (2.6), we find for $x > 0$:

$$\widetilde{F}^{-1}_\theta\left(1-\left(1-\widetilde{F}_\theta(x)\right)^{\frac{s-1}{r}}\right) = \left(F^{-1}_\theta\left(\left(F_\theta\left(\frac{1}{x}\right)\right)^{\frac{s-1}{r}}\right)\right)^{-1} = \zeta_\theta(x), \text{ say,}$$

which implies that for the MOLP of upper record values from $\widetilde{F}_\theta$

$$\pi^{(s)upper}_{MOLP}(R_\star) = \zeta_\theta(R_r)\Big|_{\theta=\hat{\theta}(R_\star)}$$

and for the MOLP of lower record values from F_θ

$$\left(\pi^{(s)lower}_{MOLP}\left(\frac{1}{R_\star}\right)\right)^{-1} = \zeta_\theta(R_r)\Big|_{\theta=\hat{\theta}(\frac{1}{R_\star})}.$$

Thus we finally arrive at

$$\pi^{(s)upper}_{MOLP}(R_\star) = \left(\pi^{(s)lower}_{MOLP}\left(\frac{1}{R_\star}\right)\right)^{-1}.$$

The other statement follows analogously. □

The same simple transformation of predictors for upper and lower record values from $\widetilde{F}$ and F, respectively, is valid for the corresponding MPSPs.

Theorem 3.2. *Let $\pi^{(s)lower}_{MPSP}$ denote the MPSP of the s-th lower record value based on a strictly increasing and absolutely continuous cdf F_θ with positive support and $\pi^{(s)upper}_{MPSP}$ the MPSP of the s-th upper record value based on the inverted cdf $\widetilde{F}_\theta$. Then, the MPSPs for lower and upper record values satisfy*

$$\pi^{(s)lower}_{MPSP}(L_\star) = \frac{1}{\pi^{(s)upper}_{MPSP}\left(\frac{1}{L_\star}\right)} \quad \text{and} \quad \pi^{(s)upper}_{MPSP}(R_\star) = \frac{1}{\pi^{(s)lower}_{MPSP}\left(\frac{1}{R_\star}\right)},$$

where $\frac{1}{x_\star} = \left(\frac{1}{x_1}, \ldots, \frac{1}{x_r}\right)$ *for* $x_\star = (x_1, \ldots, x_r) \in \mathbb{R}^r$.

Proof: With the notations of Theorem 3.1 and its proof, we deduce from (2.7) and (2.9) (with $x_0 = 0$ and $\frac{1}{x_0} = \infty$):

$$\begin{aligned}
\operatorname*{argmax}_{\theta\in\Theta} \mathcal{P}_{r,\widetilde{F}_\theta}\left(\theta \,\middle|\, x_\star\right) &= \operatorname*{argmax}_{\theta\in\Theta} \prod_{i=1}^{r} \frac{1}{\ln\left(1-\widetilde{F}_\theta(x_r)\right)} \left(\ln\left(1-\widetilde{F}_\theta(x_i)\right) - \ln\left(1-\widetilde{F}_\theta\left(x_{i-1}\right)\right)\right) \\
&= \operatorname*{argmax}_{\theta\in\Theta} \prod_{i=1}^{r} \frac{1}{\ln F_\theta\left(\frac{1}{x_r}\right)} \left(\ln F_\theta\left(\frac{1}{x_i}\right) - \ln F_\theta\left(\frac{1}{x_{i-1}}\right)\right) \\
&= \operatorname*{argmax}_{\theta\in\Theta} \mathcal{S}_{r,F_\theta}\left(\theta \,\middle|\, \frac{1}{x_\star}\right) = \hat{\theta}(x_\star), \text{ say,}
\end{aligned}$$

which means that both functions are maximized regarding θ at $\hat{\theta} = \hat{\theta}(x_\star)$. Moreover, from (2.8) and (2.10), we obtain

$$\widetilde{F}^{-1}_{\hat{\theta}}\left(1-\left(1-\widetilde{F}_{\hat{\theta}}(x)\right)^{s/r}\right) = \left(F^{-1}_{\hat{\theta}}\left(F_{\hat{\theta}}\left(\frac{1}{x}\right)\right)^{s/r}\right)^{-1}, \; x > 0.$$

Then, by using (2.8) and (2.10) and arguing as in the proof of Theorem 3.1, we finally obtain

$$\pi^{(s)upper}_{MPSP}(R_\star) = \left(\pi^{(s)lower}_{MPSP}\left(\frac{1}{R_\star}\right)\right)^{-1}.$$

The other statement follows analogously. □

The transformation results for MOLPs and MPSPs are exemplified by applying them to upper record values from exponential and Pareto distribution.

Example 3.1. We consider upper record values from an exponential distribution with cdf $F^{Exp}_{\mu,\sigma}(x) = 1 - \exp\left(-\frac{x-\mu}{\sigma}\right)$ for $x > \mu$ with $\mu \in \mathbb{R}$ and $\sigma \in (0, \infty)$. The MOLP for the s-th upper exponential record value is calculated in Volovskiy and Kamps (2020a) and given by

$$\pi^{(s)upper}_{MOLP} = R_r + (R_r - \mu)\frac{s-r-1}{r}$$

if the location parameter μ is known and

$$\pi^{(s)upper}_{MOLP}(R_\star) = R_r + (R_r - R_1)\frac{s-r-1}{r}$$

if μ is unknown and, in this case, estimated by $\widehat{\mu} = R_1$, $r \geq 2$. In order to apply the transformation result Theorem 3.1 for lower records it is necessary to restrict the support of the distribution such that the records are based on positive random variables. Therefore we restrict the parameter space of μ to $(0, \infty)$. Moreover, the exponential distribution is taken as the inverted distribution.

Now, we can obtain a MOLP for lower record values from an inverted exponential distribution with cdf

$$\widetilde{F}^{Exp}_{\mu,\sigma}(x) = 1 - F^{Exp}_{\mu,\sigma}\left(\frac{1}{x}\right) = \exp\left(\frac{\mu - \frac{1}{x}}{\sigma}\right)$$

with $0 < x < \frac{1}{\mu}$. With Theorem 3.1, the MOLP in that setting is given by

$$\pi^{(s)lower}_{MOLP}(L_\star) = \frac{L_r}{1 + (1 - L_r\mu)\frac{s-r-1}{r}}$$

if μ is known and

$$\pi^{(s)lower}_{MOLP}(L_\star) = \frac{L_r}{1 + (1 - \frac{L_r}{L_1})\frac{s-r-1}{r}}$$

if μ is unknown and, in this case, estimated by $\widehat{\mu} = \frac{1}{L_1}$ according to the estimation procedure described in (3.1).

Example 3.2. Next, we consider upper record values from a Pareto distribution with cdf $F^{Par}_{\lambda,\beta}(x) = 1 - (\lambda/x)^\beta$ with $x > \lambda$ and $\lambda, \beta > 0$. The MPSP for upper Pareto records is calculated in Volovskiy and Kamps (2020b) and given by

$$\pi^{(s)upper}_{MPSP}(R_\star) = \lambda\left(\frac{R_r}{\lambda}\right)^{\frac{s}{r}}$$

if the scale parameter λ is known and

$$\pi^{(s)upper}_{MPSP}(R_\star) = R_1\left(\frac{R_r}{R_1}\right)^{\frac{s-1}{r-1}}$$

if λ is unknown and estimated by $\widehat{\lambda} = R_1\left(\frac{R_1}{R_r}\right)$, $r \geq 2$.

In order to apply Theorem 3.2, the Pareto distribution takes the role as inverted distribution. The inverted counterpart of the Pareto distribution is the power function distribution with cdf $F^{Pow}_{\lambda,\beta}(x) = 1 - F^{Par}_{\lambda,\beta}\left(\frac{1}{x}\right) = (\lambda x)^\beta$ for $0 < x < \frac{1}{\lambda}$. With Theorem 3.2 the MPSP for lower records from a power function distribution is given by

$$\pi^{(s)lower}_{MPSP}(L_\star) = \frac{(\lambda L_r)^{\frac{s}{r}}}{\lambda}$$

if λ is known and

$$\pi^{(s)lower}_{MPSP}(L_\star) = L_1\left(\frac{L_r}{L_1}\right)^{\frac{s-1}{r-1}}$$

if λ is unknown (cf. Empacher et al., 2023).

There are other interesting examples of pairs of distributions and respective relations between MOLPs and MPSPs for upper and lower record values. For a detailed discussion concerning Weibull distributions and inverted Weibull distributions, we refer to Section 2.2.3 in Empacher (2024).

A reviewer proposed to consider upper record values based on a Pareto-type-II distribution with distribution function

$$F_\beta(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^\beta, \ x > 0, \ \beta, \lambda > 0$$

(cf. **?**, (1.3.5)).

In this case, the function ψ in (2.3) is independent of β. If λ is assumed to be known, then, by using (2.4), the MOLP of the s-th upper record value R_s is given by

$$\pi_{MOLP}^{(s)upper} = \lambda \left(1 + \frac{R_r}{\lambda}\right)^{\frac{s-1}{r}} - \lambda.$$

By means of Theorem 3.1, the MOLP of the lower record value L_s based on the respective inverted distribution is directly obtained; here, the derivation of the MOLP of L_s via (2.5) and (2.6) would need some effort.

Similarly, the function $\mathcal{P}_r$ in (2.7) does not depend on β. Thus, for a known λ, the MPSP of the s-th upper record value R_s is obtained via (2.8) and given by

$$\pi_{MPSP}^{(s)upper} = \lambda \left(1 + \frac{R_r}{\lambda}\right)^{\frac{s}{r}} - \lambda.$$

Then, the MPSP of L_s based on the inverted distribution is derived by means of Theorem 3.2.

The distribution function $\widetilde{F}_\beta$ of the inverted distribution is given by

$$\widetilde{F}_\beta(y) = \left(1 - \frac{1}{\lambda y + 1}\right)^\beta, \ y > 0,$$

which corresponds to a scaled version (w.r.t. λ) of a particular exponentiated Pareto distribution (cf. **?**, Sec. 4.2).

For the MLP, there is no transformation result similar to Theorem 3.1 and Theorem 3.2 as indicated in the next example.

Example 3.3. The MLP for a future record value from an exponential distribution as in Example 3.1 is shown in Volovskiy (2018) and given by

$$\pi_{MLP}^{(s)} = R_r + (R_r - \mu)\frac{s - r - 1}{r + 1}$$

if μ is known and

$$\pi_{MLP}^{(s)} = R_r + (R_r - R_1)\frac{s - r - 1}{r + 1}$$

if μ is unknown, $r \geq 2$.

However, the MLP for the s-th lower record values from an inverted exponential distribution has a different structure and looks quite complicated (cf. Empacher, 2024).

Moreover, the existence of the MLP in an upper record setting does not even ensure the existence in the lower record setting.

A complete overview of MLP, MOLP and MPSP for upper records from exponential, Pareto and Weibull distributions as well as lower records from the inverted distributions as their counterparts can be found in Empacher (2024).

Analogous relationships between MOLPs and MPSPs of upper and lower record values from pairs of distributions are valid for transformations other than the reciprocal as considered in this section. By inspecting the proofs of Theorem 3.1 and Theorem 3.2 and by the derivations of a MOLP (see (2.4)) and a MPSP (see (2.6)) it is seen that respective transformation results concerning predictors for upper and lower record values hold true in case of considering, e.g., $Y = -X$.

4. INTERVAL PREDICTION

In applications, it may be more useful to work with prediction intervals for future record values instead of point predictors. There are several articles in the literature dealing with this topic, mainly with prediction intervals for upper record values. We refer to, e.g., Awad and Raqab (2000), Raqab (2007), Asgharzadeh et al. (2011), Wang et al. (2015), Empacher et al. (2023, 2025), and the references, therein.

In this section we are interested in the relationship between prediction intervals of upper and lower record values from inverted distributions based on pivot statistics. Since we know that record values transformed by the cumulative hazard rate are distributed as sums of exponential random variables (see (1.1)), we consider pivotal quantities. Three of them are given by

$$\begin{aligned} T_1 &= \frac{H_\theta(R_r) - H_\theta(R_1)}{H_\theta(R_s) - H_\theta(R_1)} \sim Beta(r-1, s-r), \\ T_2 &= \frac{H_\theta(R_r)}{H_\theta(R_s)} \sim Beta(r, s-r), \\ T_3 &= 2(H_\theta(R_s) - H_\theta(R_r)) \sim \chi^2(2(s-r)), \end{aligned}$$

where $Beta(a, b)$ denotes a beta distribution with parameters a and b (cf. Awad and Raqab, 2000; Raqab et al., 2007; Asgharzadeh et al., 2011; Empacher et al., 2023). Since an analogous transformation is given by applying the cumulative reversed hazard rate, we obtain corresponding statistics for lower record values (see Empacher et al., 2023; Empacher et al., 2025), which are given by

$$\begin{aligned} \widetilde{T}_1 &= \frac{G_\theta(L_r) - G_\theta(L_1)}{G_\theta(L_s) - G_\theta(L_1)} \sim Beta(r-1, s-r), \\ \widetilde{T}_2 &= \frac{G_\theta(L_r)}{G_\theta(L_s)} \sim Beta(r, s-r), \\ \widetilde{T}_3 &= 2(G_\theta(L_s) - G_\theta(L_r)) \sim \chi^2(2(s-r)). \end{aligned}$$

Depending on the underlying distribution of the record values, the pivotal quantities may not be independent of unknown distributional parameters. Then, these parameters could be replaced by respective maximum likelihood estimators as described following Lemma 4.1.

The pivot statistics indicate that a transformation of prediction intervals for upper records to prediction intervals of lower records may be possible.

Lemma 4.1. *Let $R_1, \ldots, R_s$ be the first $s \in \mathbb{N}$ upper record values with absolutely continuous and strictly increasing baseline cdf F_θ and $L_1, \ldots, L_s$ the first $s \in \mathbb{N}$ lower record values with baseline cdf $\widetilde{F}_\theta$. Furthermore, let $\alpha, \alpha_1, \alpha_2 \in (0,1)$ with $\alpha_1 + \alpha_2 = \alpha$.*

Let $T^\star(H_\theta(R_1), \ldots, H_\theta(R_r), H_\theta(R_s))$ be a pivot statistic, which depends on the record values only through the cumulative hazard rate H_θ of F_θ, and which does not depend on ϑ, and let $T^\star(H_\theta(R_1), \ldots, H_\theta(R_r), H_\theta(R_s)) = q_\alpha$ be explicitly solvable for R_s and any quantile q_α. Furthermore, let $I_1 = (l_{\alpha_1}(R_\star), u_{\alpha_2}(R_\star))$ denote the corresponding $1-\alpha$ prediction interval of R_s with lower bound $l_{\alpha_1}(R_\star)$ and upper bound $u_{\alpha_2}(R_\star)$. Then,

$$T^\star\left(H_\theta\left(\frac{1}{L_1}\right), \ldots, H_\theta\left(\frac{1}{L_r}\right), H_\theta\left(\frac{1}{L_s}\right)\right) \stackrel{d}{=} T^\star(H_\theta(R_1), \ldots, H_\theta(R_r), H_\theta(R_s))$$

is a pivot statistic, too, and the corresponding $1-\alpha$ prediction interval for L_s based on $L_1, \ldots, L_r$ takes the form $I_2 = \left(\frac{1}{u_{\alpha_2}\left(\frac{1}{L_\star}\right)}, \frac{1}{l_{\alpha_1}\left(\frac{1}{L_\star}\right)}\right)$.

Proof: From (1.1) we obtain

$$H_\theta(R_r) \stackrel{d}{=} G_\theta(L_r), \quad r \in \mathbb{N},$$

where G_θ is the cumulative reversed hazard rate of the cdf $\widetilde{F}_\theta$. Moreover, for $x \in (\alpha(F_\theta), \omega(F_\theta))$, we find

$$G_\theta\left(\frac{1}{x}\right) = -\ln\left(\widetilde{F}_\theta\left(\frac{1}{x}\right)\right) = -\ln(1 - F_\theta(x)) = H_\theta(x).$$

Since

$$H_\theta(R_r) \stackrel{d}{=} G_\theta(L_r) = H_\theta\left(\frac{1}{L_r}\right)$$

for all $r \in \mathbb{N}$, the stated equality in distribution follows. For $\alpha_1 + \alpha_2 = \alpha$, the relationship of prediction intervals is due to

$$\begin{aligned}
1-\alpha &= P\left(l_{\alpha_1}(R_\star) \le R_s \le u_{\alpha_2}(R_\star)\right) \\
&= P\left(q_{\alpha_1} \le T^\star(H_\theta(R_1), \ldots, H_\theta(R_r), H_\theta(R_s)) \le q_{1-\alpha_2}\right) \\
&= P\left(q_{\alpha_1} \le T^\star\left(H_\theta\left(\frac{1}{L_1}\right), \ldots, H_\theta\left(\frac{1}{L_r}\right), H_\theta\left(\frac{1}{L_s}\right)\right) \le q_{1-\alpha_2}\right) \\
&= P\left(l_{\alpha_1}\left(\frac{1}{L_\star}\right) \le \frac{1}{L_s} \le u_{\alpha_2}\left(\frac{1}{L_\star}\right)\right) \\
&= P\left(\frac{1}{u_{\alpha_2}\left(\frac{1}{L_\star}\right)} \le L_s \le \frac{1}{l_{\alpha_1}\left(\frac{1}{L_\star}\right)}\right),
\end{aligned}$$

where q_α denotes the quantile of the distribution of the pivot statistic. □

If $T^\star$ is a pivot statistic as in Lemma 4.1, which, however, is not independent of the unknown parameter θ, then, in case of existence, a maximum likelihood estimator of θ could be plugged in in order to obtain an approximate prediction interval $(l_{\alpha_1}(R_\star)), u_{\alpha_2}(R_\star))$ for the upper record value R_s based on $R_\star$. The maximum likelihood estimators based on $R_\star$ and $L_\star$ are related by

$$\begin{aligned}\operatorname*{argmax}_{\theta\in\Theta} f_\theta^{L_1,\ldots,L_r}(y_\star) &= \operatorname*{argmax}_{\theta\in\Theta} \tilde{f}_\theta(y_r)\prod_{i=1}^{r-1}\frac{\tilde{f}_\theta(y_i)}{\widetilde{F}_\theta(y_i)}\\ &= \operatorname*{argmax}_{\theta\in\Theta}\left(\prod_{i=1}^{r}\frac{1}{y_i}\right) f_\theta\left(\frac{1}{x_r}\right)\prod_{i=1}^{r-1}\frac{f_\theta\left(\frac{1}{y_i}\right)}{1-F_\theta\left(\frac{1}{y_i}\right)} = \operatorname*{argmax}_{\theta\in\Theta} f_\theta^{R_1,\ldots,R_r}\left(\frac{1}{y_\star}\right),\end{aligned}$$

where $y_\star = (y_1,\ldots,y_r) \in (0,\infty)^r$ with $y_1 > \cdots > y_r$ are realizations of $L_1,\ldots,L_r$. Since the maximum likelihood estimators have the same form, we arrive at

$$P\left(l_{\alpha_1}(R_\star) \le R_s \le u_{\alpha_2}(R_\star)\right) = P\left(\frac{1}{u_{\alpha_2}\left(\frac{1}{L_\star}\right)} \le L_s \le \frac{1}{l_{\alpha_1}\left(\frac{1}{L_\star}\right)}\right),$$

and thus we obtain an approximate prediction interval for the lower s-th record value from $\widetilde{F}$.

As a first example we consider prediction intervals of exponential records which are known from the literature and transform them in order to obtain prediction intervals for lower records from an inverted exponential distribution.

Example 4.1. Prediction intervals for exponential record values constructed from the statistics T_1, T_2 and T_3 are given by

$$\begin{aligned} I_1 &= \left(R_1 + \frac{R_r - R_1}{\beta_{1-\alpha_1}(r-1,s-r)},\ R_1 + \frac{R_r - R_1}{\beta_{\alpha_2}(r-1,s-r)}\right),\\ I_2 &= \left(\mu + \frac{R_r - \mu}{\beta_{1-\alpha_1}(r,s-r)},\ \mu + \frac{R_r - \mu}{\beta_{\alpha_2}(r,s-r)}\right),\\ I_3 &= \left(R_r + \frac{R_r-\mu}{2r}\chi^2_{\alpha_2}(2(s-r)),\ R_r + \frac{R_r-\mu}{2r}\chi^2_{1-\alpha_1}(2(s-r))\right),\end{aligned}$$

with $\alpha, \alpha_1, \alpha_2 \in (0,1)$ and $\alpha_1 + \alpha_2 = \alpha$ if the location parameter μ is known. The third interval is an approximate one, where σ is estimated based on the observed record values by its maximum likelihood estimator $\widehat{\sigma}_{ML} = \frac{R_r-\mu}{r}$. If μ is unknown, it is replaced by its maximum likelihood estimate $\widehat{\mu}_{ML} = R_1$ in the intervals above and I_2 and I_3 are approximate prediction intervals. The interval I_1 can be found for example in Gupta and Kirmani (1989), Arnold et al. (1998), Raqab (2007) and Asgharzadeh and Abdi (2011). I_2 is considered in Awad and Raqab (2000) for a setting with $\mu = 0$. The approximate interval I_3 is introduced in Awad and Raqab (2000).

By applying Lemma 4.1 we can construct prediction intervals for lower records from

an inverted exponential distribution. These are given by

$$I_1 = \left(\left(\frac{1}{L_1} + \frac{\frac{1}{L_r} - \frac{1}{L_1}}{\beta_{\alpha_2}(r-1, s-r)} \right)^{-1}, \left(\frac{1}{L_1} + \frac{\frac{1}{L_r} - \frac{1}{L_1}}{\beta_{1-\alpha_1}(r-1, s-r)} \right)^{-1} \right),$$

$$I_2 = \left(\left(\mu + \frac{\frac{1}{L_r} - \mu}{\beta_{\alpha_2}(r, s-r)} \right)^{-1}, \left(\mu + \frac{\frac{1}{L_r} - \mu}{\beta_{1-\alpha_1}(r, s-r)} \right)^{-1} \right),$$

$$I_3 = \left(\left(\frac{1}{L_r} + \frac{\frac{1}{L_r} - \mu}{2r} \chi^2_{1-\alpha_1}(2(s-r)) \right)^{-1}, \left(\frac{1}{L_r} + \frac{\frac{1}{L_r} - \mu}{2r} \chi^2_{\alpha_2}(2(s-r)) \right)^{-1} \right),$$

if μ is known. As in the exponential setting, I_3 is an approximate prediction interval. If μ is unknown, the maximum likelihood estimate $\widehat{\mu}_{ML} = \frac{1}{L_1}$ is plugged in the intervals above. Then only I_1 is an exact prediction interval.

The second example refers to upper record values from a Pareto distribution and related lower record values from a power function distribution.

Example 4.2. The prediction intervals of Pareto records based on T_1, T_2 and T_3 are

$$I_1 = \left(R_1 \exp\left(\frac{\ln(R_r) - \ln(R_1)}{\beta_{1-\alpha_1}(r-1, s-r)} \right), R_1 \exp\left(\frac{\ln(R_r) - \ln(R_1)}{\beta_{\alpha_2}(r-1, s-r)} \right) \right),$$

$$I_2 = \left(\lambda \exp\left(\frac{\ln(R_r) - \ln(\lambda)}{\beta_{1-\alpha_1}(r, s-r)} \right), \lambda \exp\left(\frac{\ln(R_r) - \ln(\lambda)}{\beta_{\alpha_2}(r, s-r)} \right) \right),$$

$$I_3 = \left(R_r \exp\left(\frac{\ln(R_r) - \ln(\lambda)}{2r} \chi^2_{\alpha_2}(2(s-r)) \right), R_r \exp\left(\frac{\ln(R_r) - \ln(\lambda)}{2r} \chi^2_{1-\alpha_1}(2(s-r)) \right) \right),$$

if the scale parameter λ is known. For the third interval the shape parameter β is estimated by $\widehat{\beta}_{ML} = \frac{r}{\ln(R_r) - \ln(\lambda)}$, which yields an approximate prediction interval. If λ is unknown, it is replaced by its maximum likelihood estimator $\widehat{\lambda}_{ML} = R_1$ in the intervals above. I_1 can be found for example in Asgharzadeh et al. (2011) and I_3 in Raqab et al. (2007); see also Empacher et al. (2023). In Empacher et al. (2025) and Schmiedt et al. (2025), the two- and one-sided versions of those intervals are compared to other prediction intervals constructed from the MPSP by means of their coverage and length.

Therefore, with Lemma 4.1 the corresponding prediction intervals of lower records from a power function distribution are

$$I_1 = \left(L_1 \exp\left(\frac{\ln(L_r) - \ln(L_1)}{\beta_{\alpha_2}(r-1, s-r)} \right), L_1 \exp\left(\frac{\ln(L_r) - \ln(L_1)}{\beta_{1-\alpha_1}(r-1, s-r)} \right) \right),$$

$$I_2 = \left(\frac{1}{\lambda} \exp\left(\frac{\ln(L_r) + \ln(\lambda)}{\beta_{\alpha_2}(r, s-r)} \right), \frac{1}{\lambda} \exp\left(\frac{\ln(L_r) + \ln(\lambda)}{\beta_{1-\alpha_1}(r, s-r)} \right) \right),$$

$$I_3 = \left(L_r \exp\left(\frac{\ln(L_r) + \ln(\lambda)}{2r} \chi^2_{1-\alpha_1}(2(s-r)) \right), L_r \exp\left(\frac{\ln(L_r) + \ln(\lambda)}{2r} \chi^2_{\alpha_2}(2(s-r)) \right) \right),$$

if λ is known. If λ is unknown, it is replaced by $\widehat{\lambda}_{ML} = \frac{1}{L_1}$. Those intervals can be found in Empacher et al. (2023) as well. There, they are compared in terms of coverage and length.

Analogously to Examples 4.1 and 4.2, Lemma 4.1 can be applied to other situations for upper and lower record values and their corresponding underlying distributions.

A detailed comparison in terms of coverage and expected length of prediction intervals for records from exponential, Pareto and Weibull distribution, and their inverted counterparts can be found in Empacher (2024).

ACKNOWLEDGMENTS

The authors would like to thank a reviewer for constructive comments, in particular regarding transformations other than the reciprocal.

REFERENCES

Ahsanullah, M. (1980). Linear prediction of record values for the two parameter exponential distribution. *Annals of the Institute of Statistical Mathematics*, 32(3):363–368.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (1998). *Records*. John Wiley & Sons, Inc, Hoboken, NJ.

Asgharzadeh, A. and Abdi, M. (2011). Confidence intervals and joint confidence regions for the two-parameter exponential distribution based on records. *Communications for Statistical Applications and Methods*, 18(1):103–110.

Asgharzadeh, A., Abdi, M., and Kuş, C. (2011). Interval estimation for the two-parameter Pareto distribution based on record values. *Selçuk Journal of Applied Mathematics*, pages 149–161.

Awad, A. M. and Raqab, M. Z. (2000). Prediction intervals for the future record values from exponential distribution: comparative study. *Journal of Statistical Computation and Simulation*, 65(1-4):325–340.

Empacher, C. (2024). *Point and interval prediction of record values*. PhD thesis, RWTH Aachen University.

Empacher, C., Kamps, U., and Schmiedt, A. B. (2025). Prediction intervals for future Pareto record claims. *European Actuarial Journal*, 15:163–197.

Empacher, C., Kamps, U., and Volovskiy, G. (2023). Statistical prediction of future sports records based on record values. *Stats*, 6(1):131–147.

Gupta, R. C. and Kirmani, S. (1989). On predicting repair times in a minimal repair process. *Communications in Statistics - Simulation and Computation*, 18(4):1359–1368.

Kaminsky, K. S. and Rhodin, L. S. (1985). Maximum likelihood prediction. *Annals of the Institute of Statistical Mathematics*, 37(3):507–517.

Lehmann, E. L. and Shaffer, J. P. (1988). Inverted distributions. *The American Statistician*, 42(3):191–194.

Navarro, J. (2022). Prediction of record values by using quantile regression curves and distortion functions. *Metrika*, 85(6):675–706.

Raqab, M. Z. (2007). Exponential distribution records: different methods of prediction. In Ahsanullah, M. and Raqab, M. Z., editors, *Recent Developments in Ordered Random Variables*, pages 239–251. Nova Science Publ, New York, NY.

Raqab, M. Z., Ahmadi, J., and Doostparast, M. (2007). Statistical inference based on record data from Pareto model. *Statistics*, 41(2):105–118.

Schmiedt, A. B., Empacher, C., and Kamps, U. (2025). One- and two-sided prediction intervals for future Pareto record values with applications. *Journal of Statistical Theory and Applications*, 24:489–514.

Volovskiy, G. (2018). *Likelihood-based prediction in models of ordered data*. PhD thesis, RWTH Aachen University.

Volovskiy, G. and Kamps, U. (2020a). Maximum observed likelihood prediction of future record values. *TEST*, 29(4):1072–1097.

Volovskiy, G. and Kamps, U. (2020b). Maximum product of spacings prediction of future record values. *Metrika*, 83(7):853–868.

Wang, B. X., Yu, K., and Coolen, F. P. (2015). Interval estimation for proportional reversed hazard family based on lower record values. *Statistics & Probability Letters*, 98:115–122.

REVSTAT – Statistical Journal
Volume 23, Number 5, Special Edition 2025, 584–596
https://doi.org/10.57805/revstat.v23i5.874

Testing Regression Models via Multiple Orderings of Regressors

Authors: Artyom Kovalevskii ✉
– Sobolev Institute of Mathematics,
Novosibirsk, Russia
artyom.kovalevskii@gmail.com

Received: December 2024 Revised: June 2025 Accepted: June 2025

Abstract:

- A new method for testing linear regression dependencies in multivariate data is proposed. The null hypothesis states that one of the data columns (response) depends linearly on the other columns (regressors) up to additive noise. The regression residuals are ordered in ascending order of the different columns, and empirical bridges are calculated for each sequence of such ordered regression residuals. The test statistic is the sum of integrals of the squares of the empirical bridges. The idea of the test is to detect deviations from linear dependence of the response on any of the regressors. Theorems on the limiting behavior of the empirical bridges and the test statistic are proved under the assumption of true null hypothesis. The calculation of the limiting distribution is based on the Smirnov's formula. Theorems are illustrated with simulated examples.

Keywords:

- *regression; weak convergence; goodness-of-fit; multiple ordering.*

AMS Subject Classification:

- 62J05, 62G30, 60F05, 62E20.

✉ Corresponding author

1. INTRODUCTION

There are many ways to test the fit of linear regression to real data. The most popular methods are based on calculating recursive regression residuals. This method was proposed in Brown et al. (1975) for normal regression and developed in Sen (1982), Bischoff (2016) and Sakhanenko et al. (2021) for non-Gaussian regression noise. The idea of this method is to predict the next values based on regression estimates calculated from previous values. The advantage of this method is its universality, it does not use the structure of the regressors' dependencies, and the limit process for the sums of recursive regression residuals is the standard Wiener process. The disadvantage of this method is its ambiguity, the researcher can order the data in different ways and obtain different results. This method cannot be applied to several orderings simultaneously, since the independence of responses is lost.

Another group of statistical tests does not use the idea of recursive regression residuals and is based on ordinary regression residuals. To detect deviations from the model, normalized successive sums of the regression residuals (empirical bridges) are computed, and the statistics are functions of the empirical bridges. This method was first proposed by MacNeill (1978) for a linear regression model of time series. Bischoff (1998) significantly relaxed the assumptions of MacNeill. An analysis of results in this direction can be found in Csorgo and Horváth (1997), Chapters 2 and 3, and MacNeill et al. (2020).

This approach is modified in MacNeill and Jandhyala (1985) for a nonlinear regression model of time series and in Kovalevskii and Shatalin (2016) for simple regression on order statistics and its generalizations. The advantage of this method is that it can be applied to several orderings simultaneously, as suggested in Chebunin and Kovalevskii (2021). The disadvantage of the method is that it requires the structure of the regressor to be taken into account, and the limit process for the empirical bridge is Gaussian, but different from the Wiener process. In multiple orderings over different regressors, the vector process of empirical bridges is studied.

We calculate the limit distribution for the sum of the integrals of the squares of empirical bridges under different orderings. To calculate the limit distribution, we need the Smirnov (1937) formula, its modern presentation can be found in Martynov (1992), §18.

Section 2 contains general theorems and their proofs. Section 3 describes the application of the theorems to a specific model. Section 4 illustrates the method by simulated examples. Section 5 contains summary.

2. MULTIPLE ORDERINGS OF A REGRESSION

The null hypothesis is that the regression response Y_i depends linearly on the regressors X_{ij},

$$Y_i = \sum_{j=1}^{m} X_{ij}\theta_j + \varepsilon_i, \quad i = 1, \dots, n. \tag{2.1}$$

This can be written in the matrix form

$$\mathbf{Y} = X\theta + \varepsilon.$$

The alternative hypothesis assumes that the response depends nonlinearly on at least one regressor instead of (2.1).

Our stochastic assumptions are the following:

1. $\varepsilon_1,\ \varepsilon_2,\ \ldots$ are independent and identically distributed random variables.
2. $\mathbf{E}\,\varepsilon_1 = 0$, $\mathbf{Var}\,\varepsilon_1 = \sigma^2 > 0$.
3. Rows $\mathbf{X}_i := (X_{i1}, \ldots, X_{im})$ of matrix X are independent and identically distributed random vectors that are independent with $\{\varepsilon_i\}_{i\geq 1}$.
4. Random vector $(X_{12}, \ldots, X_{1m})$ has a joint density, first component X_{11} can be discrete.
5. Matrix $G := \mathbf{E}\mathbf{X}_1^T\mathbf{X}_1$ exists and $\det \mathrm{G} > 0$.

Our data matrix $D = (\mathbf{Y}, X)$ is assumed to be known whereas vector θ to be unknown.

We sort rows of matrix D by ascending order of different columns of matrix X.

We do not sort rows by ordering on the first column because X_{11} can take discrete values. For example, we can let $X_{i1} \equiv 1$. Orderings on all other columns are determined uniquely a.s. due to the absolute continuity of the components.

So we have $m-1$ permutations

$$\mathbf{Y}^{(j)} = X^{(j)}\theta + \varepsilon^{(j)}, \quad j = 2, \ldots, m,$$

j is the number of the column which is sorted in ascending order.

We estimate vector parameter θ as MSE

$$\widehat{\theta} = (X^TX)^{-1}X^T\mathbf{Y}$$

and calculate vectors of regression residuals for different orderings

$$\widehat{\varepsilon}^{(j)} = \mathbf{Y}^{(j)} - X^{(j)}\widehat{\theta}, \quad j = 2, \ldots, m.$$

Note that vectors $\widehat{\varepsilon}^{(2)}, \ldots, \widehat{\varepsilon}^{(m)}$ consist of the same components that are ordered in different ways.

We designate by $\widetilde{\sigma}_n$ their empirical standard deviation. It is the same for any j:

$$\widetilde{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\widehat{\varepsilon}_i^{(j)}\right)^2, \quad j = 2, \ldots, m.$$

Let

$$\Delta_k^{(j)} = \sum_{i=1}^{k}\widehat{\varepsilon}_i^{(j)}, \quad j = 1, \ldots, m, \quad k = 0, \ldots, n,$$

be sums of regression residuals with $\Delta_0^{(j)} \equiv 0$.

Let

$$Z_n^{(j)} = \{Z_n^{(j)}(t),\, 0 \le t \le 1\}, \quad j = 2, \dots, m,$$

be an empirical bridge of regression residuals under the j-th ordering, that is, a piecewise linear random function with nodes

$$\left(\frac{k}{n},\ \frac{\Delta_k^{(j)}}{\widetilde{\sigma}_n \sqrt{n}} \right), \quad k = 0, 1, \dots, n. \tag{2.2}$$

Finally, we designate vector process $\mathbf{Z}_n = (Z_n^{(2)}, \dots, \widehat{Z}_n^{(m)})$.

We study weak convergence of vector process $\mathbf{Z}_n$ as $n \to \infty$ to a centered Gaussian vector process $\mathbf{Z} = \{\mathbf{Z}(t),\ 0 \le t \le 1\}$ in the space $C([0,1], \mathbb{R}^{m-1})$ (that is, the space of continuous on $[0,1]$ vector functions) with uniform metrics. We designate weak convergence of random variables and stochastic processes in corresponding spaces by $\Rightarrow$.

We need some additional designations to describe the limiting Gaussian process $\mathbf{Z}$.

Let $F_j(x)$ be cdfs of X_{1j}, $j = 2, \dots, m$:

$$F_j(x) = \mathbf{P}(X_{1j} \le x).$$

Let $Q_j(p)$ be quantile functions of X_{1j}, $2 \le j \le m$:

$$Q_j(p) = \inf\{x \in \mathbb{R} : p \le F_j(x)\}, \quad j = 2, \dots, m, \quad p \in (0,1).$$

Next, let

$$\mathbf{h}^{(j)}(p) = \mathbf{E}\{\mathbf{X}_1 | F_j(X_{1j}) = p\}, \quad j = 2, \dots, m, \quad p \in (0,1),$$

be conditional expectations,

$$\mathbf{L}^{(j)}(p) = \int_0^p \mathbf{h}^{(j)}(s)\, ds, \quad j = 2, \dots, m, \quad p \in [0,1],$$

be *induced theoretical generalized Lorentz curves*, see Davydov and Egorov (2000).

Theorem 2.1. *If (2.1) and the stochastic assumptions hold, then* $\mathbf{Z}_n \Longrightarrow \mathbf{Z}$. *Here* $\mathbf{Z}$ *is a centered* $(m-1)$*-dimensional vector Gaussian process with continuous a.s. sample paths and covariance matrix function* $K(s,t) = (K_{ij}(s,t))_{i,j=2}^m$,

$$K_{ij}(s,t) = \mathbf{P}(F_i(X_{1i}) \le s, F_j(X_{1j}) \le t) - \mathbf{L}^{(i)}(s) G^{-1} (\mathbf{L}^{(j)}(t))^T, \quad s,t \in [0,1]. \tag{2.3}$$

Proof of Theorem 2.1: SLLN guarantees convergences

$$\frac{1}{n} X^T X \to G \text{ a.s.}, \quad \frac{1}{n} \varepsilon^2 \to \sigma^2 \text{ a.s.}, \quad \frac{1}{n} X^T \varepsilon \to \mathbf{E}\mathbf{X}_1 \varepsilon_1 = \mathbf{0} \text{ a.s.}, \tag{2.4}$$

because $\mathbf{X}_1$ and ε_1 are independent and $\mathbf{E}\varepsilon_1 = 0$. So

$$\widehat{\theta} - \theta = (X^T X)^{-1} X^T (X\theta + \varepsilon) - \theta = (X^T X)^{-1} X^T \varepsilon \to \mathbf{0} \text{ a.s.}, \tag{2.5}$$

and $\widetilde{\sigma}_n^2$ does not depend on the order of rows. We use (2.4) and (2.5) and have

$$\begin{aligned} \widetilde{\sigma}_n^2 &= \frac{1}{n}(\mathbf{Y} - X\widehat{\theta})^2 = \frac{1}{n}(X\theta + \varepsilon - X\widehat{\theta})^2 \\ &= \frac{1}{n}(\widehat{\theta} - \theta)^T X^T X(\widehat{\theta} - \theta) + \frac{2}{n}(\widehat{\theta} - \theta)^T X^T \varepsilon + \frac{1}{n}\varepsilon^2 \to \sigma^2 \text{ a.s.} \end{aligned}$$

Therefore, convergence $\mathbf{Z}_n \Longrightarrow \mathbf{Z}$ takes place if and if only $\widehat{\mathbf{Z}}_n := \frac{\sigma}{\widetilde{\sigma}_n}\mathbf{Z}_n \Longrightarrow \mathbf{Z}$.

Vector process $\widehat{\mathbf{Z}}_n = (\widehat{\mathbf{Z}}_n^{(2)}, \ldots, \widehat{\mathbf{Z}}_n^{(m)})$ consists of components $\widehat{\mathbf{Z}}_n^{(j)}$ that are piecewise linear random function with nodes

$$\left(\frac{k}{n}, \frac{\Delta_k^{(j)}}{\sigma\sqrt{n}}\right), \quad k = 0,\ 1, \ldots,\ n.$$

We use Theorem 2 in Chebunin and Kovalevskii (2021). The convergence takes place if any row $\mathbf{X}_i$ is associated with random vector $\mathbf{U}_i = (U_{11}, \ldots, U_{1,m-1})$ distributed with copula $C(u_1, \ldots, u_{m-1})$ that has copula density

$$c(u_1, \ldots, u_{m-1}) = \frac{\partial^{m-1}}{\partial u_1 \cdots \partial u_{m-1}} C(u_1, \ldots, u_{m-1}),$$

and ordering of rows $\mathbf{U}_i$ in ascending order of component $j-1$ give the desired ordering of rows $\mathbf{X}_i$ on component j. Moreover, these orderings are in one-to-one correspondence.

We have joint density of vector $(X_{12}, \ldots, X_{1m})$, so we can let

$$U_{i,j-1} = F_j(X_{ij}), \quad j = 2, \ldots, m.$$

From Sklar's theorem Sklar (1959), if

$$F(x_2, \ldots, x_m) = \mathbf{P}(X_2 \le x_2, \ldots, X_m \le x_m),$$

then

$$F(x_2, \ldots, x_m) = C\left(F_2(x_2), \ldots, F_m(x_m)\right).$$

Let f be a density of the multivariate distribution F, then

$$f(x_2, \ldots, x_m) = c(F_2(x_2), \ldots, F_m(x_m)) \cdot f_2(x_2) \cdots f_m(x_m),$$

where f_j are marginal densities, so

$$c(u_1, \ldots, u_{m-1}) = \frac{f(Q_2(u_1), \ldots, Q_m(u_{m-1}))}{f_2(Q_2(u_1)) \cdots f_m(Q_m(u_{m-1}))}.$$

So we have the copula density, and $\mathbf{Z}_n \Longrightarrow \mathbf{Z}$. □

Corollary 2.1. *If $m = 2$ and $X_{i1} \equiv 1$ then one have a simple regression model*

$$Y_i = \theta_1 + X_{i2}\theta_2 + \varepsilon_i, \quad i = 1, \dots, n,$$

and, under the assumptions of Theorem 2.1, $\mathbf{Z}_n = Z_n^{(2)} \Longrightarrow Z^{(2)}$, a centered Gaussian process with continuous a.s. sample paths and covariance function

$$K_{22}(s,t) = \min(s,t) - st - \frac{(L_2^{(2)}(s) - sL_2^{(2)}(1))(L_2^{(2)}(t) - tL_2^{(2)}(1))}{\mathbf{Var}\, X_{12}}, \quad s,t \in [0,1], \tag{2.6}$$

with

$$L_2^{(2)}(p) = \int_0^p Q_2(s)\, ds.$$

Proof of Corollary 2.1: Note that $F_2(X_{12})$ has uniform distribution on $[0,1]$, so

$$\mathbf{P}(F_2(X_{12}) \le s, F_2(X_{12}) \le t) = \min(s,t).$$

Next,

$$G = \begin{pmatrix} 1 & \mathbf{E}X_{12} \\ \mathbf{E}X_{12} & \mathbf{E}X_{12}^2 \end{pmatrix}, \quad \det G = \mathbf{Var}\, X_{12} > 0,$$

$$\mathbf{L}^{(2)}(p) = (p, L_2^{(2)}(p)), \quad L_2^{(2)}(1) = \mathbf{E}X_{12},$$

$$L_2^{(2)}(p) = \int_0^p \mathbf{E}\{X_{12} | F_2(X_{12}) = s\}\, ds = \int_0^p Q_2(s)\, ds.$$

Substituting to (2.3), we obtain (2.6). □

Let $\widetilde{Z}_n = \{Z_n(t),\, 0 \le t \le m-1\}$ be a stochastic process defined as

$$\widetilde{Z}_n(t) = Z_n^{(j)}(t - j + 2), \quad j = 2, \dots, m, \quad t \in (j-2, j-1]; \quad \widetilde{Z}_n(0) = 0.$$

That is, we write components of process $\mathbf{Z}_n$ in sequence.

Let

$$\omega_n^2 = \int_0^{m-1} \widetilde{Z}_n^2(t)\, dt.$$

Theorem 2.2. *Under the assumptions of Theorem 2.1, let $X_{i1} = 1$, $i = 1, \dots, n$. Then $\widetilde{Z}_n \Longrightarrow \widetilde{Z}$, and $\widetilde{Z}$ is a centered Gaussian process on $[0, m-1]$ with continuous a.s. sample paths and covariance function $\widetilde{K}(s,t)$, $s,t \in [0, m-1]$, defined as*

$$\widetilde{K}(s,t) = K_{ij}(s - i + 2, t - j + 2), \quad i,j = 2, \dots, m, \quad s \in (i-2, i-1], \quad t \in (j-2, j-1].$$

Moreover,

$$\widetilde{Z}(t) = \sum_{i=1}^{\infty} \eta_i \sqrt{\mu_i} e_i(t), \tag{2.7}$$

where η_i have standard normal distribution and independent, μ_i are eigenvalues and $e_i(t)$ are normalized eigenfunctions of kernel $\widetilde{K}(s,t)$, $\mu_1 \geq \mu_2 \geq \cdots > 0$;

$\omega_n^2 \Rightarrow \omega^2$. *If $\mu_1 > \mu_2 > \cdots > 0$ and terms of the series below tend to 0, then ω^2 has cdf*

$$F_{\omega^2}(x) = 1 + \frac{1}{\pi}\sum_{k=1}^{\infty}(-1)^k \int_{\lambda_{2k-1}}^{\lambda_{2k}} \frac{e^{-\lambda x/2}}{\sqrt{-D(\lambda)}} \cdot \frac{d\lambda}{\lambda}, \quad x > 0, \tag{2.8}$$

$$D(\lambda) = \prod_{k=1}^{\infty}\left(1 - \frac{\lambda}{\lambda_k}\right), \qquad \lambda_k = \mu_k^{-1}.$$

Proof of Theorem 2.2: From Theorem 2.1, Gaussian process $\widetilde{Z}$ has continuous a.s. sample paths on any interval $(j-2, j-1]$, $j = 2, \ldots, m$. Next, assumption $X_{i1} \equiv 1$ guarantees

$$\sum_{i=1}^{n} \widehat{\varepsilon}_i^{(j)} = 0, \quad j = 2, \ldots, m.$$

So $\Delta_n^{(j)} = 0$ and $\widetilde{Z}(j-1) = \widetilde{Z}(j-1+) = 0$ a.s. for any $j = 2, \ldots, m$.

Therefore the limiting Gaussian process has continuous a.s. sample paths on $[0, m-1]$.

The Karhunen—Loève theorem states (2.7), see Jin (2014), Corollary 1. Weak convergence $\omega_n^2 \Rightarrow \omega^2$ is the consequence of the fact that ω_n^2 is the continuous in the uniform metrics functional of the process $\widetilde{Z}_n$, and

$$\omega^2 = \int_0^{m-1} \widetilde{Z}^2(t)\, dt.$$

CDF of ω^2 is given as (2.8) by the Smirnov's formula, see Martynov (1992), §18. □

3. PARAMETRICAL MODEL

We consider a specific null hypothesis of the form

$$Y_i = \theta_1 + X_{i2}\theta_2 + X_{i3}\theta_3 + \varepsilon_i, \quad i = 1, \ldots, n. \tag{3.1}$$

The alternative hypothesis is that the parameters do not remain constant across all ranges of regressor values.

If the assumptions of Theorem 2.2 are satisfied then $\widetilde{Z}_n \Longrightarrow \widetilde{Z}$ as $n \to \infty$, but the kernel of the limiting Gaussian process $\widetilde{Z}$ is unknown. So we estimate it. Let $X_{1:n,j} \leq X_{2:n,j} \leq \cdots \leq X_{n:n,j}$ be components of columns $j = 2, 3$ sorted in ascending order,

$$\mathbf{L}_n^{(2)}(t) = \left(t, \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} X_{i:n,2}, \frac{1}{n}\sum_{i=1}^{n} X_{i3}\mathbf{1}(X_{i2} \le X_{\lfloor nt\rfloor:n,2}) \right),$$

$$\mathbf{L}_n^{(3)}(t) = \left(t, \frac{1}{n}\sum_{i=1}^{n} X_{i2}\mathbf{1}(X_{i3} \le X_{\lfloor nt\rfloor:n,3}), \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor} X_{i:n,3} \right),$$

$$G_n = \begin{pmatrix} 1 & \frac{1}{n}\sum_{i=1}^{n} X_{i2} & \frac{1}{n}\sum_{i=1}^{n} X_{i3} \\ \frac{1}{n}\sum_{i=1}^{n} X_{i2} & \frac{1}{n}\sum_{i=1}^{n} X_{i2}^2 & \frac{1}{n}\sum_{i=1}^{n} X_{i2}X_{i3} \\ \frac{1}{n}\sum_{i=1}^{n} X_{i3} & \frac{1}{n}\sum_{i=1}^{n} X_{i2}X_{i3} & \frac{1}{n}\sum_{i=1}^{n} X_{i3}^2 \end{pmatrix},$$

$$K_{ii}^*(s,t) = \min(s,t) - \mathbf{L}_n^{(i)}(s)G_n^{-1}(\mathbf{L}_n^{(i)}(t))^T, \quad i = 2,3,$$

$$K_{23}^*(s,t) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}(X_{i2} \le X_{\lfloor ns\rfloor:n,2}, X_{i3} \le X_{\lfloor nt\rfloor:n,3}) - \mathbf{L}_n^{(2)}(s)G_n^{-1}(\mathbf{L}^{(3)}(t))^T,$$

$$K_{32}^*(s,t) = K_{23}^*(t,s),$$

The kernel $\widetilde{K}(s,t)$, $s,t \in [0,2]$, is estimated as

$$\widetilde{K}^*(s,t) = \begin{cases} K_{22}^*(s,t) & \text{on } [0,1]^2; \\ K_{33}^*(s-1,t-1) & \text{on } [1,2]^2; \\ K_{32}^*(s-1,t) & \text{on } [1,2]\times[0,1]; \\ K_{23}^*(s,t-1) & \text{on } [0,1]\times[1,2]. \end{cases}$$

We represent the kernel $\widetilde{K}^*$ on $[0,2]^2$ in the orthonormal basis $\{\sin\frac{\pi it}{2}, i \ge 1\}$ and approximate it by the matrix $R = (r_{ij})_{i,j=1}^M$ by calculation of

$$r_{ij} = \int_0^2 \int_0^2 \widetilde{K}^*(s,t) \sin\frac{\pi is}{2} \sin\frac{\pi jt}{2}\, dsdt.$$

We calculate eigenvalues μ_k^* of matrix R to estimate eigenvalues μ_k of kernel $\widetilde{K}(s,t)$. We calculate $F_{\omega^2}^*(x)$ to estimate $F_{\omega^2}(x)$ by substitution of μ_k^* instead of μ_k in (2.8).

Our statistics ω_n^2 has the following representation

$$\omega_n^2 = \frac{1}{3n}\sum_{k=1}^{2n-1} \widetilde{Z}_n\left(\frac{k}{n}\right)\left(2\widetilde{Z}_n\left(\frac{k}{n}\right) + \widetilde{Z}_n\left(\frac{k+1}{n}\right)\right).$$

Substituting (2.2), we have

$$\omega_n^2 = \frac{1}{3n^2\widetilde{\sigma}_n^2}\sum_{j=2}^{3}\sum_{k=1}^{n-1} \Delta_k^{(j)}\left(2\Delta_k^{(j)} + \Delta_{k+1}^{(j)}\right). \tag{3.2}$$

We construct a goodness-of fit test for the null hypothesis (3.1) by calculating

$$\text{p-value} = 1 - F^*_{\omega^2}(\omega^2_{n,obs}),$$

where $\omega^2_{n,obs}$ is the observed value of ω^2_n. If all the assumptions hold then cdf of ω^2_n almost coincides with $F^*_{\omega^2}$, and the p-value is distributed approximately uniformly on [0,1].

If the model is broken in such a way that the parameters for different regions of the regressor values differ then values of ω^2_n can be dramatically greater than the desired, and it can be found by this goodness-of-fit test.

4. SIMULATIONS

We illustrate the theorems with specific examples of simulated data. Experiments 1 and 2 correspond to this null hypothesis. In Experiment 1, the regressors are independent, and in Experiment 2, they are dependent.

Experiments 3 and 4 correspond to specific alternative hypotheses. In Experiment 3, with independent regressors, parameter θ_1 changes stepwise depending on the second regressor. In Experiment 4, with dependent regressors, parameter θ_1 changes stepwise depending on the sum of the regressors. The specific design of all the experiments is given below.

Experiments 1 and 2 illustrate that when the null hypothesis holds, the empirical distribution of the statistic matches the theoretical distribution found from Theorem 2.2. Experiments 3 and 4 illustrate that when the specific alternative hypothesis holds, the empirical distribution differs significantly from the theoretical distribution, allowing us to reject the null hypothesis with sufficient confidence based on the data.

We base on the results of Section 2, so we analyze the goodness-of fit test by simulations. Values of parameters do not affect asymptotics, and we can let $\theta_1 = \theta_2 = \theta_3 = \sigma^2 = 1$ to correspond the model. Our algorithm is the next.

1. Simulate random vectors $\{(X_{i2}, X_{i3})\}_{i=1}^n$. We take $n = 200$.
2. Simulate $\{\varepsilon_i\}_{i=1}^n$ independently of $\{(X_{i2}, X_{i3})\}_{i=1}^n$.
3. Calculate $\{Y_i\}_{i=1}^n$ by (3.1).
4. Calculate $\widehat{\theta}$, $\widetilde{\sigma}_n$, $\widehat{\varepsilon}^{(2)}$, $\widehat{\varepsilon}^{(3)}$, $\{\Delta_k^{(j)}\}_{k=0}^n$, $j = 2, 3$, and ω^2_n by (3.2).
5. Repeat points (2–4) $N = 20000$ times and plot an empirical cdf $F^*_n(x)$ of ω^2_n.
6. Calculate $\mathbf{L}^{(2)}_n(t)$, $\mathbf{L}^{(3)}_n(t)$, G_n, $\widetilde{K}^*(s,t)$, R, $\{\mu_k\}_{k=1}^M$. We take $M = 42$.
7. Calculate $F^*_{\omega^2}(x)$ and compare with the empirical cdf $F^*_n(x)$.

Experiment 1: X_{i2}, X_{i3}, ε_i are independent and have standard normal distribution. We set $\theta_1 = \theta_2 = \theta_3 = 1$ in (3.1).

Experiment 2: X_{i2}, X'_{i3}, ε_i are independent and have standard normal distribution, $X_{i3} = 3X_{i2} + X'_{i3}$. We set $\theta_1 = \theta_2 = \theta_3 = 1$ in (3.1) again.

Experiment 3: the same as in the Experiment 1 but we break (3.1) by letting $\theta_1 = 1$ for $X_{i2} < 0$ and $\theta_1 = 2$ for $X_{i2} \geq 0$.

Experiment 4: the same as in the Experiment 2 but we break (3.1) by letting $\theta_1 = 1$ for $X_{i2} + X_{i3} < 0$ and $\theta_1 = 2$ for $X_{i2} + X_{i3} \geq 0$.

Fig. 1 shows the regressor components in each experiment. In Experiments 1 and 2, the regression response at all points is given by formula (3.1). In Experiments 3 and 4, one is added to the response at the points marked with circles. Experiments 1 and 3 are conducted for uncorrelated regressor components, and Experiments 2 and 4 deal with positively correlated ones.

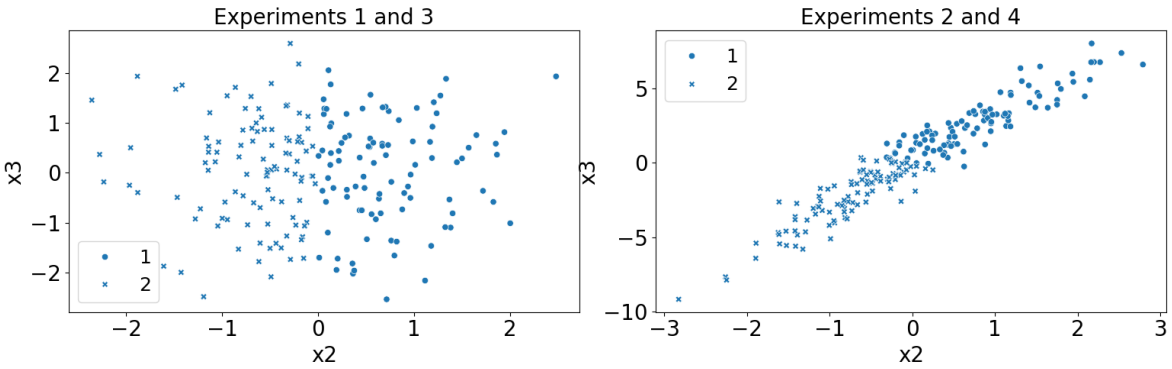


Figure 1: Components of regressors for stochastic experiments.

Fig. 2 demonstrates a very close match between the theoretical and empirical cdfs for Experiments 1 and 2, where the assumed probability model is the same as the data simulation model. Note that the distributions are almost indistinguishable (the maximum difference is about 1%), despite the fact that the regressor contains only 200 points. Different correlations of the regressor components affect the final distribution, but preserve the correspondence between the empirical and theoretical distributions. Experiments 3 and 4 demonstrate the difference between the theoretical and empirical cdfs under conditions of a difference between the assumed probability model and the data simulation model. Recall that in Fig. 1, a value equal to the standard deviation is added to the response values at the points marked with circles. The theoretical cdfs in Experiments 3 and 4 remain the same, but the empirical ones shift significantly to the right, which makes it possible to confidently diagnose the deviation of real data from the model.

All the code and calculations can be found at `https://github.com/ArtyomKovalevskii/3regression/blob/main/3regr_04.py`.

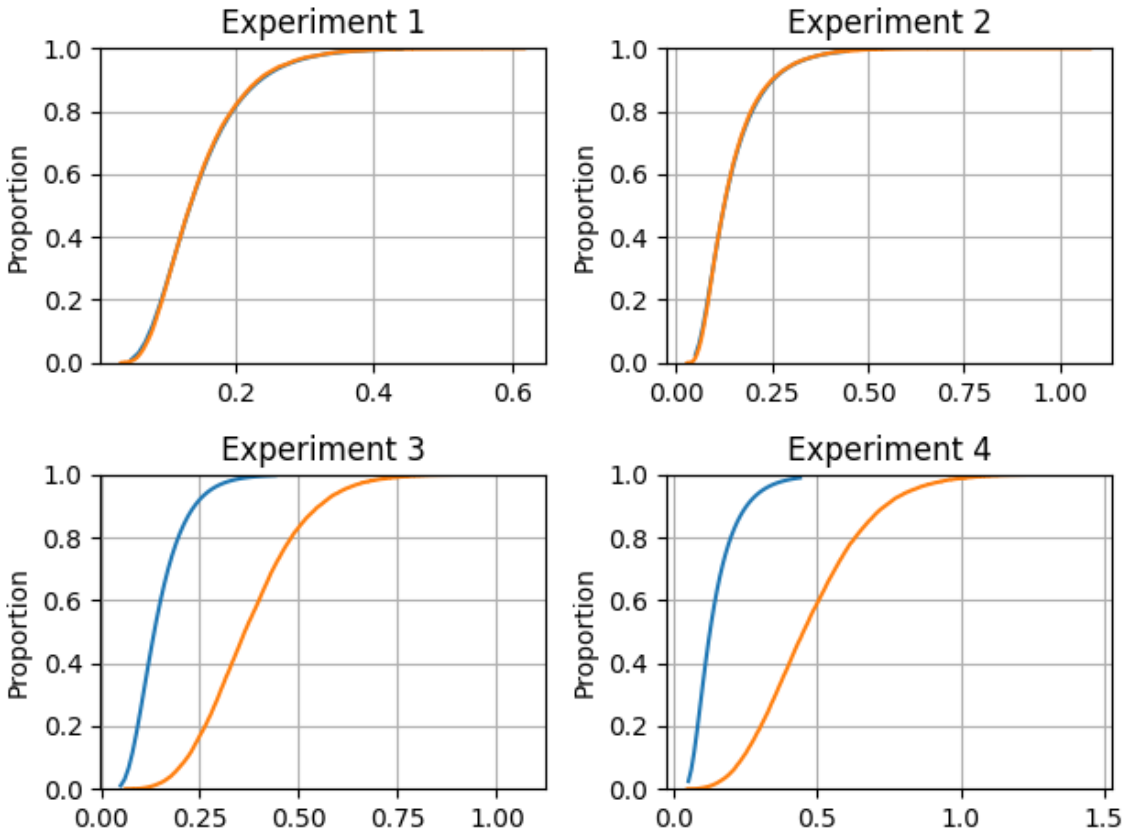


Figure 2: Theoretically estimated and empirical cdfs in stochastic experiments.

5. SUMMARY

A new statistical test is proposed to test the null hypothesis that one of the columns of a multivariate data set (the response) depends linearly on the other columns (the regressors) up to additive noise. The alternative hypothesis is that the dependence is non-linear for at least one of the columns. The test is based on ordering the entire data table by non-decreasing order over each of several regressors. For each such order, an empirical bridge is computed, i.e., a self-centered and self-normalized process of successive sums of regression residuals. Theorem 2.1 proves weak convergence of the resulting vector empirical bridge to the corresponding centered Gaussian process. The covariance function of the limit process uses the vector of induced Lorenz curves. A consequence of Theorem 2.1 is Corollary 2.1 for one order of regressors in the well-known simple regression model. Theorem 2.2 finds the limiting distribution for the omega-squared statistic, that is, for the sum of the integrals of the squares of the empirical bridges over all orders.

The preparatory stage for data analysis using Theorem 2.2 contains the estimation of all characteristics of the limit distributions, including the induced Lorenz functions. This stage is performed for a special three-parameter linear model.

The statistical test developed is illustrated with four simulated examples. Examples 1

and 2 contain data corresponding to a three-parameter linear model. They illustrate that the empirical cumulative distribution function of the omega-squared statistic matches the estimated theoretical cumulative distribution function very closely (within a difference of 0.01 for 200 points). Examples 1 and 2 differ from each other in the correlation of the regressors.

Examples 3 and 4 illustrate that the values of the omega square statistic increase significantly with deviations from the linear model. The regressors remain the same as in Examples 1 and 2, respectively. The response distribution in Examples 3 and 4 differs from the linear model. In Example 3, the distribution changes abruptly depending on the sign of the second regressor. In Example 4, the distribution changes abruptly depending on the sign of the sum of the two regressors. In both examples, the empirical cdfs are significantly to the right of the theoretical ones.

ACKNOWLEDGMENTS

The work is supported by the Fundamental scientific research of the SB RAS, project FWNF-2026-0030.

The author thanks an anonymous referee for a detailed analysis of the manuscript. The referee's helpful suggestions led to significant improvements in the presentation of the results.

REFERENCES

Bischoff, W. (1998). A functional central limit theorem for regression models. *Annals of Statistics*, 26:1398–1410.

Bischoff, W. (2016). On designs for recursive least squares residuals to detect alternatives. *Advances in Model-Oriented Design and Analysis*, pages 37–45.

Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society Series B*, 37(2):149–192.

Chebunin, M. G. and Kovalevskii, A. P. (2021). Asymptotics of sums of regression residuals under multiple ordering of regressors. *Siberian Electronic Mathematical Reports*, 18(2):1482–1492.

Csorgo, M. and Horváth, L. (1997). *Limit Theorems in Change-Point Analysis*. Wiley, New York.

Davydov, Y. and Egorov, V. (2000). Functional limit theorems for induced order statistics. *Mathematical Methods of Statistics*, 9(3):297–313.

Jin, S. (2014). *Gaussian processes: Karhunen-Loeve expansion, small ball estimates and applications in time series models*. ProQuest LLC, Ann Arbor.

Kovalevskii, A. and Shatalin, E. (2016). A limit process for a sequence of partial sums of residuals of a simple regression on order statistics with Markov-modulated noise. *Probability and Mathematical Statistics*, 36(1):113–120.

MacNeill, I. B. (1978). Limit processes for sequences of partial sums of regression residuals. *Annals of Probability*, 6:695–698.

MacNeill, I. B., Jandhyala, V., Kaul, A., and Fotopoulos, S. (2020). Multiple change-point models for time series. *Environmetrics*, 31(e2593).

MacNeill, I. B. and Jandhyala, V. K. (1985). The residual process for non-linear regression. *Journal of Applied Probability*, 22(4):957–963.

Martynov, G. (1992). Statistical tests based on empirical processes and related questions. *Journal of Soviet Mathematics*, 61:2195–2271.

Sakhanenko, A., Kovalevskii, A., and Shelepova, A. (2021). Remarks on invariance principle for one-parametric recursive residuals. *Siberian Electronic Mathematical Reports*, 18(2):1058–1074.

Sen, P. K. (1982). Invariance principles for recursive residuals. *Annals of Statistics*, 10(1):307–312.

Sklar, A. (1959). *Functions de Repartition an Dimension Set Leursmarges*. Publications de L'Institut de Statistique de L'Universite de Paris.

Smirnov, N. V. (1937). On the omega-squared distribution. *Matematicheskij Sbornik*, 2:973–993.

REVSTAT – Statistical Journal
Volume 23, Number 5, Special Edition 2025, 597–616
https://doi.org/10.57805/revstat.v23i5.921

Order-restricted Inferences on $P(X_1 < Y < X_2)$ for the Weibull Distribution Under Joint Progressive Censoring

Authors: Çağatay Çetinkaya ✉
– Department of Actuarial Sciences, Kırıkkale University, Kırıkkale, Türkiye
ccetinkaya@kku.edu.tr

Received: February 2025 Revised: October 2025 Accepted: October 2025

Abstract:

- Many experimental designs and real-world applications are susceptible to operational constraints that may result in system malfunction. This paper aims to estimate the reliability measure $\Omega = P(X_1 < Y < X_2)$, where the component strength Y lies between two stresses, X_1 and X_2. The variables X_1, Y, and X_2 are assumed to be independently distributed according to Weibull distributions with a common shape parameter and different scale parameters. The components are assumed to operate under a joint progressive type-II censoring scheme. In addition, a natural ordering constraint on the scale parameters, such as $\lambda_1 < \lambda_2 < \lambda_3$, is considered. Accordingly, the inference for Ω is derived under jointly progressive censored data subject to order-restricted scale parameters. The maximum likelihood estimates are obtained by solving the generalized isotonic regression problem. Additionally, Bayesian estimation is carried out under a gamma-Dirichlet prior distribution using the importance sampling algorithm. The asymptotic confidence intervals of the maximum likelihood estimators are presented, and the highest posterior density intervals are also derived for Bayesian inference. The performances of the proposed estimation methods are evaluated through simulation studies.

Keywords:

- *joint censoring; order-restricted inference; stress-strength; Weibull distribution.*

AMS Subject Classification:

- 60E05 , 62F15 , 62N01 , 62N02 , 62F10

✉ Corresponding author

1. INTRODUCTION

Reliability analysis plays a crucial role in engineering, quality control, and risk assessment, as it helps determine the probability that a system will perform its intended function under specified conditions. One of the important topics in reliability analysis is stress-strength models. Studies on stress-strength reliability estimation are widely covered in the statistical literature, especially in the field of reliability engineering. The simple stress-strength reliability model consists of a strength variable X and a stress variable Y acting upon it Birnbaum and McCarty (1958); Church and Harris (1970); Birnbaum et al. (1956). The system will function properly when X exceeds Y; in other words, $R = P(X > Y)$ denotes the system's reliability. Several authors have investigated many variations of this model. As an extension of the classical model, Bhattacharyya and Johnson (1974) introduced the multicomponent stress-strength reliability model involving two or more components. Furthermore, Johnson (1988) studied the reliability of systems with nonidentical component strengths, and Pandey et al. (1992) generalized this model to more than two groups of components. Recently, Mahdizadeh (2018) examined an extension of the well-known stress-strength reliability model in a non-parametric framework. These models have been addressed by several researchers under both complete and censored samples. For further details on stress-strength reliability models, the comprehensive study by Kotz et al. (2003) is recommended.

An important special case of stress-strength reliability models is the estimation of $P(X_1 < Y < X_2)$, which represents a situation where the strength Y must not only exceed the lower stress X_1 but also remain below the upper stress X_2. For example, certain devices cannot function at excessively high or extremely low temperatures. Similarly, a person's blood pressure has two boundaries, defined by systolic and diastolic pressures, within which it must remain. This type of stress-strength model can be viewed as a special case of the more general model $R_k = P\left(X_1 < X_2 < \cdots < X_k\right)$, one of the important scenarios that has not yet been fully explored. The estimation of R_k appears in isotonic regression problems, where it is essential to estimate R_k in order to determine level probabilities. The estimation of $\Omega = P(X_1 < Y < X_2)$ was introduced by Singh (1980) under the exponential distribution. Following the work of Singh (1980), many authors have contributed to this topic, and the studies included in Table 1 provide a summary of their contributions.

In reliability analysis, the lifetimes of components may not always be fully observed. To address the resulting challenges from unobserved data, censoring schemes (CS) are employed. Due to significant advancements in science and technology, it is often infeasible to observe a sufficient number of failures in life-testing experiments. Consequently, experimenters are motivated to adopt appropriate censoring strategies. Classical censoring schemes—such as Type-I, Type-II, and hybrid schemes—are widely utilized in the literature. In addition, progressive censoring schemes (PCS), which allow the systematic removal of experimental units during the study, have gained popularity in reliability research. These withdrawn units can be repurposed for further investigation, contributing to the efficiency and flexibility of the experimental process. The versatility of PCS has attracted considerable attention, leading to a broad range of studies involving reliability models under different PCS settings. For recent applications of stress-strength models under various censoring schemes, see Çetinkaya (2023), Maurya and Tripathi (2020), Kohansal (2022), and references therein.

Scheme	***Distribution***	***Author(s)***
Progressive type-II	Burr XII	Nayal et al. (2024)
Progressive type-II	Weibull	Choudhary et al. (2021)
Discrete setup	Geometric	Nayal et al. (2023)
Progressive first failure	Inverse Lomax	Ramadan et al. (2023)
Progressive first failure	Generalized inverted exponential	Saini (2024)
Progressive first failure	Length-bias exponential	Alotaibi et al. (2024)
Progressive first failure	Alpha power exponential	Almetwally et al. (2022)
Generalized prog. hybrid	Exponentiated exponential	Yousef et al. (2022)
Complete samples	Dagum	Karam and Attia (2021)
Progressive first failure	Kumaraswamy	Yousef and Almetwally (2021)
UMVUE and closed-form Bayes estimator		Rasethuntsa (2020)
Nonparametric inference		Guangming et al. (2013)

Table 1: Literature review for the estimation of $\Omega = P(X_1 < Y < X_2)$.

With the advancement of censoring plans, the concept of simultaneously censoring, referred to as joint censoring schemes (JCS), has emerged in the literature. JCS provides a useful framework for conducting comparative life tests on experimental units. In particular, it facilitates the simultaneous evaluation of components from different sources within the same system. As a result, JCS has gained attention in stress-strength reliability studies. For instance, Sultana et al. (2024) investigated two-sample balanced progressive censoring schemes for the estimation of stress-strength reliability. Similarly, Çetinkaya (2021) and Akbari Bargoshadi et al. (2024) explored the use of Type-II JCS within the stress-strength reliability context. The growing number of studies on this topic indicates the increasing interest and relevance of JCS in the field.

This study addresses the estimation of multi-stress-strength reliability, $\Omega = P(X_1 < Y < X_2)$, assuming that the involved components follow Weibull distributions with a restricted order on their scale parameters, i.e., $\lambda_1 < \lambda_2 < \lambda_3$. The units are subjected to the progressive joint censoring scheme as proposed by Balakrishnan et al. (2015). The problem is examined comprehensively throughout the study. Details of the underlying model are presented in Section 3. Subsequently, the generalized isotonic regression approach introduced by Barlow et al. (1972) is employed to derive the maximum likelihood estimator (MLE) of the reliability parameter Ω in Section 4. In addition to the point estimation, an approximate confidence interval for $\hat{\Omega}$ is constructed using the asymptotic distribution of the MLE, as detailed in Section 4.1. Further, an order-restricted Bayesian estimation procedure is developed in Section 5, utilizing the prior framework of Peña and Gupta (1990), along with the corresponding highest posterior density (HPD) credible intervals. Finally, the theoretical results are assessed through simulation studies, which are discussed in Section 7.

2. PRELIMINARIES

The isotonic regression problem seeks to minimize, $\sum_{i=1}^{k} w_i(g_i - x_i)^2$, subject to the constraint $x_i < x_j$ whenever $i \preceq j$, where $w_i > 0$ and the values g_i $(i = 1, 2, \ldots, k)$ are given. Here, $\preceq$ denotes a specified partial order over the index set $\{1, 2, \ldots, k\}$. The resulting

solution is referred to as the isotonic regression of g. Building upon this framework, Barlow et al. (1972) introduced a generalized loss function to extend the scope of isotonic regression methods

$$\sum_{i=1}^{k} \Delta_i w_i \tag{2.1}$$

where $\Delta_i = \Psi_i - g_i x_i$ for proper convex $\Psi : \mathbf{R} \to \mathbf{R}$ which minimized subject to some isotonicity constraints as discussed above. This problem is known as the generalized isotonic regression (GIR) problem, which involves minimizing an objective function under order restrictions. Specifically, let g and w, with $w_i > 0$, be given functions defined on the index set $A = \{1, 2, \ldots, k\}$, and regarded as elements of $\mathbb{R}^k$; that is, $g = (g_1, g_2, \ldots, g_k)$ and $w = (w_1, w_2, \ldots, w_k)$. Let $\Psi : \mathbb{R} \to \mathbb{R}$ be a proper convex function, and denote by $\mathbf{J}$ the cone of increasing vectors in $\mathbb{R}^k$, namely, $\mathbf{J} = \{x \in \mathbb{R}^k \mid x_1 \leq x_2 \leq \cdots \leq x_k\}$. The derivative of Ψ, denoted by Φ, is assumed to exist and be finite on $\mathbf{J}$; as expected for convex functions, Φ is non-decreasing. Now, suppose that for two real numbers a and b, the quantity Δ_i is defined as follows

$$\Delta(a,b) = \Delta_\Phi(a,b) = \begin{cases} \Psi(a) - \Psi(b) - (a-b)\Phi(b) & \text{if } a, b \in \mathbf{J} \\ \infty & \text{if } b \in \mathbf{J}, a \notin \mathbf{J}. \end{cases} \tag{2.2}$$

Accordingly, Barlow et al. (1972) introduced the following theorem to minimize equation 2.1 over the cone $\mathbf{J}$.

Theorem 2.1. *If $f(\cdot)$ is isotonic in A and if the range of $f(\cdot)$ is in $\mathbf{J}$, then g^*, the generalized isotonic regression of g with weights w minimizes*

$$\sum_{x \in A} \Delta[g(x), f(x)] w(x) \quad \textit{subject to} \quad x_i \leq x_j \qquad \textit{for all} \quad (i,j) \in A$$

in the class of isotonic functions $f(\cdot)$ with range in J and is given by

$$g^*(x_i) = \min_{t \geq i} \max_{s \leq i} \frac{\sum_{r=s}^{t} g(x_r)\, w(x_r)}{\sum_{r=s}^{t} w(x_r)} \tag{2.3}$$

Proof: See Chapter 1 in Barlow et al. (1972). □

3. MODEL DESCRIPTION

Let X_1, Y and X_2 be i.i.d. random variables from Weibull distributions with a common shape parameter and different scale parameters. That is, $\boldsymbol{X_2} \sim \text{Weibull}(\boldsymbol{\lambda_1, \beta})$, $\boldsymbol{Y} \sim \text{Weibull}(\boldsymbol{\lambda_2, \beta})$ and $\boldsymbol{X_1} \sim \text{Weibull}(\boldsymbol{\lambda_3, \beta})$ with the corresponding sample sizes n_1, n_2 and n_3. In this setup, higher Ω values can be obtained under restriction $\lambda_1 < \lambda_2 < \lambda_3$ since the mean time to failure of each component in the system is equal to

$$\lambda_j^{-\beta} \Gamma(1 + 1/\beta)$$

for $j = 1, 2, 3$. The distribution and density functions of the Weibull distributed components, $Z = (X_1, Y, X_2)$, are given as

$$F_Z(z; \lambda_j, \beta) = 1 - \exp\left\{-\lambda_j z^\beta\right\} \quad \text{and} \quad f_Z(z; \lambda_j, \beta) = \lambda_j \beta z^{\beta-1} \exp\left\{-\lambda_j z^\beta\right\}, \tag{3.1}$$

where $z > 0$, $\lambda_j, \beta > 0$ for $j = 1, 2, 3$ as λ_j's are scale parameters and β is the shape parameter. The formula to obtain $\Omega = P(X_1 < Y < X_2)$ is given by Singh (1980) as follows:

$$\Omega = P(X_1 < Y < X_2) = P(X_1 < Y) - P(X_1 < Y, X_2 < Y) = \int_{-\infty}^{\infty} F_{X_1}(y)\left[1 - F_{X_2}(y)\right] dF_Y(y)$$

In the case of $\boldsymbol{X_2} \sim$ Weibull$(\boldsymbol{\lambda_1, \beta})$, $\boldsymbol{Y} \sim$ Weibull$(\boldsymbol{\lambda_2, \beta})$ and $\boldsymbol{X_1} \sim$ Weibull$(\boldsymbol{\lambda_3, \beta})$ we get the reliability function as

$$\Omega = \int_0^\infty \left[1 - e^{-\lambda_3 y^\beta}\right] e^{-\lambda_1 y^\beta} \lambda_2 \beta y^{\beta-1} e^{-\lambda_2 y^\beta}\, dy = \frac{\lambda_2 \lambda_3}{(\lambda_1 + \lambda_2)(\lambda_1 + \lambda_2 + \lambda_3)} \tag{3.2}$$

The 3D plots presented in Figure 1 indicate that larger values of Ω are achieved as λ_3 increases. Then, all components are assumed to be subjected to the jointly progressive type-

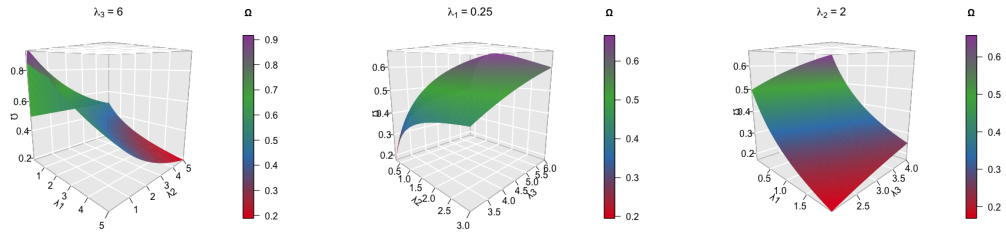


Figure 1: 3D representations of Ω under different parameter values.

II censoring scheme (JPCS) as defined by Balakrishnan et al. (2015). Considering the number of components as $k = 3$, the JPCS can be described as follows.

- The variables $X_{11}, X_{12}, \ldots, X_{1n_1}$ denote the lifetimes of n_1 units of component X_1.
- The variables $Y_1, Y_2, \ldots, Y_{n_2}$ denote the lifetimes of n_2 units of component Y.
- The variables $X_{21}, X_{22}, \ldots, X_{2n_3}$ denote the lifetimes of n_3 units of component X_2.

Let the samples be denoted by $Z = (X_1, Y, X_2)$. The total sample size is then given by $N = \sum_{i=1}^{3} n_i$, and the total number of observed failures is denoted by r, with $r < N$. Let $w_1 \leq w_2 \leq \cdots \leq w_N$ represent the order statistics of the N random variables $\{Z_{ij}\,;\, 1 \leq i \leq 3,\ 1 \leq j \leq n_i\}$. Therefore, under the Type-II JPCS for three samples, the observable data consist of (δ, w), where

$$w = (w_1, w_2, \ldots, w_r)\,, \qquad w_i \in \{Z_{h_i 1}, Z_{h_i 2}, \ldots, Z_{h_i n_i}\}$$

for $1 \leq h_1, h_2, \ldots, h_r \leq 3$, $(h_i \in Z^+)$ with r being a pre-fixed integer. Thus, associated to $(h_1, h_2, \ldots, h_r)$, let us define $\boldsymbol{\delta} = (\delta_1(h), \delta_2(h), \ldots, \delta_r(h))$ as

$$\delta_i(h) = \begin{cases} 1, & \text{if } h = h_i \\ 0, & \text{otherwise} \end{cases} \tag{3.3}$$

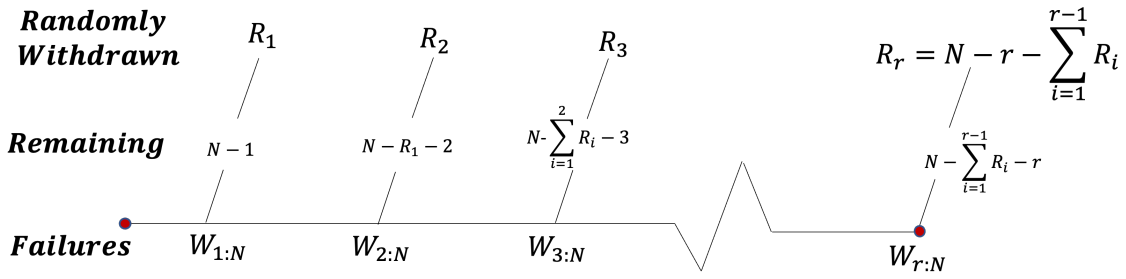


Figure 2: Schematic representation of the JPCS.

where $\sum_{h=1}^{3} \delta_i(h) = 1$ and $W_i \in Z_{h_i}$ for $1 \leq h_i \leq 3$, (1≤i≤r). In this censoring scheme (CS), the schematic representation is given in Figure 2. The vector $R = (R_1, R_2, \ldots, R_r)$ is specified at the beginning of the experiment and has the decomposition $R_i = \sum_{h=1}^{3} s_i(h), \quad 1 \leq i \leq r$, where $s_i(h)$ denotes the number of units withdrawn at the time of the i^{th} failure from the Z_h-sample. These quantities are unknown and are considered random variables. The observed data in this case consist of $(\boldsymbol{\delta}, R, w)$. Furthermore, let $M_r(h) = \sum_{i=1}^{r} \delta_i(h)$ denote the number of Z_h-failures among the w observations, for $1 \leq h \leq 3$ and

$$r = \sum_{h=1}^{3} M_r(h), \quad \sum_{h=1}^{3} s_i(h) = R_i, \quad \sum_{i=1}^{r} s_i(h) + M_r(h) = n_h, \quad \sum_{i=1}^{r} R_i = N - r.$$

4. MAXIMUM LIKELIHOOD ESTIMATION

In this section, the maximum likelihood estimate of Ω is derived for Weibull-distributed components under the Type-II JPCS. Accordingly, the corresponding likelihood function of the three components under the Type-II JPCS for the parameter vector $\theta = (\lambda_1, \lambda_2, \lambda_3, \beta)$ is expressed as

$$L(\theta, \boldsymbol{\delta}, \mathrm{s}, \mathrm{w}) \propto \prod_{i=1}^{r} \prod_{h=1}^{3} \left(f_h(w_i)\right)^{\delta_i(h)} \prod_{i=1}^{r} \prod_{h=1}^{3} \left(\bar{F}_h(w_i)\right)^{s_i(h)} \tag{4.1}$$

When the components follow Weibull distributions with the cumulative distribution functions given in Equation 3.1, the likelihood function can be expressed as

$$L(\theta, \boldsymbol{\delta}, \mathrm{s}, \mathrm{w}) \propto \beta^r \left(\prod_{h=1}^{3} \lambda_h^{M_r(h)} \right) \exp\left\{ - \sum_{h=1}^{3} \xi_h(\beta) \lambda_h \right\} \left(\prod_{i=1}^{r} \prod_{h=1}^{3} w_i^{\delta_i(h)(\beta-1)} \right) \tag{4.2}$$

where

$$\xi_h(\beta) = \sum_{i=1}^{r} \left[\delta_i(h) + s_i(h) \right] w_i^{\beta}.$$

and the corresponding log-likelihood function of the observed sample is obtained as

(4.3)
$$\ell\left(\theta, \boldsymbol{\delta}, \mathrm{s}, \mathrm{w}\right) \propto \sum_{h=1}^{3} M_r(h) \log(\lambda_h) + r \log(\beta) - \sum_{h=1}^{3} \lambda_h \xi_h(\beta) + (\beta - 1) \sum_{i=1}^{r} \sum_{h=1}^{3} \delta_i(h) \log(w_i)$$

The MLEs of the parameters can be obtained by maximizing Equation 4.3 with respect to λ_1, λ_2, λ_3, and β. According to Equation 3.2, lower reliability values are observed when $\lambda_3 < \lambda_2 < \lambda_1$, whereas higher reliability values are achieved when $\lambda_1 < \lambda_2 < \lambda_3$. A higher value of Ω means there is a higher probability that the component will successfully withstand the applied stresses without failure. In other words, the closer Ω is to 1, the more reliable the component or system is, since it is more likely to survive under operational conditions. Therefore, in engineering and reliability analysis, higher Ω values are more preferable, because they directly indicate better performance, greater safety, and lower risk of failure. Therefore, we consider the case of $\lambda_1 < \lambda_2 < \lambda_3$ to ensure the system operates under higher reliability conditions. Thus, the following theorem obtained following the GIR problem is proposed for the MLEs of the parameters.

Theorem 4.1. *For the fixed shape parameter β, if $L\left(\theta, \boldsymbol{\delta}, \mathrm{s}, \mathrm{w}\right)$ is maximized under the restriction on the scale parameters $\lambda_1 < \lambda_2 < \lambda_3$, the MLE of $\tilde{\lambda}_h^{(\beta)}$ of λ_h for $h = 1, 2, 3$ is given by*

(4.4)
$$\tilde{\lambda}_h^{(\beta)} = \left[\min_{u \le h} \max_{v \ge h} \frac{\sum_{r=u}^{v} \xi_r(\beta)}{\sum_{r=u}^{v} M_r(h)}\right]^{-1}, \quad u, v \in \tilde{I} = I.$$

Proof: Based on the order restriction,

$$\max_{\lambda_1 < \lambda_2 < \lambda_3} L\left(\theta, \boldsymbol{\delta}, \mathrm{s}, \mathrm{w}\right) = \max_{\lambda_1 < \lambda_2 < \lambda_3} \left\{\prod_{h \in I} \lambda_h^{M_r(h)} e^{-\lambda_h \xi_h(\beta)}\right\}$$

As $log(w)$ is a monotonically increasing function, this is equivalent to

$$\begin{aligned}
&\max_{\lambda_1 < \lambda_2 < \lambda_3} \left\{\sum_{h \in I} \left[M_r(h) \log \lambda_h - \lambda_h \xi_h(\beta)\right]\right\} \\
&= \max_{\lambda_1 < \lambda_2 < \lambda_3} \left\{\sum_{h \in I} \left[-M_r(h) \log\left(\frac{1}{\lambda_h}\right) - \frac{\xi_h(\beta)}{\frac{1}{\lambda_j}}\right]\right\} \\
&= \max_{\frac{1}{\lambda_1} > \frac{1}{\lambda_2} > \frac{1}{\lambda_3}} \left\{\sum_{h \in I} \left[-\log\left(\frac{1}{\lambda_h}\right) - \frac{\xi_h(\beta)}{\frac{M_r(h)}{\lambda_j}}\right] M_r(h)\right\} \\
&= \min_{\frac{1}{\lambda_1} > \frac{1}{\lambda_2} > \frac{1}{\lambda_3}} \left\{\sum_{h \in I} \left[\log\left(\frac{1}{\lambda_h}\right) + \frac{\xi_h(\beta)}{\frac{M_r(h)}{\lambda_j}}\right] M_r(h)\right\}.
\end{aligned}$$

Considering the convex function given in Eq (1), we take $\Psi(u) = -\log u$

(4.5)
$$\Delta_\phi(g(x), f(x)) = -\log g(x) + \log f(x) + (g(x) - f(x))/f(x)$$

Now the problem

$$\text{Minimize} \sum_{x \in A} \Delta_\Phi[g(x), f(x)] w(x)$$

subject to f is isotonic, is equivalent to

$$\text{Minimize} \sum_{x \in A} \left[\log f(x) + \frac{g(x)}{f(x)} \right] w(x)$$

subject to f is isotonic. Let, $A = 1, 2, 3.$ and

$$g(h) = \xi_h(\beta)/M_r(h), \qquad f(h) = 1/\lambda_h, \qquad w(h) = M_r(h), j \in I.$$

From Theorem 2.1, g^*,the GIR of g solves the problem and is given by

$$\frac{1}{\tilde{\lambda}_h^{(\beta)}} = \left[\min_{u \le h} \max_{v \ge h} \frac{\sum_{r=u}^{v} \xi_r(\beta)}{\sum_{r=u}^{v} M_r(h)} \right], \quad u, v \in \tilde{I} = I.$$

□

For the MLE of β, we need to maximize the profile log-likelihood function of β, say $\psi(\beta)$ where

$$\psi(\beta) = \sum_{h=1}^{3} M_r(h) \log(\tilde{\lambda}_h) + r \log(\beta) - \sum_{h=1}^{3} \tilde{\lambda}_h \xi_h(\beta) + (\beta - 1) \sum_{i=1}^{r} \sum_{h=1}^{3} \delta_i(h) \log(w_i) \tag{4.6}$$

The MLEs of the parameters can be obtained by replacing λ_1, λ_2 and λ_3 in Equation 4.6. Then, the MLE of β is obtained by solving the non-linear Equation 4.6. Following, $\hat{\beta}$ is replaced in Equation 4.4 with its estimate and $\tilde{\lambda}_1$, $\tilde{\lambda}_2$ and $\tilde{\lambda}_3$ are obtained respectively. After all, the MLE of Ω can be obtained by replacing the parameters in Equation 3.2 from the invariance property of the MLEs. Consequently, the MLE of Ω, $\hat{\Omega}$ is obtained as

$$\tilde{\Omega} = \frac{\tilde{\lambda}_2 \tilde{\lambda}_3}{(\tilde{\lambda}_1 + \tilde{\lambda}_2)(\tilde{\lambda}_1 + \tilde{\lambda}_2 + \tilde{\lambda}_3)} \tag{4.7}$$

4.1. Asymptotic Confidence Interval

The asymptotic confidence interval (ACI) for Ω can be proposed as an approximate confidence interval. In order to construct an ACI for $\hat{\Omega}$, the observed Fisher information matrix of the parameter vector $\theta = (\lambda_1, \lambda_2, \lambda_3, \beta)$, denoted by $I(\theta)$, is required. In this study, it is clearly seen that there are significant difficulties in obtaining the expected values in the Fisher information matrix $I(\theta)$, specifically $E\left(-\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}\right)$ for $i, j = 1, 2, 3$. The main reason for these difficulties is the lack of the exact distribution of the jointly censored samples $(w \,|\, \delta, s)$, and therefore the exact expectations cannot be evaluated. To overcome this issue, the observed Fisher information matrix, $I(\theta) = -\frac{\partial^2}{\partial\theta_i\,\partial\theta_j}$ is employed. The elements of the Fisher information matrix, which are the second partial derivatives of the log-likelihood

function (8) with respect to $\lambda_1, \lambda_2, \lambda_3$, and β, are obtained as follows:

$$\begin{aligned}
&\frac{\partial^2 \ell}{\partial \lambda_1^2} = -\frac{M_r(1)}{\lambda_1^2}, \qquad \frac{\partial^2 \ell}{\partial \lambda_2^2} = -\frac{M_r(2)}{\lambda_2^2}, \qquad \frac{\partial^2 \ell}{\partial \lambda_3^2} = -\frac{M_r(3)}{\lambda_3^2} \\
&\frac{\partial^2 \ell}{\partial \beta^2} = -\frac{r}{\beta^2} - \sum_{h=1}^{3} \lambda_h \xi_h''(\beta), \qquad \frac{\partial^2 \ell}{\partial \lambda_1 \partial \lambda_4} = \frac{\partial^2 \ell}{\partial \lambda_4 \partial \lambda_1} = -\xi_1'(\beta) \\
&\frac{\partial^2 \ell}{\partial \lambda_2 \partial \lambda_4} = \frac{\partial^2 \ell}{\partial \lambda_4 \partial \lambda_2} = -\xi_2'(\beta), \qquad \frac{\partial^2 \ell}{\partial \lambda_3 \partial \lambda_4} = \frac{\partial^2 \ell}{\partial \lambda_4 \partial \lambda_3} = -\xi_3'(\beta) \\
&\frac{\partial^2 \ell}{\partial \lambda_1 \partial \lambda_2} = \frac{\partial^2 \ell}{\partial \lambda_2 \partial \lambda_1} = \frac{\partial^2 \ell}{\partial \lambda_1 \partial \lambda_3} = \frac{\partial^2 \ell}{\partial \lambda_3 \partial \lambda_1} = \frac{\partial^2 \ell}{\partial \lambda_2 \partial \lambda_3} = \frac{\partial^2 \ell}{\partial \lambda_3 \partial \lambda_2} = 0
\end{aligned}$$

where

$$\xi_h'(\beta) = \sum_{i=1}^{r} \Big[\delta_i(h) + s_i(h)\Big] w_i^{\beta} \log(w_i) \quad \text{and} \quad \xi_h'(\beta) = \sum_{i=1}^{r} \Big[\delta_i(h) + s_i(h)\Big] w_i^{\beta} \log^2(w_i).$$

Therefore, according to the asymptotic normality of the MLEs, $\tilde{\Omega}$ is asymptotically normally distributed with mean Ω and variance

$$\sigma_{\tilde{\Omega}}^2 = \sum_{j=1}^{3} \sum_{i=1}^{3} \frac{\partial \Omega}{\partial \theta_i} \frac{\partial \Omega}{\partial \theta_j} I_{ij}^{-1} \tag{4.8}$$

where I_{ij}^{-1} is the (i, j) th element of the inverse of the $I(\theta)$. Then,

$$\begin{aligned}
\sigma_{\tilde{\Omega}}^2 = &\left(\frac{\partial \Omega}{\partial \lambda_1}\right)^2 I_{11}^{-1} + \left(\frac{\partial \Omega}{\partial \lambda_2}\right)^2 I_{22}^{-1} + \left(\frac{\partial \Omega}{\partial \lambda_3}\right)^2 I_{33}^{-1} + 2\left(\frac{\partial \Omega}{\partial \lambda_1}\right)\left(\frac{\partial \Omega}{\partial \lambda_2}\right) I_{12}^{-1} \\
&+ 2\left(\frac{\partial \Omega}{\partial \lambda_1}\right)\left(\frac{\partial \Omega}{\partial \lambda_3}\right) I_{13}^{-1} + 2\left(\frac{\partial \Omega}{\partial \lambda_2}\right)\left(\frac{\partial \Omega}{\partial \lambda_3}\right) I_{23}^{-1}
\end{aligned} \tag{4.9}$$

where

$$\frac{\partial \Omega}{\partial \lambda_1} = -\frac{\lambda_2 \lambda_3 \big(2\lambda_1 + 2\lambda_2 + \lambda_3\big)}{(\lambda_1 + \lambda_2)^2 (\lambda_1 + \lambda_2 + \lambda_3)^2},$$

$$\frac{\partial \Omega}{\partial \lambda_2} = \frac{\lambda_3 \big(\lambda_1^2 + \lambda_1 \lambda_3 - \lambda_2^2\big)}{(\lambda_1 + \lambda_2)^2 (\lambda_1 + \lambda_2 + \lambda_3)^2},$$

$$\frac{\partial \Omega}{\partial \lambda_3} = \frac{\lambda_2 \big(\lambda_1 + \lambda_2\big)^2}{(\lambda_1 + \lambda_2)^2 (\lambda_1 + \lambda_2 + \lambda_3)^2}.$$

As a result, we construct the $100(1-\gamma)\%$ ACI for Ω by

$$\Omega \in \big(\tilde{\Omega} \pm z_{\gamma/2} \sigma_{\tilde{\Omega}}\big)$$

where z_γ is 100γth percentile of standard normal distribution $N(0, 1)$. However, approximate confidence intervals for the reliability parameter Ω are not always constrained within the unit interval $(0, 1)$. In some instances, they may produce an upper bound greater than 1 or a lower bound less than 0. To address this, the logit transformation $h(\Omega) = \log\left(\frac{\Omega}{1-\Omega}\right)$ can be applied. Using the delta method (see, e.g., Xu and Long (2005)), the asymptotic distribution of the maximum likelihood estimator of Ω, denoted by $\tilde{\Omega}$, satisfies the following convergence:

$$\sqrt{N}\left\{h(\tilde{\Omega}) - h(\Omega)\right\} \to \mathbf{N}\left(0, \frac{\partial^T(\theta) I^{-1}(\theta) \partial(\theta)}{\Omega^2 (1-\Omega)^2}\right).$$

Consequently, the $100(1-\gamma)\%$ asymptotic confidence interval for $h(\Omega)$ is given by

$$\left(h(\tilde{\Omega}) - z_{\frac{\gamma}{2}} \frac{\sqrt{\tilde{\sigma}^2_{\Omega}}}{\tilde{\Omega}(1-\tilde{\Omega})}; h(\tilde{\Omega}) + z_{\frac{\gamma}{2}} \frac{\sqrt{\tilde{\sigma}^2_{\Omega}}}{\tilde{\Omega}(1-\tilde{\Omega})}\right) \equiv (L, U).$$

Therefore, the $100(1-\gamma)\%$ asymptotic confidence interval for Ω is constructed as follows

$$\left(\frac{e^L}{1+e^L}; \frac{e^U}{1+e^U}\right). \tag{4.10}$$

5. ORDER-RESTRICTED BAYESIAN INFERENCE

An order-restricted Bayesian inference procedure can be employed for the estimation of Ω. We first assume a dependent prior distribution for the parameter vector $\theta = (\lambda_1, \lambda_2, \lambda_3, \beta)$ following the approach of Peña and Gupta (1990). To this end, we define $\lambda_0 = \lambda_1 + \lambda_2 + \lambda_3$ and assume that λ_0 follows a Gamma prior distribution $GA(a_0, b_0)$ with the probability density function (pdf)

$$\pi_0(\lambda_0 \mid a_0, b_0) = \frac{b_0^{a_0}}{\Gamma(a_0)} \lambda_0^{a_0-1} e^{-b_0\lambda_0}$$

where $a_0 > 0, b_0 > 0, \lambda_0 > 0$. For a given λ_0, $\left(\lambda_1/\lambda_0, \lambda_2/\lambda_0, \lambda_3/\lambda_0 \middle| \lambda_0, c_1, c_2, c_3\right)$ follows a Dirichlet prior, $\pi_1(\cdot|c_1, c_2, c_3)$ and it is given by

$$\pi_1\left(\frac{\lambda_1}{\lambda_0}, \frac{\lambda_2}{\lambda_0}, \frac{\lambda_3}{\lambda_0} \middle| \lambda_0, c_1, c_2, c_0\right) = \frac{\Gamma(c_0)}{\Gamma(c_1)\Gamma(c_2)\Gamma(c_3)} \prod_{i=0}^{3} \left(\frac{\lambda_i}{\lambda_0}\right)^{c_i}, \tag{5.1}$$

where $\lambda_i > 0$ for $i = 1, 2, 3$ and $c_i > 0$ are the parameters of the Dirichlet distribution with $c_0 = c_1 + c_2 + c_3$. Then, the joint prior of $(\lambda_1, \lambda_2, \lambda_3)$, namely $\pi(\lambda_1, \lambda_2, \lambda_3 \mid a_0, b_0, c_1, c_2, c_3)$, can be expressed as

$$\pi_1(\lambda_1, \lambda_2, \lambda_3) = \frac{\Gamma(c_0)}{\Gamma(a_0)} \prod_{i=1}^{3} \frac{b_0^{a_0}}{\Gamma(c_i)} \lambda_i^{c_i-1} e^{-b_0\lambda_i} (\lambda_0 b_0)^{a_0-c_0} \tag{5.2}$$

In most cases, this prior is dependent; however, if $a_0 = c_0$, then the parameters λ_i $(i = 1, 2, 3)$ are assumed to be independent gamma priors with parameters b_0 and c_i, respectively. Additionally, we propose a gamma prior for the common shape parameter β, with its probability density function given as follows

$$\beta \sim GA(a_1, b_1), \quad \pi(\beta) \propto \beta^{a_1-1} \exp\{-\beta b_1\}$$

where a_1 and b_1, are the non-negative hyper-parameters. Under the order restriction on the scale parameters as $\lambda_1 < \lambda_2 < \lambda_3$. Then, we obtain the the joint prior of $(\lambda_1, \lambda_2, \lambda_3)$ as

$$\pi_{or}(\lambda_1, \lambda_2, \lambda_3) = \frac{\Gamma(c_0)}{\Gamma(a_0)\Gamma(c_1)\Gamma(c_2)\Gamma(c_3)} b_0^{a_0} e^{-b_0\lambda_0} \lambda_0^{a_0-c_0} \sum_{\mathcal{P}} \left(\lambda_{i_1}^{c_1-1} \lambda_{i_2}^{c_2-1} \lambda_{i_3}^{c_3-1}\right) \tag{5.3}$$

where $0 < \lambda_1 < \lambda_2 < \lambda_3 < \infty$. Here, $\pi_{or}(\cdots)$ is order restricted gamma-Dirichlet (OGD) prior and $\mathcal{P}$ denotes the set of all the 3! permutations of I. Hence, the joint posterior density function of $\beta > 0$ and $0 < \lambda_1 < \lambda_2 < \lambda_3 < \infty$ can be written as

$$\begin{aligned} L_{or}(\beta, \boldsymbol{\lambda}, \boldsymbol{W}) \propto \lambda_0^{a_0-c_0} \prod_{h=1}^{3} \lambda_h^{M_r(h)} \exp\Big\{ -\lambda_h \left(b_0 + \xi_h(\beta)\right) \Big\} \beta^{a_1+r-1} \\ \times \exp\left\{ -\beta\Big(b_1 - \sum_{i=1}^{r}\sum_{h=1}^{3} \delta_i(h) \log(w_i)\Big)\right\} \sum_{(i_1,i_2,i_3)\in\mathcal{P}} \left(\lambda_{i_1}^{c_1-1}\lambda_{i_2}^{c_2-1}\lambda_{i_3}^{c_3-1}\right). \end{aligned} \tag{5.4}$$

We can obtain the following joint posterior density function after modifying Equation 5.2

$$\begin{aligned} L_{or}(\beta, \boldsymbol{\lambda}, \boldsymbol{W}) \propto \lambda_0^{a_0-c_0} \prod_{h=1}^{3} \lambda_h^{M_r(h)-J} \exp\Big\{ -\lambda_h \left(\xi_h(\beta) - S(\beta)\right) \Big\} \\ \times \exp\Big\{ -\lambda_0(b_0 + S(\beta))\Big\} \exp\left\{ -\beta\Big(b_1 - \sum_{i=1}^{r}\sum_{h=1}^{3} \delta_i(h) \log(w_i)\Big)\right\} \\ \times \beta^{a_1+r-1} \sum_{(i_1,i_2,i_3)\in\mathcal{P}} \left(\lambda_{i_1}^{c_1+J-1}\lambda_{i_2}^{c_2+J-1}\lambda_{i_3}^{c_3+J-1}\right), \end{aligned} \tag{5.5}$$

where $J = \min\left(M_r(h)\right)$ and $S(\beta) = \min\left(\xi_h(\beta)\right)$ for $h = 1, 2, 3$. Then, we can obtain the posterior densities of the parameters as

$$\pi(\lambda, \beta | data) \propto \pi_1(\lambda|\beta)\pi_2(\beta|\lambda)\mathcal{H}(\lambda, \beta)$$

where

$$\pi_1(\lambda|\beta) = OGD(a_0 + 3J, b_0 + S(\beta), c_1 + J, c_2 + J, c_3 + J)$$

$$\pi_2(\beta|\lambda) = \beta^{a_1+r-1} \exp\left\{ -\beta\Big(b_1 - \sum_{i=1}^{r}\sum_{h=1}^{3} \delta_i(h) \log(w_i)\Big)\right\}$$

and

$$\mathcal{H}(\lambda, \beta) = \frac{\prod_{h=1}^{3} \lambda_h^{M_r(h)-J} \exp\Big\{ -\lambda_h \left(\xi_h(\beta) - S(\beta)\right) \Big\}}{\Big[b_0 + S(\beta)\Big]^{a_0+3J}}.$$

In the case where $b_1 > \sum_{i=1}^{r}\sum_{h=1}^{3} \delta_i(h) \log(w_i)$, the conditional density $\pi_2(\beta \mid \lambda)$ is a proper and well-known log-concave density. Moreover, sampling from the conditional distribution $\pi_1(\lambda \mid \beta)$ can be conveniently performed using the algorithm proposed by Pal et al. (2021). Therefore, we implement the following importance sampling algorithm

Step 1: Generate λ_i from $\pi_1(\lambda|\beta, \mathbf{data})$ and β from $\pi_2(\beta|\lambda, \mathbf{data})$.

Step 2: Find $\mathcal{H} = \mathcal{H}(\lambda, \beta)$.

Step 3: Calculate Ω using Equation 3.2.

Step 4: Generate M estimates of Ω_j and $\mathcal{H}_j$, for $j = 1, 2, \ldots, M$ by repeating steps 1-3 M times.

Thus, an approximate Bayes estimate of Ω under SELF can be obtained as

$$\tilde{\Omega}_B = \frac{\frac{1}{M}\sum_{j=1}^{M}\Omega_j\mathcal{H}_j}{\frac{1}{M}\sum_{i=1}^{M}\mathcal{H}_j}. \tag{5.6}$$

Then, the credible interval for Ω can be obtained following the approach of Chen and Shao (1999). Assume that $\varphi_j = \mathcal{H}(\beta_j, \lambda_{ij})$, where $\{\beta_j\}_{j=1}^{M}$ and $\{\lambda_{ij}\}_{j=1}^{M}$ are posterior samples generated respectively from the conditional distributions $\pi_2(\beta \mid \lambda)$ and $\pi_1(\lambda \mid \beta)$ for β and λ_i. Define

$$\omega_j = \frac{\mathcal{H}(\beta_j, \lambda_{ij})}{\sum_{j=1}^{M}\mathcal{H}(\beta_j, \lambda_{ij})}.$$

To construct a $100(1-\gamma)\%$ credible interval for Ω, arrange the samples $\{\Omega_j\}_{j=1}^{M}$ in ascending order to obtain $(\Omega_{(1)}, \Omega_{(2)}, \ldots, \Omega_{(M)})$. Arrange the corresponding values $\{\omega_j\}_{j=1}^{M}$ accordingly to get $(\omega_{(1)}, \omega_{(2)}, \ldots, \omega_{(M)})$. Note that the sequence $(\omega_{(1)}, \omega_{(2)}, \ldots, \omega_{(M)})$ may not be ordered. A $100(1-\gamma)\%$ credible interval can be obtained as $(\Omega_{j_{L_0}}, \Omega_{j_{U_0}})$ where L_0 and U_0 such that,

$$L_0 < U_0, \quad L_0, U_0 \in \{1, \ldots, M\} \quad \text{and} \quad \sum_{i=L_0}^{U_0}\omega_i \le 1-\gamma < \sum_{i=L_o}^{U_0+1}\omega_i. \tag{5.7}$$

The $100(1-\gamma)\%$ HPD credible interval can be obtained as $(\Omega_{j^*_{L_0}}, \Omega_{j^*_{U_0}})$, such that $\Omega_{j^*_{U_0}} - \Omega_{j^*_{L_0}} \le \Omega_{j_{U_0}} - \Omega_{j_{L_0}}$ and the indices L_0^* and U_0^* satisfy Equation (5.7) for all L_0 and U_0 that satisfy Equation (5.7).

6. UNRESTRICTED INFERENCE

In this section, inferences regarding Ω are presented under the unrestricted parameter setting. Consequently, the estimators of the proposed model can be compared with the classical inference procedures. In this purpose, the classical MLEs of the parameters can be obtained by taking partial derivatives of the log-likelihood function (4.3) with respect to the parameters and equating them to zero as given in the following

$$\hat{\lambda}_h = \frac{M_r(h)}{\xi_h(\hat{\beta})} \tag{6.1}$$

where $\hat{\beta}$ is the MLE of β and it can be obtained with the solution of the following nonlinear equation

$$\frac{r}{\beta} - \sum_{h=1}^{3}\frac{M_r(h)\xi_h(\beta)'}{\xi_h(\beta)} + \sum_{i=1}^{r}\sum_{h=1}^{3}\delta_i(h)\log(w_i) = 0.$$

Then, MLE of Ω, that is $\hat{\Omega}$, can be obtained using the invariance property of the MLE as replacing the MLEs of the parameters in Equation (3.2). On the other hand, Bayesian inferences can be obtained considering independent gamma priors such that $\lambda_h \sim GA(a_h, b_h)$, and $\beta \sim GA(a_4, b_4)$ $(h = 1, 2, 3)$. Then, the conditional posterior density functions of the parameters can be obtained easily as in the following

$$\pi(\lambda_h \mid \beta) \propto GA\big\{M_r(h) + a_h, \xi_h(\beta) + b_h\big\}, \quad h = 1, 2, 3.$$

and

$$\pi(\beta \mid \lambda_h) \propto \beta^{r+a_4-1} \exp\left\{-\beta\left[b_4 - \sum_{i=1}^{r}\sum_{h=1}^{3} \delta_i(h) \log(w_i)\right] - \sum_{h=1}^{3} \xi_h(\beta)\lambda_h\right\}.$$

It is clearly seen from that the samples for λ_h can be easily obtained by using gamma densities. However, the density from $\pi(\beta \mid \lambda_h)$ can not be reduced analytically to well known distributions and therefore it is not possible to generate sample directly by standard methods. Hence, the Metropolis-Hasting (M-H) algorithm with normal proposal distribution as suggested by Tierney (1994) can be used for the Bayesian estimation of β. Thus, the Bayesian estimation of Ω, $\hat{\Omega}_B$, can be obtained Gibbs sampling with M-H method.

7. SIMULATION STUDIES

This section aims to demonstrate the performance of the estimation methods employed in this study. Both order-restricted and unrestricted estimates of Ω have been computed to compare the proposed method with existing procedures. To this end, various simulation schemes are conducted. Initially, three different censoring schemes (CS) are considered, as follows:

- $R = (1_{(N-r)}, 0_{(2r-N)}) \longrightarrow CS - I$
- $R = (0_{(2r-N)}, 1_{(N-r)}) \longrightarrow CS - II$
- $R = (0_{(1:(\tilde{m}-1))}, N-r, 0_{(\tilde{m}+1:r)}) \longrightarrow CS - III$

where $\tilde{m}$ denotes the sample median. The various total sample size are determined as $N = (24, 36, 54)$. We divide each sample size into three sets of units as equal and unequal, respectively. In each sample, we determine the censoring rates %30 and %15. The samples are generated with the actual parameter values $\lambda_1 = 1.25$, $\lambda_2 = 2.00$, $\lambda_3 = 3.50$ and $\beta = 2.00$ as $\Omega \approx 0.32$. In Bayesian computations, we consider both informative and non-informative hyperparameter values. The informative hyperparameters are determined to match the expected value of the prior distribution. Specifically, we set $(a_0, b_0) = (13.5, 2.00)$, $(a_1, b_1) = (2.00, 1.00)$, and $(c_1, c_2, c_3) = (2.50, 4.00, 7.00)$. For non-informative hyperparameters, we use $a_0 = b_0 = a_1 = b_1 = 0.001$ and $c_1 = c_2 = c_3 = 1$. We use the R (R Core Team (2023)) software for the computations. The simulations are repeated 1000 times, employing Markov chains with a sample size of 3500 in each iteration. A burn-in number of $K = 500$ is used in the Markov chains, with a thinning procedure that involves selecting every third variate as uncorrelated samples. This process enables the derivation of absolute biases (AB) and mean squared errors (MSE) for the point estimations. In addition, the average lengths (AL) of the approximate confidence intervals are computed along with their corresponding coverage probabilities (CP). The bias values and MSEs are reported in Tables 2- 6. In addition, the ALs and CPs of the approximate confidence intervals are presented in Tables 3- 7. The following results are observed.

- The results indicate that, in all cases examined, the bias and MSE values were acceptably small.

- Consistent results were obtained from the simulation studies, regardless of the censoring scheme or sample size.
- Results indicate that the Bayesian method outperforms maximum likelihood estimation. The performance gap between the two methods narrowed with increasing sample sizes.
- Order-restricted Bayesian estimators yield the smallest MSE values. Especially in small samples, the order-restricted Bayesian method outperforms the classical inference method. As the sample size increases, although the difference between the order-restricted MLE and the classical MLE diminishes, the order-restricted Bayesian estimator maintains its superiority across all sample sizes.
- Within Bayesian inference, informative priors contribute to improved estimation accuracy in both classical and order-restricted contexts. The superiority of Bayesian estimators employing informative priors becomes particularly pronounced in small sample scenarios. In contrast, the use of non-informative priors often results in poorer performance, especially under order restrictions. This finding is theoretically consistent, as Bayesian approaches fundamentally depend on incorporating natural parameter constraints into the estimation process.
- When the prior assumptions are misspecified, the resulting estimations may exhibit poor performance, particularly in large sample settings. Although this issue was not pronounced in our study, the ALs of the HPD intervals yielded unsatisfactory results under misspecified prior assumptions, especially in the case of non-informative priors.
- In all cases, the estimators exhibit very small bias values. It is clearly seen that the estimation methods produce at most %3 bias from the true parameter value. Although the order-restricted methods demonstrate superiority in terms of MSE, the classical method provides smaller bias values.
- The average lengths (ALs) of the approximate confidence intervals decrease with larger sample sizes. Their coverage probabilities (CPs) are mostly very close to the nominal value of 0.95 in most cases. However, in some schemes, the ACI intervals have smaller CP values than expected. The Bayesian HPD credible intervals exhibit shorter lengths, especially in the order-restricted case; however, their CPs are slightly higher than the nominal level of 0.95.

In addition, graphical techniques are employed to assess the convergence of the Markov chains generated in the simulation studies. First, a trace plot is used, which depicts the number of iterations t versus the corresponding values of $\Omega(t)$. Subsequently, the posterior density plot of Ω is drawn. Furthermore, a running mean diagram is utilized to display the mean of the samples up to iteration t; the accuracy of the ergodic average depends on the autocorrelation of the chain. For each sample size in the simulation schemes, convergence diagnostics of the Markov chains are evaluated, and all relevant plots are provided in Figure 3. The results indicate that the convergence of the Markov chains is satisfactory based on these evaluations. The trace plots fluctuate around their central values with similar variation, and the ranges tend to decrease as the sample size increases. Symmetric and unimodal posterior density plots are observed, with the bandwidths of the densities decreasing in parallel with increasing sample sizes. Additionally, the running mean and the autocorrelation function (ACF) plots confirm that the chains have reached stationarity. These observations provide

			Classical Inference						Order-Restricted Inference					
			Bias			MSE			Bias			MSE		
CS	n_i	r	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$
I	(8,8,8)	17	0.0093	0.0072	0.0108	0.0159	0.0059	0.0057	0.0301	0.0921	0.0392	0.0129	0.0106	0.0029
	(8,8,8)	20	0.0063	0.0099	0.0129	0.0121	0.0051	0.0050	0.0224	0.0933	0.0434	0.0100	0.0101	0.0030
	(10,6,8)	17	0.0038	0.0135	0.0167	0.0126	0.0053	0.0051	0.0259	0.0901	0.0374	0.0102	0.0101	0.0028
	(10,6,8)	20	0.0026	0.0142	0.0159	0.0115	0.0050	0.0049	0.0210	0.0867	0.0400	0.0096	0.0092	0.0029
II	(8,8,8)	17	0.0331	0.0260	0.0271	0.0147	0.0076	0.0074	0.0087	0.0799	0.0318	0.0104	0.0110	0.0039
	(8,8,8)	20	0.0253	0.0229	0.0242	0.0127	0.0068	0.0068	0.0055	0.0799	0.0353	0.0095	0.0097	0.0036
	(10,6,8)	17	0.0345	0.0292	0.0304	0.0144	0.0076	0.0075	0.0065	0.0821	0.0311	0.0100	0.0108	0.0037
	(10,6,8)	20	0.0294	0.0282	0.0292	0.0111	0.0060	0.0059	0.0080	0.0817	0.0354	0.0082	0.0092	0.0032
III	(8,8,8)	17	0.0023	0.0090	0.0111	0.0132	0.0054	0.0052	0.0231	0.0832	0.0341	0.0103	0.0096	0.0028
	(8,8,8)	20	0.0011	0.0118	0.0140	0.0120	0.0054	0.0053	0.0184	0.0893	0.0418	0.0098	0.0099	0.0032
	(10,6,8)	17	0.0008	0.0150	0.0173	0.0130	0.0057	0.0055	0.0226	0.0847	0.0344	0.0105	0.0097	0.0029
	(10,6,8)	20	0.0077	0.0105	0.0125	0.0110	0.0049	0.0048	0.0241	0.0816	0.0362	0.0092	0.0085	0.0027

Table 2: Bias and MSEs of the estimations in the case of $N = 24$.

			Classical Inference						Order-Restricted Inference					
			AL			CP			AL			CP		
CS	n_i	r	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$
I	(8,8,8)	17	0.4419	0.3311	0.3305	0.96	0.97	0.98	0.4106	0.3896	0.3580	0.96	0.95	0.99
	(8,8,8)	20	0.3957	0.3094	0.3091	0.96	0.97	0.97	0.3754	0.3500	0.3285	0.96	0.95	0.99
	(10,6,8)	17	0.4217	0.3184	0.3184	0.97	0.98	0.98	0.3989	0.3840	0.3436	0.98	0.95	0.99
	(10,6,8)	20	0.3787	0.2972	0.2965	0.95	0.97	0.97	0.3616	0.3518	0.3229	0.96	0.96	0.98
II	(8,8,8	17	0.4092	0.3265	0.3265	0.96	0.94	0.94	0.3883	0.3798	0.3511	0.97	0.91	0.97
	(8,8,8)	20	0.3718	0.3059	0.3056	0.96	0.94	0.95	0.3568	0.3471	0.3208	0.96	0.93	0.97
	(10,6,8)	17	0.3907	0.3139	0.3140	0.95	0.93	0.93	0.3745	0.3658	0.3341	0.97	0.90	0.97
	(10,6,8)	20	0.3549	0.2939	0.2936	0.95	0.96	0.96	0.3428	0.3426	0.3138	0.97	0.94	0.97
III	(8,8,8)	17	0.4460	0.3344	0.3342	0.97	0.98	0.98	0.4113	0.3921	0.3627	0.98	0.96	0.99
	(8,8,8)	20	0.3953	0.3108	0.3105	0.96	0.96	0.96	0.3754	0.3514	0.3274	0.96	0.95	0.99
	(10,6,8)	17	0.4158	0.3192	0.3188	0.96	0.97	0.97	0.3953	0.3787	0.3437	0.96	0.94	0.98
	(10,6,8)	20	0.3768	0.2990	0.2983	0.96	0.97	0.97	0.3630	0.3538	0.3226	0.96	0.96	0.99

Table 3: ALs and CP values of the approximate CIs in the case of $N = 24$.

evidence of the convergence of the Markov chains. The `mcmcplot` function from the **mcmcplots** package Curtis (2018) in the `R` software R Core Team (2023) was used to obtain these plots.

8. CONCLUSIONS

In this study, it is demonstrated that a stress-strength model incorporating both lower and upper stresses can be effectively analyzed under simultaneous censoring schemes applied to all components. Imposing order restrictions on the scale parameters to achieve higher reliability values is another critical aspect of this problem. It is shown that such problems can be addressed using the GIR framework introduced by Barlow et al. (1972). In the

			Classical Inference						Order-Restricted Inference					
			Bias			MSE			Bias			MSE		
CS	n_i	r	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$
I	(12,12,12)	25	0.0008	0.0107	0.0143	0.0094	0.0047	0.0044	0.0126	0.0897	0.0483	0.0077	0.0096	0.0035
	(12,12,12)	31	0.0034	0.0084	0.0115	0.0077	0.0043	0.0041	0.0121	0.0891	0.0522	0.0065	0.0091	0.0037
	(10,12,14)	25	0.0030	0.0060	0.0078	0.0098	0.0049	0.0048	0.0153	0.0812	0.0451	0.0080	0.0087	0.0034
	(10,12,14)	31	0.0089	0.0020	0.0032	0.0075	0.0042	0.0042	0.0161	0.0782	0.0464	0.0065	0.0075	0.0032
II	(12,12,12)	25	0.0253	0.0223	0.0233	0.0089	0.0056	0.0055	0.0113	0.0702	0.0340	0.0067	0.0080	0.0034
	(12,12,12)	31	0.0214	0.0217	0.0229	0.0079	0.0051	0.0051	0.0103	0.0756	0.0422	0.0063	0.0077	0.0034
	(10,12,14)	25	0.0273	0.0231	0.0235	0.0103	0.0064	0.0063	0.0113	0.0703	0.0361	0.0076	0.0083	0.0039
	(10,12,14)	31	0.0254	0.0232	0.0250	0.0075	0.0050	0.0048	0.0140	0.0698	0.0406	0.0058	0.0072	0.0035
III	(12,12,12)	25	0.0012	0.0105	0.0146	0.0091	0.0047	0.0046	0.0105	0.0830	0.0440	0.0075	0.0091	0.0035
	(12,12,12)	31	0.0027	0.0088	0.0103	0.0075	0.0042	0.0041	0.0109	0.0838	0.0495	0.0064	0.0084	0.0036
	(10,12,14)	25	0.0018	0.0061	0.0079	0.0092	0.0047	0.0046	0.0151	0.0761	0.0416	0.0073	0.0082	0.0034
	(10,12,14)	31	0.0016	0.0060	0.0079	0.0079	0.0045	0.0045	0.0107	0.0749	0.0453	0.0066	0.0073	0.0033

Table 4: Bias and MSEs of the estimations in the case of $N = 36$.

			Classical Inference						Order-Restricted Inference					
			AL			CP			AL			CP		
CS	n_i	r	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^I$
I	(12,12,12)	25	0.3486	0.2882	0.2874	0.96	0.96	0.97	0.3381	0.3178	0.3018	0.96	0.96	0.99
	(12,12,12)	31	0.3094	0.2633	0.2631	0.95	0.96	0.97	0.3023	0.2636	0.2568	0.95	0.93	0.99
	(10,12,14)	25	0.3632	0.2971	0.2969	0.96	0.97	0.97	0.3460	0.3308	0.3164	0.96	0.97	0.99
	(10,12,14)	31	0.3162	0.2730	0.2731	0.95	0.97	0.97	0.3104	0.2846	0.2782	0.96	0.97	0.99
II	(12,12,12)	25	0.3335	0.2881	0.2880	0.97	0.95	0.95	0.3256	0.3175	0.3050	0.98	0.96	0.98
	(12,12,12)	31	0.2949	0.2613	0.2611	0.94	0.95	0.95	0.2913	0.2751	0.2624	0.96	0.97	0.98
	(10,12,14)	25	0.3531	0.2981	0.2982	0.96	0.94	0.95	0.3380	0.3243	0.3177	0.97	0.95	0.97
	(10,12,14)	31	0.3063	0.2694	0.2692	0.96	0.95	0.96	0.2991	0.2904	0.2846	0.97	0.97	0.98
III	(12,12,12)	25	0.3482	0.2901	0.2890	0.96	0.96	0.96	0.3377	0.3207	0.3041	0.97	0.96	0.98
	(12,12,12)	31	0.3067	0.2644	0.2641	0.96	0.96	0.97	0.3024	0.2720	0.2617	0.96	0.96	0.99
	(10,12,14)	25	0.3692	0.3020	0.3019	0.97	0.97	0.98	0.3513	0.3288	0.3187	0.97	0.97	0.99
	(10,12,14)	31	0.3190	0.2727	0.2725	0.95	0.95	0.95	0.3094	0.2918	0.2839	0.95	0.98	0.99

Table 5: ALs and CP values of the approximate CIs in the case of $N = 36$.

Bayesian inference framework, ordered gamma-Dirichlet (OGD) priors are adopted, following the approach of Peña and Gupta (1990). Both inference methods exhibited consistent and satisfactory performance. Regarding bias and mean squared error (MSE) values, the estimations were notably small. Additionally, approximate confidence intervals derived from both models demonstrated short lengths and favorable coverage probabilities. Bayesian credible intervals tended to exhibit slightly higher coverage probabilities than nominal levels. The Markov chains employed in the Bayesian method exhibited satisfactory convergence behavior. As a recommendation for future research, the extended form of this reliability model, namely the estimation of $P(X_1 < X_2 < \cdots < x_k)$, could be investigated under a joint censoring scheme involving k populations. Such an extension can be performed for Weibull-distributed populations.

			Classical Inference						Order-Restricted Inference					
			Bias			MSE			Bias			MSE		
CS	n_i	r	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$
I	(18,18,18)	43	0.0024	0.0068	0.0088	0.0047	0.0031	0.0030	0.0066	0.0884	0.0586	0.0042	0.0087	0.0042
	(18,18,18)	49	0.0047	0.0029	0.0059	0.0042	0.0029	0.0028	0.0080	0.0875	0.0602	0.0039	0.0087	0.0043
	(16,18,20)	43	0.0013	0.0060	0.0088	0.0053	0.0035	0.0033	0.0069	0.0822	0.0561	0.0045	0.0078	0.0040
	(16,18,20)	49	0.0042	0.0028	0.0054	0.0042	0.0029	0.0029	0.0075	0.0805	0.0567	0.0038	0.0073	0.0039
II	(18,18,18)	43	0.0152	0.0157	0.0168	0.0049	0.0037	0.0036	0.0100	0.0626	0.0381	0.0042	0.0056	0.0029
	(18,18,18)	49	0.0162	0.0178	0.0189	0.0050	0.0038	0.0037	0.0109	0.0761	0.0526	0.0042	0.0069	0.0038
	(16,18,20)	43	0.0137	0.0138	0.0152	0.0052	0.0038	0.0038	0.0079	0.0597	0.0376	0.0044	0.0055	0.0030
	(16,18,20)	49	0.0217	0.0219	0.0231	0.0048	0.0036	0.0036	0.0169	0.0687	0.0474	0.0041	0.0061	0.0035
III	(18,18,18)	43	0.0023	0.0060	0.0081	0.0049	0.0033	0.0032	0.0068	0.0816	0.0535	0.0043	0.0078	0.0038
	(18,18,18)	49	0.0047	0.0035	0.0057	0.0043	0.0029	0.0029	0.0075	0.0839	0.0570	0.0039	0.0079	0.0041
	(16,18,20)	43	0.0024	0.0043	0.0066	0.0055	0.0037	0.0036	0.0067	0.0756	0.0512	0.0049	0.0070	0.0037
	(16,18,20)	49	0.0026	0.0037	0.0059	0.0045	0.0031	0.0031	0.0056	0.0775	0.0554	0.0041	0.0071	0.0039

Table 6: Bias and MSEs of the estimations in the case of $N = 54$.

			Classical Inference						Order-Restricted Inference					
			AL			CP			AL			CP		
CS	n_i	r	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$	$\hat{\Omega}$	$\hat{\Omega}_B^{NI}$	$\hat{\Omega}_B^{I}$
I	(18,18,18)	43	0.2592	0.2322	0.2320	0.96	0.96	0.96	0.2582	0.2142	0.2124	0.97	0.86	0.96
	(18,18,18)	49	0.2417	0.2191	0.2185	0.95	0.96	0.96	0.2411	0.1818	0.1836	0.95	0.68	0.89
	(16,18,20)	43	0.2641	0.2360	0.2355	0.95	0.96	0.96	0.2609	0.2302	0.2265	0.95	0.93	0.98
	(16,18,20)	49	0.2462	0.2237	0.2231	0.95	0.96	0.96	0.2450	0.2034	0.2013	0.96	0.86	0.96
II	(18,18,18)	43	0.2509	0.2317	0.2313	0.95	0.95	0.95	0.2500	0.2309	0.2254	0.96	0.99	0.99
	(18,18,18)	49	0.2351	0.2170	0.2168	0.94	0.93	0.93	0.2341	0.1962	0.1927	0.95	0.85	0.96
	(16,18,20)	43	0.2591	0.2382	0.2377	0.95	0.95	0.95	0.2569	0.2348	0.2303	0.96	0.98	0.99
	(16,18,20)	49	0.2366	0.2194	0.2190	0.94	0.93	0.93	0.2351	0.2161	0.2127	0.95	0.97	0.99
III	(18,18,18)	43	0.2592	0.2326	0.2321	0.95	0.96	0.96	0.2579	0.2247	0.2205	0.95	0.93	0.98
	(18,18,18)	49	0.2424	0.2201	0.2195	0.95	0.96	0.96	0.2418	0.1902	0.1890	0.95	0.77	0.94
	(16,18,20)	43	0.2639	0.2378	0.2372	0.95	0.95	0.95	0.2619	0.2364	0.2336	0.95	0.98	0.99
	(16,18,20)	49	0.2457	0.2240	0.2234	0.94	0.95	0.96	0.2447	0.2089	0.2080	0.95	0.93	0.98

Table 7: ALs and CP values of the approximate CIs in the case of $N = 54$.

ACKNOWLEDGMENTS

The author sincerely thank the anonymous reviewers and associate editor for their constructive feedback, which greatly improved the manuscript. This work was supported by a grant from the Office of Scientific Research Projects at Kırıkkale University, Türkiye (2024/070).

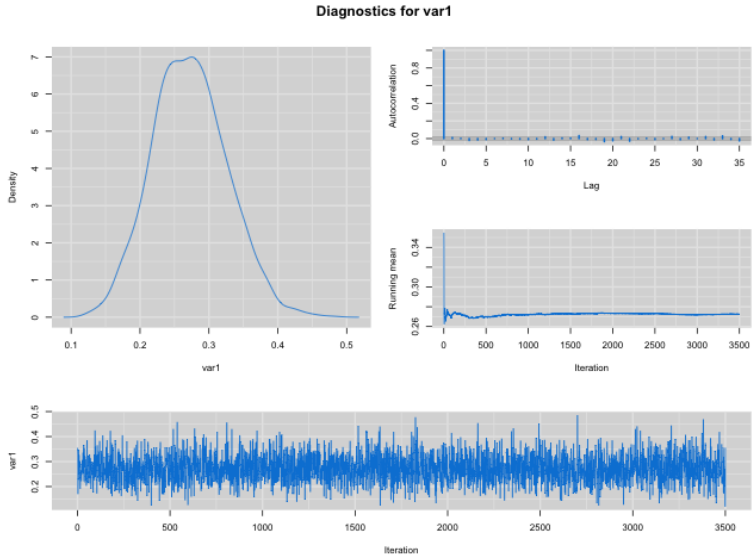


Figure 3: Plots for diagnostics for convergence of the Markov chains for the case $N = 24(16, 18, 20)$, $r = 20$, $CS - III$. (Density plot (top left), ACF (top right), running mean (center right) and trace plots (below)).

REFERENCES

Akbari Bargoshadi, S., Bevrani, H., and Arabi Belaghi, R. (2024). Statistical inference under joint type$-II$ censored data from two Burr$-XII$ populations. *Communications in Statistics-Simulation and Computation*, 53(12):6305–6325.

Almetwally, E. M., Alotaibi, R., Mutairi, A. A., Park, C., and Rezk, H. (2022). Optimal plan of multi-stress–strength reliability bayesian and non-bayesian methods for the alpha power exponential model using progressive first failure. *Symmetry*, 14(7):1306.

Alotaibi, R., Almetwally, E. M., Ghosh, I., and Rezk, H. (2024). The optimal multi-stress–strength reliability technique for the progressive first failure in the length-bias exponential model using bayesian and non-bayesian methods. *Journal of Statistical Computation and Simulation*, 94(11):2436–2461.

Balakrishnan, N., Su, F., and Liu, K.-Y. (2015). Exact likelihood inference for k exponential populations under joint progressive type$-II$ censoring. *Communications in Statistics-Simulation and Computation*, 44(4):902–923.

Barlow, R., Bartholomew, D., Bremner, J., and Brunk, H. (1972). *Statistical Inference Under Order Restrictions: Theory and Application of Isotonic Regression*. John Wiley & Sons Ltd, New York.

Bhattacharyya, G. and Johnson, R. A. (1974). Estimation of reliability in a multicomponent stress-strength model. *Journal of the American Statistical Association*, 69(348):966–970.

Birnbaum, Z. et al. (1956). On a use of the mann-whitney statistic. In *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, volume 1, pages 13–17. University of California Press Berkeley, CA, USA.

Birnbaum, Z. and McCarty, R. (1958). A distribution-free upper confidence bound for $Pr\{Y < X\}$, based on independent samples of x and y. *The Annals of Mathematical Statistics*, pages 558–562.

Çetinkaya, Ç. (2021). Reliability estimation of the stress–strength model with non-identical jointly type$-II$ censored weibull component strengths. *Journal of Statistical Computation and Simulation*, 91(14):2917–2936.

Çetinkaya, Ç. (2023). Inference of $P(X > Y)$ for the Burr$-XII$ model under generalized progressive hybrid censored data with binomial removals. *Concurrency and Computation: Practice and Experience*, 35(7):e7609.

Chen, M.-H. and Shao, Q.-M. (1999). Monte carlo estimation of bayesian credible and hpd intervals. *Journal of computational and Graphical Statistics*, 8(1):69–92.

Choudhary, N., Tyagi, A., and Singh, B. (2021). Estimation of $R = P[Y < X < Z]$ under progressive type$-II$ censored data from weibull distribution. *Lobachevskii Journal of Mathematics*, 42(2):318–335.

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

Curtis, S. M. (2018). *mcmcplots: Create Plots from MCMC Output*. R package version 0.4.3.

Guangming, P., Xiping, W., and Wang, Z. (2013). Nonparametric statistical inference for $P(X < Y < Z)$. *Sankhya A*, 75:118–138.

Johnson, R. A. (1988). 3 stress-strength models for reliability. *Handbook of statistics*, 7:27–54.

Karam, N. S. and Attia, A. M. (2021). Stress-strength reliability for $P(T < X < Z)$ using dagum distribution. In *Journal of Physics: Conference Series*, volume 1879, page 032004. IOP Publishing.

Kohansal, A. (2022). Inference on stress-strength model for a kumaraswamy distribution based on hybrid progressive censored sample. *REVSTAT-Statistical Journal*, 20(1):51–83.

Kotz, S., Lumelskii, Y., and Pensky, M. (2003). *The stress–strength model and its generalizations: theory and applications*. World Scientific.

Mahdizadeh, M. (2018). On estimating a stress-strength type reliability. *Hacettepe Journal of Mathematics and Statistics*, 47(1):243–253.

Maurya, R. K. and Tripathi, Y. M. (2020). Reliability estimation in a multicomponent stress-strength model for burr xii distribution under progressive censoring. *Brazilian Journal of Probability and Statistics*, 34(2):345–369.

Nayal, A. S., Singh, B., Tripathi, V., and Tyagi, A. (2024). Analyzing stress-strength reliability $\delta = P[U < V < W]$: a bayesian and frequentist perspective with Burr$-XII$ distribution under progressive type-ii censoring. *International Journal of System Assurance Engineering and Management*, 15(6):2453–2472.

Nayal, A. S., Singh, B., Tyagi, A., and Chesneau, C. (2023). Classical and bayesian inferences on the stress-strength reliability $R = P[Y < X < Z]$ in the geometric distribution setting. *AIMS Mathematics*, 8(9):20679–20699.

Pal, A., Mitra, S., and Kundu, D. (2021). Bayesian order-restricted inference of a weibull multi-step step-stress model. *Journal of statistical theory and practice*, 15(2):44.

Pandey, M., Uddin, M. B., and Ferdous, J. (1992). Reliability estimation of an s-out-of-k system with non-identical component strengths: the weibull case. *Reliability Engineering & System Safety*, 36(2):109–116.

Peña, E. A. and Gupta, A. K. (1990). Bayes estimation for the marshall–olkin exponential distribution. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 52(2):379–389.

R Core Team (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Ramadan, D. A., Almetwally, E. M., and Tolba, A. H. (2023). Statistical inference for multi stress–strength reliability based on progressive first failure with lifetime inverse lomax distribution and analysis of transformer insulation data. *Quality and Reliability Engineering International*, 39(6):2558–2581.

Rasethuntsa, T. R. (2020). On the umvue and closed-form bayes estimator for $Pr(X < Y < Z)$ and its generalizations. *arXiv preprint arXiv:2012.06487*.

Saini, S. (2024). Estimation of multi-stress strength reliability under progressive first failure censoring using generalized inverted exponential distribution. *Journal of Statistical Computation and Simulation*, 94(14):3177–3209.

Singh, N. (1980). On the estimation of $Pr(X_1 < Y < X_2)$. *Communications in Statistics-Theory and Methods*, 9(15):1551–1561.

Sultana, F., Çetinkaya, Ç., and Kundu, D. (2024). Estimation of the stress-strength parameter under two-sample balanced progressive censoring scheme. *Journal of Statistical Computation and Simulation*, 94(6):1269–1299.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *the Annals of Statistics*, pages 1701–1728.

Xu, J. and Long, J. S. (2005). Using the delta method to construct confidence intervals for predicted probabilities, rates, and discrete changes. *Stata J*, 5(4):537–559.

Yousef, M. M. and Almetwally, E. M. (2021). Multi stress-strength reliability based on progressive first failure for kumaraswamy model: Bayesian and non-bayesian estimation. *Symmetry*, 13(11):2120.

Yousef, M. M., Hassan, A. S., Alshanbari, H. M., El-Bagoury, A.-A. H., and Almetwally, E. M. (2022). Bayesian and non-bayesian analysis of exponentiated exponential stress–strength model based on generalized progressive hybrid censoring process. *Axioms*, 11(9):455.

REVSTAT – Statistical Journal
Volume 23, Number 5, Special Edition 2025, 617–639
https://doi.org/10.57805/revstat.v23i5.1010

An Easily Verifiable Dispersion Order for Discrete Distributions

Authors: Andreas Eberl
– Institute of Statistics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
andreas.eberl@kit.edu

Bernhard Klar
– Institute of Stochastics, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
bernhard.klar@kit.ed

Alfonso Suárez-Llorens ✉
– Facultad de Ciencias, Dpto. Estadística e IO., Universidad de Cádiz, Cádiz, Spain
alfonso.suarez@uca.es

Received: July 2025 Revised: November 2025 Accepted: November 2025

Abstract:

- Dispersion is a fundamental concept in statistics, yet standard approaches - especially via stochastic orders - face limitations in the discrete setting. In particular, the classical dispersive order, well-established for continuous distributions, becomes overly restrictive for discrete random variables due to support inclusion requirements. To address this, we propose a novel weak dispersive order for discrete distributions. This order retains desirable properties while relaxing structural constraints, thereby broadening applicability. We further introduce a class of variability measures based on probability concentration, offering robust and interpretable alternatives that conform to classical axioms. Empirical illustrations highlight the practical relevance of this framework.

Keywords:

- *dispersion; variability; discrete distribution; univariate stochastic order; Lévy concentration function; robust dispersion measure.*

AMS Subject Classification:

- 60E05, 60E15.

✉ Corresponding author

1. INTRODUCTION

The concept of dispersion - or variability - is a cornerstone of statistics, playing a pivotal role in both theoretical developments and applied methodologies. However, when we speak of "dispersion", it is essential to clarify what precisely is being measured and how this measurement is defined. In this context, two foundational questions arise: Given a random variable X and a functional $\nu(X)$, under what conditions can $\nu(X)$ be considered a valid measure of dispersion? Furthermore, given two random variables X and Y, when can we say that X is more dispersed than Y?

The first question was systematically addressed by Bickel and Lehmann (1976, 1979), who identified a set of core properties that any dispersion functional should satisfy. Specifically, a measure of dispersion $\nu(X)$ should be non-negative, i.e., $\nu(X) \geq 0$, with equality if X is degenerate. It should be translation invariant, so that $\nu(X+k) = \nu(X)$ for any constant k, and it should satisfy absolute homogeneity: $\nu(\lambda X) = |\lambda|\,\nu(X)$. A classical example fulfilling these conditions is the standard deviation. In some contexts, absolute homogeneity is relaxed to positive homogeneity, allowing rescaling only by positive factors.

The second question – how to compare the dispersion of two random variables – has been addressed through stochastic ordering. Several types of stochastic orders have been introduced for this purpose (see, e.g., Shaked and Shanthikumar, 2006; Müller and Stoyan, 2002). Among these, the *dispersive order*, denoted $\leq_{\text{disp}}$, introduced by Bickel and Lehmann and further developed by Lewis and Thompson (1981), plays a central role. According to this order, a random variable X is more dispersed than Y if the differences between corresponding quantiles of X consistently exceed those of Y. This order is particularly strong, implying a range of dispersion-related comparisons and preserving important measures such as the standard deviation. For these reasons, Bickel and Lehmann (1979) and Oja (1981) viewed the dispersive order as a canonical tool for formalizing dispersion comparisons.

In the continuous case, the dispersive order has proven to be a powerful framework for comparing well-known distribution families and has been shown to preserve various classical dispersion measures, including the standard deviation, the Gini mean difference, differential entropy, and the L_2-norm of the density. These properties have made it valuable in fields such as finance, actuarial science, and reliability theory.

Although stochastic orderings for discrete distributions have generally received considerable attention in the literature (see, e.g., Giovagnoly and Wynn, 2008; Klenke and Mattner, 2010), the application of the dispersive order to such distributions presents serious limitations. As discussed in Eberl and Klar (2025) and formalized in Theorem 1.7.3 of Müller and Stoyan (2002), a necessary condition for $F \leq_{\text{disp}} G$ is that the support of F must be contained within that of G. While this condition is benign for continuous distributions, it becomes prohibitively restrictive in the discrete case. It effectively excludes most lattice distributions and empirical distributions with ties or differing sample sizes. These challenges stem from the discontinuities in the cumulative distribution function and the irregular spacing of the quantile levels, which complicate quantile-based comparisons in the discrete domain.

To address these difficulties, Eberl and Klar proposed a new dispersive order tailored to

univariate discrete distributions. This refined order preserves many appealing features of the continuous dispersive order while adapting them to the discrete context. However, despite this advancement, the Eberl-Klar order remains relatively strong and may not accommodate all practical needs.

The goal of this paper is to go further by introducing a more flexible and weaker form of dispersion comparison for discrete random variables. Our approach builds on the notion of probability concentration originally formulated by Lévy (1937), and we establish connections to classical concepts such as majorization, entropy, and randomness. Through a detailed study of the properties of this proposed order, we underscore the necessity for dispersion measures that are well-suited to the structure of discrete data.

This line of reasoning also motivates the introduction of a new family of dispersion measures specifically designed for discrete random variables. These measures fulfill the axiomatic framework of Bickel and Lehmann while offering the adaptability required to capture the peculiarities of discrete distributions. Altogether, our contribution offers a unified and interpretable framework for analyzing dispersion in discrete settings, enriching and extending existing methodologies.

The structure of the paper is as follows. In the next section, we introduce the weak dispersive order for discrete distributions. Section 3 examines its main properties, while Section 4 demonstrates that the proposed order is strictly weaker than the discrete dispersive order of Eberl and Klar. In Section 5, we define new measures of variability based on the concentration function and analyze their properties. Finally, several empirical illustrations and concluding remarks are presented in Sections 6 and 7.

2. THE WEAK DISPERSIVE ORDER FOR DISCRETE DISTRIBUTIONS

In Fernández-Ponce and Suárez-Llorens (2003), the authors introduced the weak dispersive order specifically for continuous distributions, providing several characterizations and analyzing its relationship with both the classical dispersive order and the majorization order. We recall its definition in what follows.

Let X and Y be two random variables with continuous distribution functions F and G, respectively. The random variable X is said to be less weakly dispersive than Y, denoted $X \leq_{\text{wd}} Y$ or equivalently $F \leq_{\text{wd}} G$, if for all $\varepsilon > 0$ the following inequality holds:

$$\sup_{x_0} [F(x_0 + \varepsilon) - F(x_0)] \geq \sup_{y_0} [G(y_0 + \varepsilon) - G(y_0)] .$$

The authors also say that X and Y are equally dispersed in the weakly dispersive sense, denoted by $X =_{wd} Y$, if both $X \leq_{\text{wd}} Y$ and $Y \leq_{\text{wd}} X$ hold. It is easily seen that the relation $\leq_{\text{wd}}$ defines a partial order on the set of continuous distribution functions of real-valued random variables.

As mentioned above, Fernández-Ponce and Suárez-Llorens (2003) considered only continuous distribution functions. In the continuous case, it is well known that

$$\sup_{x_0} [F(x_0 + \varepsilon) - F(x_0)] = Q_X(\varepsilon) = \sup_{x_0} \Pr\{X \in [x_0, x_0 + \varepsilon]\},$$

where $Q_X(\varepsilon)$ is the Lévy concentration function, a widely used measure of probability concentration in the literature.

Remark 2.1. a) Generally speaking, if $X \leq_{\text{wd}} Y$ holds, then whenever there exists an interval of length ε in the support of Y, there also exists an interval of the same length in the support of X that accumulates at least as much probability. The supremum is well-defined, since the expression $F(x+\varepsilon) - F(x)$ is bounded.

(b) Intuitively, this reflects a different perspective compared to the classical dispersive order. In the classical setting, we fix a probability level $q - p$ and search for an interval that is more widely separated. Here, we fix the interval length ε and look for an interval that accumulates more probability.

(c) In the general case (i.e., not necessarily continuous distributions), the supremum may not be attained at a point in the support due to the right-continuity of F. However, this technicality is not crucial for the interpretation or application of the order.

In the discrete case, a natural way to define the weak dispersive order is based directly on the Lévy concentration function, as follows:

Definition 2.1. Let X and Y be two discrete random variables with distribution functions F and G, respectively. Then X is said to be less weakly dispersive than Y, denoted $X \leq_{\text{wd}} Y$, or equivalently $F \leq_{\text{wd}} G$, if for all $\varepsilon > 0$, it holds that

$$Q_X(\varepsilon) = \sup_{x_0} \Pr\{X \in [x_0, x_0 + \varepsilon]\} \;\geq\; Q_Y(\varepsilon) = \sup_{y_0} \Pr\{Y \in [y_0, y_0 + \varepsilon]\}.$$

It is well known that the concentration function $Q_X(\varepsilon)$ is right-continuous and non-decreasing in ε. Moreover, it is easy to verify that $Q_X(\varepsilon) \to 1$ as $\varepsilon \to \infty$. Finally, $Q_X(0)$ can be defined as the limit from the right:

$$Q_X(0) := \lim_{\varepsilon \to 0^+} Q_X(\varepsilon) = \sup\{p_i\},$$

where $p_i = \Pr[X = x_i], i = 1, 2, \ldots,$ denotes the point masses of X.

Definition 2.1 can also be expressed in a more constructive form. Let X and Y be two discrete random variables with supports R_X and R_Y, and associated probability masses $p_i = \Pr[X = x_i]$ and $q_i = \Pr[Y = y_i]$, respectively. The supports may be finite or countably infinite, and they may differ both in cardinality and in the location of their support points.

Given two support points $x_i, x_j \in R_X$, denote their distance by $d^X_{i,j} = |x_j - x_i|$. Analogously, define $d^Y_{i,j} = |y_j - y_i|$ for $y_i, y_j \in R_Y$. Then, $X \leq_{\text{wd}} Y$ if and only if, for every $\varepsilon > 0$,

$$\max_{i,j:d^X_{i,j} \leq \varepsilon} \sum_{k: x_i \leq x_k \leq x_j} p_k \;\geq\; \max_{i,j:d^Y_{i,j} \leq \varepsilon} \sum_{k: y_i \leq y_k \leq y_j} q_k. \tag{2.1}$$

For two lattice distributions on $\mathbb{N}_0$, we have $X \leq_{\text{wd}} Y$ if and only if, for all $m \in \mathbb{N}_0$,

$$\sup_{i \in \mathbb{N}_0} \left\{ \sum_{k=i}^{i+m} p_k \right\} \geq \sup_{j \in \mathbb{N}_0} \left\{ \sum_{k=j}^{j+m} q_k \right\}. \tag{2.2}$$

For any i_0 such that $\sum_{k=i_0}^{\infty} p_k < \max_k\{p_k\}$, we can replace the supremum over $i \in \mathbb{N}_0$ by a maximum over the finite set $\{0, \ldots, i_0\}$. The same holds for Y with a corresponding index j_0. Setting $l_0 = \max\{i_0, j_0\}$, the condition in (2.2) becomes equivalent to

$$\max_{i\in\{0,\ldots,l_0\}} \left\{ \sum_{k=i}^{i+m} p_k \right\} \geq \max_{j\in\{0,\ldots,l_0\}} \left\{ \sum_{k=j}^{j+m} q_k \right\}. \tag{2.3}$$

Hence, the supremum in (2.2) is actually attained. The same argument applies to distributions on $\mathbb{Z}$ and to lattice distributions with arbitrary step size.

3. PROPERTIES OF THE WEAK DISPERSIVE ORDER

This section outlines several properties of the weak dispersive order. We begin with two elementary observations:

Remark 3.1. a) Suppose there exists a fixed $\varepsilon > 0$ that is smaller than the distance between any two distinct points in the support of both X and Y; that is, $\varepsilon < d^X_{i,j}$ and $\varepsilon < d^Y_{l,r}$ for all $i < j$ and $l < r$. In this case, the condition $X \leq_{\text{wd}} Y$ implies that $\max\{p_i\} \geq \max\{q_i\}$ This situation always occurs when both X and Y have finite supports.

b) The supremum defining the Lévy concentration function may be attained at multiple intervals. Since the set of rational numbers is dense in $\mathbb{R}$, it suffices to evaluate the weak dispersive order on a countable collection of ε values. That is, the order relation can be verified by considering only a countable subset of positive ε.

We next present a consequence of the weak dispersive order in terms of the range of the support:

Property 3.1. Let X and Y be discrete random variables with bounded supports. Denote the minimum and maximum elements of the support of X by x_l and x_u, respectively, and those of Y by y_l and y_u. Then, if $X \leq_{\text{wd}} Y$, it follows that

$$x_u - x_l \leq y_u - y_l.$$

Proof: The proof follows by contradiction, setting $\varepsilon = y_u - y_l$. □

The following result allows many comparisons. First, we recall the definition of a contraction function.

Definition 3.1. Let $\phi : \mathbb{R} \mapsto \mathbb{R}$ be a real function. We will say that ϕ is a contraction if $|\phi(y) - \phi(x)| \leq |y - x|$, for all $x, y \in \mathbb{R}$.

Theorem 3.2. *Let X and Y be two random variables such that $X =_{st} \phi(Y)$, where ϕ is a monotone contraction function. Then $X \leq_{wd} Y$.*

Proof: Let us suppose first that ϕ is a non-decreasing contraction. Then

$$\begin{aligned} Pr\{Y \in [x_0, x_0+\varepsilon]\} &\leq Pr\{\phi(Y) \in [\phi(x_0), \phi(x_0+\varepsilon)]\} \\ &= Pr\{X \in [\phi(x_0), \phi(x_0+\varepsilon)]\} \\ &\leq Pr\{X \in [\phi(x_0), \phi(x_0)+\varepsilon]\}, \end{aligned}$$

where the first inequality follows by just taking in account that ϕ is non-decreasing, and the second equality by considering that $X =_{st} \phi(Y)$. For the last inequality, we use both assumptions ϕ is a non-decreasing function and ϕ is a contraction function. Then $|\phi(x_0 + \varepsilon) - \phi(x_0)| = \phi(x_0+\varepsilon) - \phi(x_0) \leq \varepsilon$, which leads that $[\phi(x_0), \phi(x_0+\varepsilon)] \subseteq [\phi(x_0), \phi(x_0) + \varepsilon]$. Therefore, the probabilities are ordered. The proof easily concludes by considering the supremum. The proof for a non-increasing contraction function follows a similar argument. □

Remark 3.2. The inverse of a strictly monotone contraction function is an expansion function, i.e., a function where two images are more widely separated than the corresponding values of the images. Therefore if $\phi(X) =_{st} Y$ for a strictly monotone expansion function then $X \leq_{\text{wd}} Y$ holds.

Corollary 3.1. *If $X \leq_{disp} Y$ in the classical Lewis and Thompson sense, then $X \leq_{wd} Y$ holds.*

Proof: The proof follows directly from using jointly Theorem 1.7.4 of Müller and Stoyan (2002) and Theorem 3.2. □

Remark 3.3. Note, however, that the classical dispersive order is incompatible with almost all discrete distributions, including lattice and most empirical distributions (Eberl and Klar, 2025). The following simple example shows that the weak dispersive order is weaker than the classical dispersive order: Let $U\{1, n\}$ denote the uniform distribution on $\{1, \dots, n\}$. Obviously, $U\{1, m\} \leq_{\text{wd}} U\{1, n\}$ for $m \leq n$, but they are not generally ordered with respect to the dispersive order in the L-T sense. For example, $U\{1, 2\}$ and $U\{1, 5\}$ are not ordered with respect to $\leq_{\text{disp}}$.

As another example, we compare Bernoulli distributions.

Example 3.1. Let $X \sim Be(p_1)$ and $Y \sim Be(p_2)$ be two Bernoulli random variables. A straightforward calculation shows that $X \leq_{\text{wd}} Y$ if and only if $\max\{p_1, 1-p_1\} \geq \max\{p_2, 1-p_2\}$. Then, $Be(p) \leq_{\text{wd}} Be(0.5)$, for all $p \in [0, 1]$.

On the other hand, it follows directly from Prop. 2.5 in Eberl and Klar (2025) that neither $X \leq_{\text{disp}} Y$ nor $Y \leq_{\text{disp}} X$ holds for $p_1 \neq p_2$.

The next result directly follows from Theorem 3.2.

Corollary 3.2. *Given a discrete random variable X, we obtain that*

a) $X \leq_{wd} (\geq_{wd}) aX + b$, $\forall |a| > (<) 1$, $b \in \mathbb{R}$.

b) $X =_{wd} -X + b =_{wd} X + b$, $\forall b \in \mathbb{R}$.

Remark 3.4. Corollary 3.2 points to an important difference between the weak and the usual dispersive order. The equivalence class of a distribution with cdf F with respect to $=_{disp}$ is given by all shifts of F, i.e. $\{F(\cdot - b) : b \in \mathbb{R}\}$ (see Oja, 1981, p. 157), so distributions that are equivalent with respect to dispersion can only differ in location. Consequently, $-X =_{disp} X$ does not hold in general. The same holds for its discrete counterpart considered in Section 4.

Corollary 3.2 b) shows that, besides shifts, also reflections belong to the equivalence class of a distribution with respect to $=_{wd}$. But this is not the complete equivalence class, as the following example shows. Consider random variables X and Y with support $\{1, \ldots, 5\}$ and probability mass functions $p = (0.1, 0.4, 0.05, 0.3, 0.15)$ and $q = (0.4, 0.1, 0.25, 0.15, 0.1)$, respectively. Then, $X =_{wd} Y$, but X and Y are not location shifts of one another.

The following proposition and its proof closely follow Lévy (1937) and hold for all types of random variables—discrete, continuous, or otherwise.

Proposition 3.1. *Let X and Y be two independent random variables. Then,*

$$Q_{X+Y}(\varepsilon) \leq Q_X(\varepsilon), \forall \varepsilon > 0.$$

So $X + Y \geq_{wd} X$ holds.

Proof: By considering the convolution, we obtain that

$$\begin{aligned} \Pr\{X + Y \in [x_0, x_0 + \varepsilon]\} &= \int_{\mathbb{R}} \Pr\{X \in [x_0 - y, x_0 - y + \varepsilon]\} \, dF_Y(y) \\ &\leq \int_{\mathbb{R}} Q_X(\varepsilon) dF_Y(y) \\ &= Q_X(\varepsilon), \forall \varepsilon > 0, \end{aligned}$$

where the first inequality follows from the fact that the interval has length ε. Taking the supremum, the proof follows easily. □

Example 3.2. By Prop. 3.1, many well-known distributions can be ordered with respect to $\geq_{wd}$.

a) Let X and Y have Poisson distributions with parameters λ and μ. Then, $X \leq_{\text{wd}} Y$ if and only if $\lambda \leq \mu$.

b) Let X and Y have binomial distributions $Bin(m,p)$ and $Bin(n,p)$. Then, $X \leq_{\text{wd}} Y$ if and only if $m \leq n$.

c) Let X and Y have negative binomial distributions $NB(r,p)$ and $NB(s,p)$. Then, $X \leq_{\text{wd}} Y$ if and only if $r \leq s$.

d) As an example of distributions with an arbitrary number of modes, consider the Hermite distributions $Herm(a,b), a,b > 0$: if U and V are independent Poisson random variables with parameters a and b, then $Z = U + 2V \sim Herm(a,b)$. Now, if $X \sim Herm(a_1,b_1), Y \sim Herm(a_2,b_2)$, and X and Y are independent, then $X+Y \sim Herm(a_1+a_2, b_1+b_2)$, where $a_1,a_2,b_1,b_2 > 0$. It follows that $Herm(a,b) \leq_{\text{wd}} Herm(c,d)$ if $a \leq c$ and $b \leq d$. It is worth recalling that the Hermite discrete random variable can model overdispersion in count data (see, e.g., Johnson et al., 2005; Kemp and Kemp, 1965).

As two examples, Figure 1 shows the probability mass functions of three Hermite distributed random variables X, Y and Z with $X <_{wd} Y <_{wd} Z$. In the left panel, $a_1 = a_2 = 0.10, a_3 = 0.15, b_1 = 0.10, b_2 = b_3 = 0.15$; in the right panel, $a_1 = a_2 = 0.10, a_3 = 0.15, b_1 = 1.0, b_2 = b_3 = 1.1$.

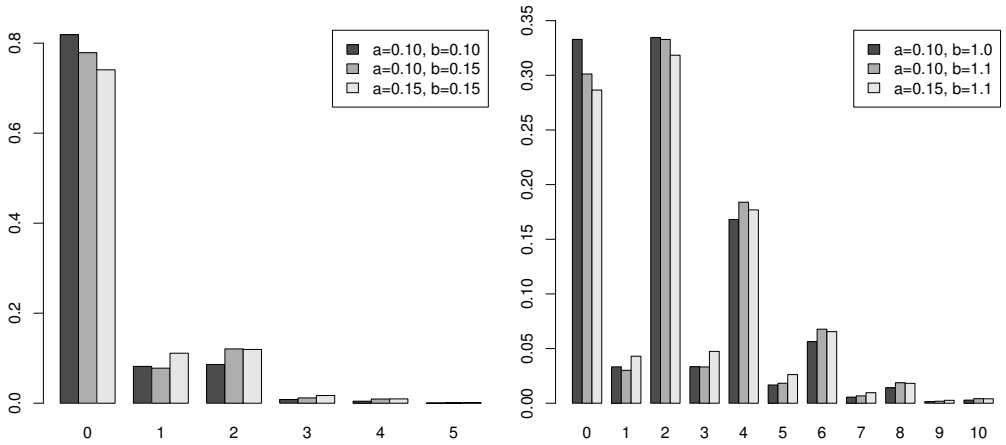


Figure 1: Both panels show probability mass functions of three Hermite distributed random variables X (black), Y (darkgrey) and Z (lightgrey) with $X <_{wd} Y <_{wd} Z$.

When the probability mass functions of two discrete distributions are decreasing, the following result can help to establish the weak dispersive order between the two distributions.

Proposition 3.2. *Let X and Y be two random variables on $\mathbb{N}$ with decreasing probability mass functions. Further, assume $X \leq_{st} Y$, where $\leq_{st}$ denotes the usual stochastic order. Then, $X \leq_{wd} Y$ holds.*

Proof: We just need to take into account that the supremum is always achieved at $k = 1$, i.e.,

$$Q_X(\varepsilon) = \sup_{x_0} \Pr\{X \in [x_0, x_0 + \varepsilon]\} = F_X(1+\varepsilon) = F_X(k),$$

where k is the maximum integer equal to or less than $1+\varepsilon$. The proof then concludes with the assumption that the two distributions are ordered in the stochastic sense. □

Example 3.3. Let us consider two applications of Proposition 3.2.

a) Let $X \sim Geom(p_1)$ and $Y \sim Geom(p_2)$ be two geometric random variables, i.e. the number of trials needed to get one success. By Proposition 3.2, $X \leq_{we} Y$ if and only if $p_1 \geq p_2$.

b) Here, we consider the logarithmic distribution, which has many applications. For example, it has long been used to model species abundances (Fisher et al., 1943) and intensities (see, e.g., Williams, 1964). Let X have a logarithmic distribution with probability mass function

$$f(k,p) = \frac{-1}{\log(1-p)} \frac{p^k}{k}, \quad k \geq 1,$$

and let Y be logarithmically distributed with probability mass function $f(k,q), k \geq 1$, where $0 < p, q < 1$. For the logarithmic distribution, it is easy to see that $X \leq_{lr} Y$, where $\leq_{lr}$ denotes the likelihood ratio order, if $p \leq q$. Since the likelihood ratio order implies the stochastic order (see, e.g., Müller and Stoyan, 2002, p. 12), we obtain $X \leq_{st} Y$ if $p \leq q$. On the other hand, if $X \leq_{st} Y$, then $EX = -[\log(1-p)]^{-1} p/(1-p) \leq EY$, which implies $p \leq q$. Thus, $X \leq_{st} Y$ if and only if $p \leq q$.

Using Proposition 3.2, it immediately follows that $X \leq_{\text{wd}} Y$ if and only if $p \leq q$.

4. RELATION TO THE DISCRETE DISPERSIVE ORDER OF EBERL AND KLAR

We begin with some definitions and results necessary to understand the discrete dispersive order of Eberl and Klar (2025), which serves as a fundamental order for defining measures of dispersion for discrete distributions in an axiomatic way.

Definition 4.1. Let $\mathcal{D}$ denote the set of discrete distributions, and let $F \in \mathcal{D}$ be a cdf with probability mass function f and let $X \sim F$. The class $\mathcal{D}_0 \subseteq \mathcal{D}$ is defined by

$$\begin{aligned} F \in \mathcal{D}_0 \Leftrightarrow\ & \text{supp}(F) \text{ is order-isomorphic to a subset of } \mathbb{Z} \text{ with at least two} \\ & \text{elements} \\ \Leftrightarrow\ & \exists \text{ a subset } A \subseteq \mathbb{Z}, |A| \geq 2, \text{and a bijection } \varphi : \text{supp}(F) \to A \\ & \text{such that } x \leq y \Leftrightarrow \varphi(x) \leq \varphi(y)\ \forall x, y \in \text{supp}(F). \end{aligned}$$

Proposition 4.1. *Define* $\mathbb{I} = \{\mathbb{Z}, \mathbb{N}, -\mathbb{N}\} \cup \{\{1, \ldots, n\} : n \in \mathbb{N}_{\geq 2}\}$ *and*

$$\mathbb{S}_A = \Big\{(x_j, p_j)_{j \in A} \subseteq \mathbb{R} \times (0,1] : x_i < x_j \text{ for } i < j, p_j > 0 \text{ for } j \in A, \sum_{j \in A} p_j = 1\Big\}$$

for $A \in \mathbb{I} \setminus \mathbb{Z}$ *as well as*

$$\mathbb{S}_{\mathbb{Z}} = \Big\{(x_j, p_j)_{j \in \mathbb{Z}} \subseteq \mathbb{R} \times (0,1] : x_i < x_j \text{ for } i < j, p_j > 0 \text{ for } j \in \mathbb{Z}, \sum_{j \in \mathbb{Z}} p_j = 1, \inf\{j \in \mathbb{Z} : \sum_{i \leq j} p_i \geq \tfrac{1}{2}\} = 0\Big\}.$$

For any $F \in \mathcal{D}_0$, there exists a unique index set $A \in \mathbb{I}$ that is order-isomorphic to $\operatorname{supp}(F)$*, and there exists a unique sequence $(x_j, p_j)_{j\in A} \in \mathbb{S}_A$ such that $\Pr(X = x_j) = p_j$ for all $j \in A$. This unique association is denoted by $F \mathrel{\hat{=}} (A, (x_j, p_j)_{j\in A})$. A is said to be the indexing set of F and $(x_j, p_j)_{j\in A}$ is said to be the identifying sequence of F.*

In the following, let $F \mathrel{\hat{=}} (A, (x_j, p_j)_{j\in A})$ and $G \mathrel{\hat{=}} (B, (y_j, q_j)_{j\in B})$. Furthermore, we use the conventions $x_a = -\infty$ and $F(x_a) = 0$ for $a < \min A$ as well as $x_a = \infty$ and $F(x_a) = 1$ for $a > \max A$, provided that the minimum and the maximum exist, respectively. Eberl and Klar (2025) defined the following relations, where $\underline{A} = A \setminus \{\min A\}$.

Definition 4.2. Let $F, G \in \mathcal{D}_0$.

a) The relation $\leftrightharpoons$ on the set $A \times B$ is defined by

$$a \leftrightharpoons b \Leftrightarrow (F(x_{a-1}), F(x_a)) \cap (G(y_{b-1}), G(y_b)) \neq \emptyset.$$

for $a \in A, b \in B$. The set $R(\leftrightharpoons)$ of all $(a, b) \in A \times B$ with $a \leftrightharpoons b$ is said to be the set of (F, G)-*dispersion-relevant* pairs of indices.

b) The relation $\leftrightharpoons_\wedge$ on the set $\underline{A} \times \underline{B}$ is defined by

$$a \leftrightharpoons_\wedge b \Leftrightarrow (a \leftrightharpoons b) \wedge (a - 1 \leftrightharpoons b - 1)$$

for $a \in \underline{A}, b \in \underline{B}$. The set $R(\leftrightharpoons_\wedge)$ of all $(a, b) \in \underline{A} \times \underline{B}$ with $a \leftrightharpoons_\wedge b$ is said to be the set of (F, G)-$\wedge$-*dispersion-relevant* pairs of indices.

Eberl and Klar (2025) proposed the following discrete dispersive order.

Definition 4.3. Let $F, G \in \mathcal{D}_0$. G is said to be *at least as $\wedge$-discretely dispersed* as F, denoted by $F \leq_{\text{disp}}^{\wedge\text{-disc}} G$, if the following two conditions are satisfied:

(i) $q_b \leq p_a \quad \forall (a, b) \in R(\leftrightharpoons)$,

(ii) $x_a - x_{a-1} \leq y_b - y_{b-1} \quad \forall (a, b) \in R(\leftrightharpoons_\wedge)$.

Now, we can formulate the main result of this section. It shows that for distributions in $\mathcal{D}_0$, the weak dispersive order is indeed weaker than the discrete dispersive order.

Theorem 4.1. *Let $F, G \in \mathcal{D}_0$, where $\mathcal{D}_0$ is defined as in Eberl and Klar (2025, Def. 3.1). Then, $F \leq_{\text{disp}}^{\wedge\text{-disc}} G$ implies $F \leq_{\text{wd}} G$.*

Proof: In this proof, we utilize the notation from Eberl and Klar (2025). In particular, let $F \mathrel{\hat{=}} (A, (x_j, p_j)_{j\in A})$ and $G \mathrel{\hat{=}} (B, (y_j, q_j)_{j\in B})$.

In the following, we will prove that, for all $y^* \in \mathbb{R}$ and all $\varepsilon > 0$, there exists a $x^* \in \mathbb{R}$ such that $F(x^* + \varepsilon) - F(x^*) \geq G(y^* + \varepsilon) - G(y^*)$. This then directly implies $F \leq_{\text{wd}} G$.

Let $y^* \in \mathbb{R}$ and $\varepsilon > 0$. Choose $b, \beta \in B \cup \{\min B - 1\}$ with $b \leq \beta$ such that $y_{b+1} = \sup\{y \in \mathbb{R} : G(y) = G(y^*)\}$ and $y_\beta = \inf\{y \in \mathbb{R} : G(y) = G(y^* + \varepsilon)\}$. Since the case

$b = \beta$ is trivial, we assume $b < \beta$. Furthermore, choose $a \in A \cup \{\min A - 1\}$ such that $F(x_a) = \max([0, G(y_b)] \cap F(\mathbb{R}))$. Then, we have $F(x_a) \leq G(y_b) < F(x_{a+1})$.

Let $k_{a+1} = b + 1$ and note that $a + 1 \leftrightharpoons_\wedge k_{a+1}$. Now, exactly one of the following two conditions is satisfied:

(i) $G(y_{b+1}) \leq F(x_{a+1})$,

(ii) $G(y_{b+1}) > F(x_{a+1})$.

In case (i) under the condition $F(x_{a+1}) < 1$, we choose $k_{a+2} \in B \cap [k_{a+1} + 1, \infty)$ such that $G(y_{k_{a+2}-1}) = \min \mathrm{NN}_F^G(a + 2)$. This implies $G(y_{k_{a+2}-1}) \leq F(x_{a+1})$ and $a + 2 \leftrightharpoons_\wedge k_{a+2}$. If $F(x_{a+2}) < 1$, we choose $k_{a+3} \in B$ such that $G(y_{k_{a+3}-1}) = \min \mathrm{NN}_F^G(a + 3)$. According to Eberl and Klar (2025, Lemma A.2 in the supplement), we have $k_{a+3} > k_{a+2}$. As long as $F(x_{a+i}) < 1, i = 2, 3, \ldots$, we can keep on choosing $k_{a+i+1} \in B \cap [k_{a+i} + 1, \infty)$ such that $G(y_{k_{a+i+1}-1}) = \min \mathrm{NN}_F^G(a + i + 1)$.

In case (ii), we choose $k_{a+2} = b+2$, which is equivalent to $G(y_{k_{a+2}-1}) = \max \mathrm{NN}_F^G(a+2)$. (Note that $G(y_{b+1}) < 1$, because assuming $G(y_{b+1}) = 1$ implies $F(x_{a+2}) = G(y_{b+1}) = 1$ and thus $p_{a+2} < q_{b+1}$ in spite of $a + 2 \leftrightharpoons b + 1$, which contradicts $F \leq_{\text{disp}}^{\wedge\text{-disc}} G$.) This implies $G(y_{k_{a+2}-1}) > F(x_{a+1})$ and $a + 2 \leftrightharpoons_\wedge k_{a+2}$ as well as

$$\begin{aligned} 0 &< G(y_{k_{a+2}-1}) - F(x_{a+1}) \\ &= G(y_{b+1}) - F(x_{a+1}) \\ &= (G(y_b) - F(x_a)) + (q_{b+1} - p_{a+1}) \\ &\leq G(y_b) - F(x_a). \end{aligned}$$

Then, we make the same distinction as previously, just increasing the indices by one. Again, exactly one of the following conditions is true:

(i') $G(y_{b+2}) \leq F(x_{a+2})$,

(ii') $G(y_{b+2}) > F(x_{a+2})$.

In case (i') under the condition $F(x_{a+2}) < 1$, we proceed as in case (i), just starting with one index higher, i.e. by choosing $k_{a+3} \in B \cap [k_{a+2}+1, \infty)$ such that $G(y_{k_{a+3}-1}) = \min \mathrm{NN}_F^G(a+3)$. In case (ii'), we proceed as in case (ii), just starting with one index higher, i.e. by choosing $k_{a+3} = b + 3$.

Overall, we iteratively obtain

$$\forall j \in A \cap [a + 1, \infty)\ \exists \text{pairwise distinct } k_j \in B \cap [b + 1, \infty) : j \leftrightharpoons_\wedge k_j. \tag{4.1}$$

In this iteration, we either remain in cases (ii), (ii'), (ii"), ... or we at some point get into cases (i), (i'), (i"), ... and remain there from that point on out. Let $\tilde{\beta} = \max\{k \in B \cap (-\infty, \beta] \mid \exists j \in A : k = k_j\}$ and define $\alpha \in A$ by $\tilde{\beta} = k_\alpha$. Then, (4.1) can be made more specific in the following sense:

$$\forall j \in \{a + 1, \ldots, \alpha\}\ \exists \text{pairwise distinct } k_j \in \{b + 1, \ldots, \tilde{\beta}\} : j \leftrightharpoons_\wedge k_j.$$

It follows that

$$\begin{aligned} x_\alpha - x_{a+1} &= \sum_{j=a+2}^{\alpha} x_j - x_{j-1} \\ &\leq \sum_{j=a+2}^{\alpha} y_{k_j} - y_{k_j-1} \leq \sum_{k=b+2}^{\tilde{\beta}} y_k - y_{k-1} \\ &= y_{\tilde{\beta}} - y_{b+1} \leq y_\beta - y_{b+1} < \varepsilon. \end{aligned}$$

Choose $\delta > 0$ small enough such that $x^* = x_{a+1} - \delta$ and $x^* + \varepsilon \geq x_\alpha$.

If the pair of α and $k_\alpha = \tilde{\beta}$ is in case (i'...'), then we have $G(y_{\tilde{\beta}-1}) \leq F(x_{\alpha-1}) < G(y_{\tilde{\beta}}) \leq G(y_\beta)$. If $F(x_\alpha) = 1$, then $G(y_\beta) \leq F(x_\alpha)$ obviously holds; if $F(x_\alpha) < 1$, then $G(y_{\beta-1}) < \min \mathrm{NN}_F^G(\alpha+1)$ and thus $G(y_\beta) = \min \mathrm{NN}_F^G(\alpha+1) \leq F(x_\alpha)$ holds because of the maximality of $\tilde{\beta}$. Overall, it follows that

$$F(x^* + \varepsilon) - F(x^*) \geq F(x_\alpha) - F(x_a) \geq G(y_\beta) - G(y_b) = G(y^* + \varepsilon) - G(y^*),$$

concluding the proof if the pair of α and $k_\alpha = \tilde{\beta}$ is in case (i'...').

If the pair of α and $k_\alpha = \tilde{\beta}$ is in case (ii'...'), then we have $F(x_{\alpha-1}) < G(y_{\tilde{\beta}-1}) < F(x_\alpha)$ and $G(y_{\tilde{\beta}-1}) - F(x_{\alpha-1}) \leq G(y_b) - F(x_a)$. If $F(x_\alpha) = 1$, then $G(y_\beta) \leq F(x_\alpha)$ obviously holds. If $F(x_\alpha) < 1$, then either $\alpha + 1$ and $k_{\alpha+1} > \beta$ are in case (i'...'), which implies $G(y_\beta) \leq G(y_{k_{\alpha+1}-1}) \leq F(x_\alpha)$; or $\alpha + 1$ and $k_{\alpha+1} = \beta + 1$ are in case (ii'...'), which implies $0 < G(y_\beta) - F(x_\alpha) \leq G(y_b) - F(x_a)$. Overall, it follows that

$$\begin{aligned} F(x^* + \varepsilon) - F(x^*) &\geq F(x_\alpha) - F(x_a) \\ &= (G(y_\beta) - G(y_b)) + (G(y_b) - F(x_a)) - (G(y_\beta) - F(x_\alpha)) \\ &\geq G(y_\beta) - G(y_b) = G(y^* + \varepsilon) - G(y^*), \end{aligned}$$

concluding the proof if the pair of α and $k_\alpha = \tilde{\beta}$ is in case (ii'...'). □

The next example shows that several common dispersion measures do not preserve the weak dispersive order.

Example 4.1. The interquartile range (IQR) does not preserve the order $\leq_{\text{disp}}^{\wedge\text{-disc}}$ (Eberl and Klar, 2025, Theorem 4.7). Since, by Theorem 4.1, the order $\leq_{\text{wd}}$ is weaker than $\leq_{\text{disp}}^{\wedge\text{-disc}}$, it follows that the IQR also does not preserve $\leq_{\text{wd}}$.

Other common measures of dispersion do preserve $\leq_{\text{disp}}^{\wedge\text{-disc}}$ (Eberl and Klar, 2025). However, the following three-point distributions illustrate that this is not generally true for $\leq_{\text{wd}}$. Consider two distributions:

$$p = \{p_0 = 0.6,\ p_1 = 0.2,\ p_2 = 0.2\}, \quad q = \{q_0 = 0.3,\ q_1 = 0.5,\ q_2 = 0.2\}.$$

It is straightforward to verify that $p \leq_{\text{wd}} q$. On the other hand, the dispersion measures for these distributions yield:

- Standard deviations: $\mathrm{SD}_p = 0.80$, $\mathrm{SD}_q = 0.65$;

- Mean absolute deviations from the mean: $\mathrm{MAD}_p = 0.72$, $\mathrm{MAD}_q = 0.54$;
- Gini mean differences: $\mathrm{GMD}_p = 0.80$, $\mathrm{GMD}_q = 0.74$.

All differences $\mathrm{SD}_p - \mathrm{SD}_q$, $\mathrm{MAD}_p - \mathrm{MAD}_q$, and $\mathrm{GMD}_p - \mathrm{GMD}_q$ are strictly positive. Based on these classical measures, the distribution p would be considered to be more dispersed than q, despite $p \leq_{\mathrm{wd}} q$.

As the empirical illustrations in Section 6 show, such discrepancies are relatively rare but stem from deeper conceptual differences, which are discussed in further detail in the next section.

5. NEW MEASURES OF VARIABILITY BASED ON THE CONCENTRATION FUNCTION

The preceding example highlights an important conceptual distinction: the Lévy concentration function captures concentration rather than deviation from a central location. Traditional dispersion measures such as the standard deviation (SD) and the median absolute deviation (MAD) quantify variability with respect to a fixed location parameter, typically the mean or median. In contrast, the Gini mean difference (GMD) assesses variability by averaging the absolute differences between all pairs of independent copies of the random variable. Formally, for a random variable X, the GMD is defined as:

$$\mathrm{GMD}_X = \mathbb{E}_{X'}\left[\mathbb{E}_{X|X'}\left[|X - X'|\right]\right],$$

where X' is an independent copy of X.

In the example under discussion, although all three classical dispersion measures are larger for distribution p than for distribution q, both distributions share the same support. Thus, the higher concentration of p indicates a lower level of uncertainty or randomness. From a practical perspective, if one had to predict a likely value, choosing for p would be more straightforward than for q, suggesting that p is more "predictable". This interpretation is supported by the Shannon entropy values: $H(p) = -\sum p_i \log_2(p_i) = 1.371$, while $H(q) = 1.485$. The lower entropy for p reflects reduced uncertainty.

This relationship is underpinned by the fact that the vector p majorizes q (i.e. $q \preceq p$); that is, $\sum_{i=1}^{k} x_{[i]} \geq \sum_{i=1}^{k} y_{[i]}$ for all k, where $x_{[i]}$ denotes the i-th largest entry of x. Since entropy is Schur-concave - that is, $f(q) \geq f(p)$ if $q \preceq p$ - it decreases as the distribution becomes more concentrated. Hence, in this example, greater concentration corresponds to lower entropy, providing a compelling case where classical dispersion measures may fail to capture the perceived variability in discrete distributions. This underscores the relevance of concentration-based and entropy-based approaches, particularly when comparing distributions with identical support.

However, concentration can be viewed both as a lower bound for several classical measures of variability - those that quantify deviation from a central point - and as a dispersion

measure in its own right. Using the formula

$$\mathbb{E}[Z^r] = \int_0^\infty r t^{r-1} \Pr(Z > t)\, dt$$

for a non-negative random variable Z and $r \geq 1$, we obtain

$$E[|X - a|^r] \geq r \int_0^\infty t^{r-1}(1 - Q_X(2t))\, dt, \quad \forall a \in \mathbb{R}, \forall r \geq 1.$$

Therefore,

$$E[|X - a_r|^r] \geq \frac{r}{2^r} \int_0^\infty \varepsilon^{r-1}(1 - Q_X(\varepsilon))\, d\varepsilon, \quad \forall r \geq 1,$$

where $a_r = \arg\min_a E[|X - a|^r]$, $r \geq 1$. In fact, any functional $\nu_r(X)$ defined as

$$\nu_r(X) = \frac{\sqrt[r]{r}}{2} \left(\int_0^\infty \varepsilon^{r-1}(1 - Q_X(\varepsilon))\, d\varepsilon \right)^{1/r},$$

can also be used as a measure of the variability of X, preserving the weak dispersive order $\leq_{\text{wd}}$.

To describe the properties of this functional, recall that a measure of variability ν is a map from the set of random variables to $\mathbb{R}$, such that given a random variable X, $\nu(X)$ quantifies the variability of X. Next, we list a number of properties that a measure of variability should reasonably satisfy:

- P1 Law invariance: if X and Y have the same distribution, then $\nu(X) = \nu(Y)$.
- P2 Translation invariance: $\nu(X + k) = \nu(X)$ for all X and all constant k.
- P3 Absolute homogeneity: $\nu(\lambda X) = |\lambda|\nu(X)$ for all X and all $\lambda \in \mathbb{R}$.
- P4 Non-negativity: $\nu(X) \geq 0$, with equality if X is degenerate at $c \in \mathbb{R}$.
- P5 Consistency with some dispersion order $\leq_D$: if $X \leq_D Y$, then $\nu(X) \leq \nu(Y)$.

A functional ν satisfying the properties P1-P5 - using $\leq_{\text{disp}}$ as the dispersion order in P5 - is called a measure of variability or spread in the sense of Bickel and Lehmann; see Bickel and Lehmann (1979). It is obvious that $\nu_r(X)$ satisfies P1, P2, and P4. To prove P3, we need only observe that $Q_{\lambda X}(\varepsilon) = Q_X(\varepsilon/|\lambda|)$, from which $\nu_r(\lambda X) = |\lambda|\nu_r(X)$, $\lambda \in \mathbb{R}$, follows. Property P5 - whether using $\leq_{\text{disp}}$ or the discrete dispersive order $\leq_{\text{disp}}^{\wedge\text{-disc}}$ - follows directly from Corollary 3.1 and Theorem 4.1, respectively.

Note that for many discrete distributions, $\nu_r(X)$ is not difficult to compute. For a random variable X with values in $\mathbb{N}_0$, we obtain

$$\nu_r(X) = \frac{1}{2} \left(\sum_{k=0}^\infty \left((k+1)^r - k^r \right) (1 - Q_X(k)) \right)^{1/r}$$

and, in particular, $\nu_1(X) = 1/2 \sum_{k=0}^\infty (1 - Q_X(k))$. Thus, ν_r is essentially the sum over the minimum probability outside an interval of length k; if this is small, the distribution is highly concentrated and its variability is low.

Example 5.1. a) If $X \sim Be(p)$, we obtain $\nu_1(X) = \min\{p, 1-p\}/2,\ \nu_2(X) = \sqrt{\min\{p, 1-p\}}/2$.

b) Let $X \sim Geom(p)$ with $F_X(k) = 1 - (1-p)^k, k \in \mathbb{N}$. Then, $1 - Q_X(k) = (1-p)^{k+1}$, which entails $\nu_1(X) = (1-p)/(2p),\ \nu_2(X) = \sqrt{(1-p)(2-p)}/(2p)$.

Definition 5.1. a) Let X and Y be discrete random variables and let p and q be their probability mass functions (defined on the union of their supports). X is said to be *less random* than Y (written $X \leq_{rand} Y$) if $q \preceq p$, where $\preceq$ denotes the order of majorization (Hickey, 1983).

b) A distribution p with support on the lattice of integers is unimodal if there exists at least one integer M such that

$$p_k \geq p_{k-1},\ \forall k \leq M, \qquad p_{k+1} \leq p_k,\ \forall k \geq M.$$

Closely related characterizations of discrete unimodality are studied in Bertin and Theoderescu (1984). To compare distributions in terms of randomness, Hickey (1984) and Hickey (1986) introduce a notion of majorization for absolutely continuous distributions via the decreasing rearrangement of their density functions. For the absolutely continuous case, Fernández-Ponce and Suárez-Llorens (2003) shows that comparison in terms of randomness is equivalent to the weak dispersive order for unimodal distributions, as defined in Sudhakar and Kumar (1988). The following proposition serves an analogous purpose for discrete distributions.

Proposition 5.1. *Let p and q be unimodal distributions on the non-negative integers, and let X and Y be distributed according to p and q. Then,*

$$X \leq_{wd} Y \Leftrightarrow X \leq_{rand} Y.$$

Proof: By the unimodality assumption,

$$\max_{i \in \mathbb{N}_0} \left\{ \sum_{k=i}^{i+m} p_k \right\} = p_{[0]} + \cdots + p_{[m]},$$

where $p_{[0]} \geq p_{[1]} \geq \ldots$ are the components of p in decreasing order, and the assertion follows. □

Remark 5.1. a) With the help of Prop. 5.1, one can recover the results in Example 3.2 a)-c) and Example 3.3 a), since it is known that the corresponding distributions are ordered with respect to randomness (Hickey, 1983).

b) Consider two unimodal distributions on the non-negative integers with $p_{[i]} = q_{[i]}, i \in \mathbb{N}_0$. Then, $X =_{rand} Y$, and $X =_{wd} Y$ follows by Prop. 5.1. Hence, all unimodal distributions with the same ordered probabilities are equivalent with respect to $=_{wd}$.

c) Of course, without the unimodality assumption, the orders $\leq_{rand}$ and $\leq_{\mathrm{wd}}$ are not equivalent. Consider the distributions $X \equiv \{p_0 = 0.1, p_1 = 0.4, p_2 = 0.4, p_3 = 0.1\}$ and $Y \equiv \{q_0 = 0.4, q_1 = 0.1, q_2 = 0.1, q_3 = 0.4\}$. It is clear that $X =_{rand} Y$, but $X \leq_{\mathrm{wd}} Y$ while $Y \not\leq_{\mathrm{wd}} X$.

d) Consider the class of unimodal distributions whose support is a subset of $\{1, \ldots, M\}$. In this class, the discrete uniform distribution $U\{1, M\}$ is the unique maximizer with respect to ν_1, with

$$\nu_1 = \frac{1}{2} \sum_{k=0}^{M-2} (1 - Q_X(k)) = \frac{1}{2} \sum_{k=1}^{M-1} \left(1 - \frac{k}{M}\right) = \frac{M-1}{4}.$$

Without the unimodality assumption, there is no maximal element: the two-point distribution with mass 1/2 at 1 and M and the uniform distribution are not comparable with respect to $\leq_{\mathrm{wd}}$. Note that for the latter distribution $\nu_1 = (M-1)/4$ holds, as it does for the uniform distribution.

e) The measure ν_r is finite if $\mathbb{E}|X|^r < \infty$, i.e., it is not robust with respect to outliers. However, it is straightforward to define robust dispersion measures based on a similar construction. For a random variable X taking values in $\mathbb{N}_0$, consider the measure

$$\nu_{\mathrm{rob}}(X) = \sum_{k=0}^{\infty} \frac{1 - Q_X(k)}{1 + k^2},$$

which satisfies properties P1, P2, P4, and property P5 with respect to $\leq_{\mathrm{disp}}$, $\leq_{\mathrm{disp}}^{\wedge\text{-disc}}$ and $\leq_{\mathrm{wd}}$ (note that P3 is not meaningful for lattice distributions).

Let $x_1, \ldots, x_n$ be an independent sample from X, and define $M = \max\{x_1, \ldots, x_n\}$. Then, $1 - Q(k) = 0$ for all $k \geq M$. Suppose now that a new observation $x > M$ is added. The empirical influence function satisfies

$$n\left(\nu_{\mathrm{rob}}(x_1, \ldots, x_{n-1}, x) - \nu_{\mathrm{rob}}(x_1, \ldots, x_n)\right) = c + n \sum_{k=M}^{x} \frac{1/n}{1 + k^2} \leq c + \sum_{k=0}^{\infty} \frac{1}{1 + k^2},$$

for some constant c. That is, the empirical influence function is bounded, demonstrating the robustness of ν_{rob} with respect to outliers.

6. EMPIRICAL ILLUSTRATIONS

6.1. Counts of swimbladder nematodes in Japanese eels

The first example compares counts of larvae of swimbladder nematodes in two populations of the Japanese eel (*Anguilla japonica*) from southwest Taiwan. Münderle et al. (2006) compared wild eels from the Kao-Ping River (sample 2, $n = 168$) with cultured eels from an adjacent aquaculture farm (sample 1, $n = 71$). All recorded nematodes belong to the species *Anguillicoloides crassus*. Figure 2 shows bar plots of the relative frequencies p and q for the

two datasets. Table 1 presents the absolute frequencies $h_i (i = 1, 2)$ alongside the relative frequencies for both datasets. We see that $p_0 < q_0$,but $p_1 > q_1$, $p_2 > q_2$, etc., and it is not possible to deduce anything regarding the weak dispersive ordering between the two samples based directly on these values.

k	0	1	2	3	4	5	6	7	8	12	16	21	42	64
h_1	32	15	8	4	1	1	3	2	1	0	1	1	1	1
p	0.45	0.21	0.11	0.06	0.01	0.01	0.04	0.03	0.01	0.00	0.01	0.01	0.01	0.01
h_2	134	19	9	0	1	2	0	1	1	1	0	0	0	0
q	0.80	0.11	0.05	0.00	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.00

Table 1: Absolute frequencies h_1 and h_2 and relative frequencies p and q for the datasets in Example 6.1.

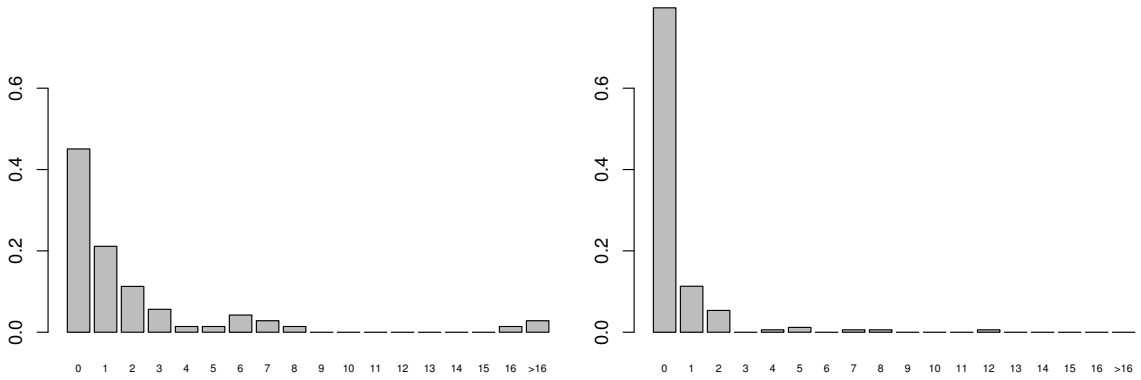


Figure 2: Bar plots of the relative frequencies for the datasets in Example 6.1.

Findings: For these two samples, both $X >_{st} Y$ and $X >_{wd} Y$ hold. Consequently, for any location measure μ, we have $\mu(X) \geq \mu(Y)$. Similarly, for the dispersion measures ν_r and $\widehat{\nu}_{\text{rob}}$ introduced in Section 5, the inequality $\nu(X) \geq \nu(Y)$ also holds. Since the two samples are not comparable with respect to $\leq_{\text{disp}}^{\wedge\text{-disc}}$, no definitive conclusion can be drawn regarding the ordering of SD, MAD, and GMD. Nevertheless, Table 2 shows that even for these classical measures, as well as for the interquartile range, the inequality $\nu(X) \geq \nu(Y)$ still holds:

	SD	MAD	GMD	IQR	$\widehat{\nu}_1$	$\widehat{\nu}_2$	$\widehat{\nu}_{\text{rob}}$
sample 1	9.39	4.29	5.48	2	1.65	4.95	0.90
sample 2	1.43	0.74	0.83	0	0.23	0.75	0.51

Table 2: Dispersion measures for the datasets in Example 6.1.

The following two examples consider swimbladder and intestinal nematodes in European eels (*Anguilla anguilla*). The differences between the two distributions are rather small in Example 6.2 and large in Example 6.3.

6.2. Counts of swimbladder nematodes in European eels

As second example, we consider counts of larvae of swimbladder nematodes in European eels from two different locations: sample 1 (n = 196) from the River Rhine near Karlsruhe, and sample 2 ($n = 100$) from the River Rhine near Sulzbach (Münderle et al., 2006). Again, all recorded nematodes were A. crassus. Klar et al. (2010) concluded that both samples can be assumed to come from the same distribution. For example, a two-sample Kolmogorov-Smirnov test yields a p-value of 0.8. Figure 3 displays bar plots of the relative frequencies p and q for the two datasets. Table 3 presents the absolute frequencies $h_i\,(i = 1, 2)$ alongside the relative frequencies for both datasets.

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	16	22	23
h_1	104	47	16	13	5	3	2	1	0	1	0	0	0	1	1	1	1
p	0.53	0.24	0.08	0.07	0.03	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
h_2	61	16	10	1	2	1	1	1	2	0	2	2	1	0	0	0	0
q	0.61	0.16	0.10	0.01	0.02	0.01	0.01	0.01	0.02	0.00	0.02	0.02	0.01	0.00	0.00	0.00	0.00

Table 3: Absolute frequencies h_1 and h_2 and relative frequencies p and q for the datasets in Example 6.2.

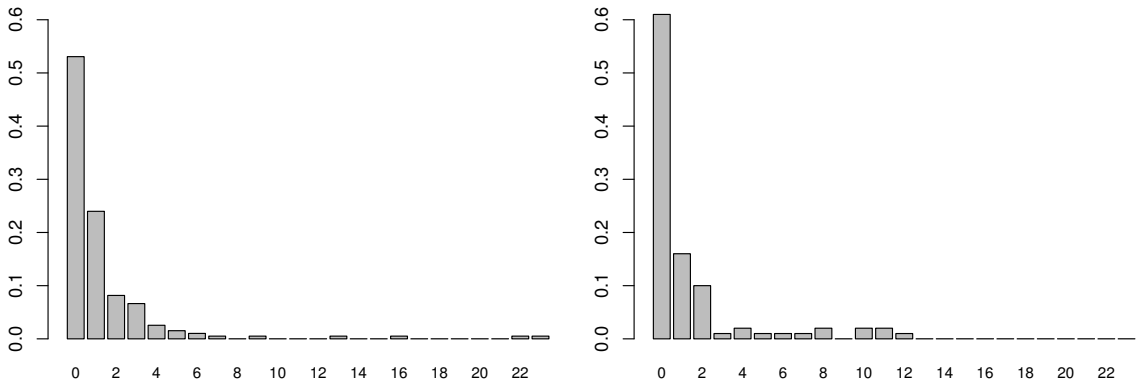


Figure 3: Bar plots of the relative frequencies for the datasets in Example 6.2.

Findings: Under the assumption that both samples originate from the same distribution, one would expect the differences

$$d_m = \sup_{i\in\mathbb{N}_0}\left\{\sum_{k=i}^{i+m} p_k\right\} - \sup_{j\in\mathbb{N}_0}\left\{\sum_{k=j}^{j+m} q_k\right\}$$

to be close to zero for all m, which is indeed the case. Since the sign of d_m varies with m, the samples are not ordered with respect to $\leq_{\text{wd}}$ and, consequently, also not with respect to $\leq_{\text{disp}}^{\wedge\text{-disc}}$. Similarly, the differences in the dispersion measures between the two samples (see Table 4) are small and also vary in sign:

	SD	MAD	GMD	IQR	$\widehat{\nu}_1$	$\widehat{\nu}_2$	$\widehat{\nu}_{\text{rob}}$
sample 1	2.94	1.53	1.99	1	0.65	1.61	0.79
sample 2	2.73	1.76	2.18	1	0.68	1.52	0.75

Table 4: Dispersion measures for the datasets in Example 6.2.

6.3. Counts of intestinal parasites in European eels

The third example compares two samples of counts of intestinal parasites in European eels at the same location (River Rhine near Karlsruhe), but from different years. Sample 1 ($n = 40$) was recorded in summer 1999, and sample 2 ($n = 20$) in summer 2005 (Thielen, 2006). The two samples come from quite different distributions, see Klar et al. (2010) for an explanation. Figure 4 displays bar plots of the relative frequencies p and q for the two datasets. Table 5 presents the absolute frequencies $h_i (i = 1, 2)$ alongside the relative frequencies for both datasets.

k	0	1	2	3	4	5	6	7	8	9	10	11	12	14	27	37	39	≥ 40
h_1	14	3	5	4	2	1	1	1	0	0	1	1	0	0	1	1	1	4
p	0.35	0.07	0.12	0.10	0.05	0.02	0.02	0.02	0.00	0.00	0.02	0.02	0.00	0.00	0.02	0.02	0.02	0.10
h_2	1	0	2	1	3	2	1	0	1	2	0	2	2	1	1	0	1	0
q	0.05	0.00	0.10	0.05	0.15	0.10	0.05	0.00	0.05	0.10	0.00	0.10	0.10	0.05	0.05	0.00	0.05	0.00

Table 5: Absolute frequencies h_1 and h_2 and relative frequencies p and q for the datasets in Example 6.3.

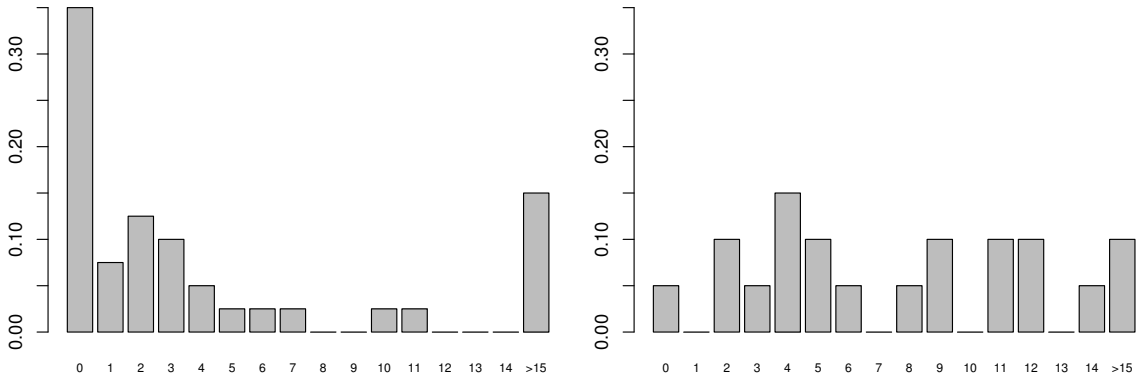


Figure 4: Bar plots of the relative frequencies for the datasets in Example 6.3.

Findings: Here, $d_m > 0$ for $m = 0, \ldots, 10$, while $d_m < 0$ for $m = 12, \ldots, 173$. Consequently, neither $X \leq_{\text{wd}} Y$ nor $Y \leq_{\text{wd}} X$ holds. Table 6 shows that the standard deviation, mean absolute deviation, and Gini mean difference are clearly larger for the first sample; the same applies to $\widehat{\nu}_1$ and $\widehat{\nu}_2$. In contrast, the differences in the interquartile ranges and the robust measure $\widehat{\nu}_{\text{rob}}$ are negative.

	SD	MAD	GMD	IQR	$\widehat{\nu}_1$	$\widehat{\nu}_2$	$\widehat{\nu}_{\text{rob}}$
sample 1	35.83	21.48	25.68	6.25	7.60	19.25	1.06
sample 2	9.19	6.06	8.93	7.25	4.07	6.23	1.23

Table 6: Dispersion measures for the datasets in Example 6.3.

6.4. Amount of aggression attributed to film characters

To illustrate their test procedure, Siegel and Tukey (1960) considered data from a study comparing the amount of aggression attributed to film characters by members of two populations, A and B, both with a sample size of 9. Based on the test result, they concluded that population A is more spread out than population B. Figure 5 displays bar plots of the relative frequencies p and q for the two datasets. Table 7 presents the absolute frequencies $h_i (i = 1, 2)$ alongside the relative frequencies for both datasets.

k	0	3	5	6	8	10	11	12	13	14	15	16	17	19	25
h_1	1	0	1	0	2	0	0	0	0	1	1	0	1	1	1
p	0.11	0.00	0.11	0.00	0.22	0.00	0.00	0.00	0.00	0.11	0.11	0.00	0.11	0.11	0.11
h_2	0	1	0	1	0	2	1	1	2	0	0	1	0	0	0
q	0.00	0.11	0.00	0.11	0.00	0.22	0.11	0.11	0.22	0.00	0.00	0.11	0.00	0.00	0.00

Table 7: Absolute frequencies h_1 and h_2 and relative frequencies p and q for the datasets in Example 6.4.

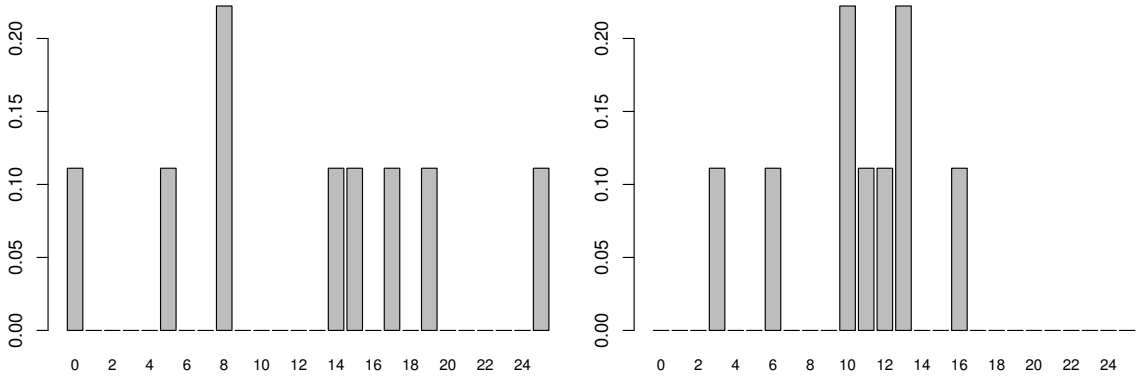


Figure 5: Bar plots of the relative frequencies for the datasets in Example 6.4.

Findings: As in the first example, $X >_{wd} Y$ holds. Therefore, we have $\nu(X) \geq \nu(Y)$ for the dispersion measures introduced in Section 5. Once again, the two samples are not comparable with respect to $\leq_{\text{disp}}^{\wedge\text{-disc}}$, so no conclusion can be drawn regarding the other measures. However, Table 8 shows that, in this case as well, the inequality $\nu(X) > \nu(Y)$ holds for all measures considered:

Remark 6.1. Eberl and Klar (2025) present in Examples 2.7a) and 4.9a) two real-world datasets for which $X \leq_{\text{disp}}^{\wedge\text{-disc}} Y$ holds. Consequently, the classical dispersion measures

	SD	MAD	GMD	IQR	$\widehat{\nu}_1$	$\widehat{\nu}_2$	$\widehat{\nu}_{\text{rob}}$
sample 1	7.75	6.30	9.28	9	4.78	6.28	1.23
sample 2	3.91	2.84	4.50	3	2.11	3.02	1.14

Table 8: Dispersion measures for the datasets in Example 6.4.

SD, MAD, and GMD as well as the dispersion measures introduced in Section 5, agree in their ranking: $\nu(X) \leq \nu(Y)$. In this case, we can confidently state that Y exhibits higher variability than X. However, such a clear ordering is rarely encountered in real-world data.

Subsections 6.1 and 6.4 provide examples where $X >_{wd} Y$ holds. As a result, we have $\nu(X) \geq \nu(Y)$ for the dispersion measures introduced in Section 5. Thus, if we agree on using the order $\leq_{\text{wd}}$ to assess dispersion, we can conclude that the variability of X is greater than that of Y. Even though this ordering does not directly imply a relationship for classical dispersion measures, the examples show that these measures typically follow the same trend in such cases.

In contrast, the example in subsection 6.3 demonstrates that relying on a specific dispersion measure can be misleading, even when the observed differences appear substantial - for instance, standard deviations of 36 and 9. Since the samples are not ordered with respect to $\leq_{\text{wd}}$, neither sample should be deemed more variable than the other in a general sense.

7. CONCLUDING REMARKS

This paper contributes to the ongoing development of dispersion concepts in statistics by proposing a weak dispersive order that is specifically designed for discrete distributions. This new order overcomes the limitations of existing approaches, particularly the classical dispersive order, which is overly restrictive in discrete settings due to its support-inclusion requirement. It also addresses the limitations of the discrete adaptation of this order introduced by Eberl and Klar (2025), which is useful as a foundational concept but imposes strict conditions that are often difficult to verify. The proposed weak dispersive order is more flexible while preserving key interpretability features, and it enables meaningful comparisons of variability between a broader class of discrete distributions.

In addition to the theoretical framework, we introduced a family of variability measures derived from the Lévy concentration function. These measures satisfy the classical axioms of dispersion formulated by Bickel and Lehmann, offering robust and interpretable alternatives to standard dispersion measures in the discrete domain. Several empirical examples demonstrate the practical usefulness of the proposed ordering and associated measures, even in situations where classical dispersion measures may yield conflicting or unintuitive results.

There are several areas that could be explored in future research. For example, a detailed characterization of the equivalence classes under the weak dispersive order for non-unimodal distributions appears to be a challenging and interesting problem. Furthermore, the constructive form of the weak dispersive order - based on comparing probabilities over intervals of bounded length - suggests a natural pathway for extending the concept to mul-

tivariate discrete settings. One possibility would be to replace one-dimensional intervals with Euclidean balls or other metric-based neighborhoods, thereby enabling comparisons of spatially distributed count data, such as point patterns in ecology or materials science.

An additional point of interest lies in the relationship between the concepts of concentration and variability - where variability, in a narrower sense, refers to the spread around a central location. Both notions aim to capture aspects of dispersion; however, the examples suggest that standard measures of variability generally do not respect the weak dispersive order. By contrast, we have seen that concentration can serve as a lower bound for certain variability measures. This raises the following questions: Is there a more rigorous framework that formally clarifies the similarities and differences between these two concepts? For pairs of distributions ordered in variability according to weaker criteria than the dispersive order, such as the convex or dilation orders (see, e.g., Shaked and Shanthikumar, 2006), can we ensure that they are also ordered in concentration?

In summary, this work broadens the conceptual and methodological foundations for comparing variability in discrete data. It provides both theoretical insights and practical tools for statistical analysis in applied fields.

ACKNOWLEDGMENTS

A.S.L. acknowledges support from the Ministerio de Ciencia, Innovación y Universidades (Spain) under grants PID2020-116216GB-I00 and PID2024-157464NB-I00. The authors also thank the anonymous referees and the associate editor for their positive comments and suggestions, which helped improve the manuscript.

REFERENCES

Bertin, E. M. J. and Theoderescu, R. (1984). Some characterizations of discrete unimodality. *Statistics & Probability Letters*, 2(1):23–30.

Bickel, P. J. and Lehmann, E. L. (1976). Descriptive statistics for nonparametric models. III. Dispersion. *The Annals of Statistics*, 4:1139–1158.

Bickel, P. J. and Lehmann, E. L. (1979). Descriptive statistics for nonparametric models IV. Spread. in: J. Jurečková, Ed. In *Contributions to Statistics, Jaroslav Hájek Memorial Volume*. 33-40, Academia, Prague.

Eberl, A. and Klar, B. (2025). Defining dispersion: A fundamental order for univariate discrete distributions. *ALEA - Latin American Journal of Probability and Mathematical Statistics*, 22(2):859–888.

Fernández-Ponce, J. M. and Suárez-Llorens, A. (2003). An aging concept based on majorization. *Probability in the Engineering and Informational Sciences*, 17(1):107–117.

Fisher, R. A., Corbet, A. S., and Williams, C. B. (1943). The relation between the number of species and the number of individuals in a random sample of an animal population. *Journal of Animal Ecology*, 12(1):42–58.

Giovagnoly, A. and Wynn, H. P. (2008). Stochastic orderings for discrete random variables. *Statistics & Probability Letters*, 78(16):2827–2835.

Hickey, R. J. (1983). Majorisation, randomness and some discrete distributions. *Journal of Applied Probability*, 20(4):897–902.

Hickey, R. J. (1984). Continuous majorisation and randomness. *Journal of Applied Probability*, 21:924–929.

Hickey, R. J. (1986). Concepts of dispersion in distributions: A comparative note. *Journal of Applied Probability*, 23:914–921.

Johnson, N. L., Kemp, A. W., and Kotz, S. (2005). *Univariate Discrete Distributions, 3rd edition*. Wiley Series in Probability and Statistics. Wiley.

Kemp, C. D. and Kemp, A. W. (1965). Some properties of the Hermite distribution. *Biometrika*, 52:381–393.

Klar, B., Petney, T. N., and Taraschewski, H. (2010). Quantifying differences in parasite numbers between samples of hosts. *Journal of Parasitology*, 96:856–861.

Klenke, A. and Mattner, L. (2010). Stochastic ordering of classical discrete distributions. *Advances in Applied Probability*, 42(2):392–410.

Lewis, T. and Thompson, J. W. (1981). Dispersive distributions, and the connection between dispersivity and strong unimodality. *Journal of Applied Probability*, 18(1):76–90.

Lévy, P. (1937). *Théorie de l'addition des variables aléatoires*. Gauthier-Villars, Paris.

Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. Wiley Series in Probability and Statistics. Wiley.

Münderle, M., Taraschewski, H., Klar, B., Chang, C. W., Shiao, J. C., Shen, K. N., He, J. T., Lin, S. H., and Tzeng, W. N. (2006). Occurrence of Anguillicola crassus (Nematoda: Dracunculoidea) in Japanese eels Anguilla japonica from a river and an aquaculture unit in southwest Taiwan. *Diseases of Aquatic Organisms*, 71:101–108.

Oja, H. (1981). On location, scale, skewness and kurtosis of univariate distributions. *Scandinavian Journal of Statistics*, 8:154–168.

Shaked, M. and Shanthikumar, J. G. (2006). *Stochastic Orders*. Springer Series in Statistics. Springer.

Siegel, S. and Tukey, J. W. (1960). A nonparametric sum of ranks procedure for relative spread in unpaired samples. *Journal of the American Statistical Association*, 55(291):429–445.

Sudhakar, D. and Kumar, J. (1988). *Unimodality, Convexity, and Applications*. Academic Press.

Thielen, F. (2006). *Der Einfluss einwandernder Amphipodenarten auf die Parasitozönose des europäischen Aals (Anguilla anguilla)*. PhD thesis, University of Karlsruhe.

Williams, C. B. (1964). *Patterns in the Balance of Nature*. Academic Press.

REVSTAT – Statistical Journal
Volume 23, Number 5, Special Edition 2025, 640–672
https://doi.org/10.57805/revstat.v23i5.875

Design of Generalized Hybrid Censoring Schemes With an Application to Likelihood Prediction

Authors: Thomas van Bentum
– Institute of Statistics, RWTH Aachen University, D-52056 Aachen, Germany
bentum@isw.rwth-aachen.de

Erhard Cramer ✉
– Institute of Statistics, RWTH Aachen University, D-52056 Aachen, Germany
erhard.cramer@rwth-aachen.de

Received: November 2024 Revised: January 2026 Accepted: January 2026

Abstract:

- A general approach to the design of hybrid censoring schemes is proposed. It is shown that the hybrid censoring schemes discussed in the literature fit into this scheme. Assuming generalized order statistics as the underlying sample, a representation of the joint density of hybrid censored generalized order statistics is obtained which yields a simple expression for the predictive likelihood function. As an application of the approach, we consider maximum likelihood prediction and maximum observed likelihood prediction. The results are illustrated for exponential and Weibull distributed lifetimes as well as for order statistics, record data/minimal repair data, progressively Type-II censored data, and generalized order statistics, respectively. In particular, the prediction methods are applied to order statistics and minimal repair data illustrating the usefulness of the approaches.

Keywords:

- *likelihood inference; exponential distribution; hybrid censoring; generalized order statistics; minimal repair; point prediction.*

AMS Subject Classification:

- 62N01, 62F99, 62N05.

✉ Corresponding author

1. INTRODUCTION

Hybrid censoring schemes as an extension and alternative to classical Type-I and Type-II censoring have received great attention in the last decade. Dating back to initial discussions in Epstein (1954), who introduced Type-I hybrid censoring for samples with and without replacement, many hybrid censoring schemes have been proposed and inferential results have been established. For a recent survey, we refer to the monograph Balakrishnan et al. (2023) where also a detailed discussion and justification of hybrid censoring is provided (see also the review paper Balakrishnan and Kundu (2013)). Furthermore, it includes an introductory discussion of designing hybrid censoring schemes as well as a couple of data sets and real-world scenarios where hybrid censoring is present. Some applications of the methodology developed in this paper to real data are presented in Section 4. A recent application to one-shot device data can be found in Ling et al. (2026). As pointed out in Cramer (2025), the models can be decomposed in two parts: the design of the hybrid censoring scheme and the underlying data generation process which is subject to the censoring scheme. Based on Epstein's fundamental work on Type-I hybrid censoring, various extensions and adaptations of this censoring design have been presented. Of course, it should also be mentioned that a variety of distributions have been studied, whereby especially the exponential distribution has attracted a great deal of attention, as the distributions of the maximum likelihood estimators can be explicitly determined, for instance.

The first follow up hybrid censoring scheme was Type-II hybrid censoring introduced in Childs et al. (2003), where instead of a maximum test duration, a minimum sample size is guaranteed by the censoring procedure. Afterwards generalized versions have been proposed and analyzed in detail (see Balakrishnan et al. (2023)). From a construction point of view, the design of all these hybrid censoring schemes is based on both time thresholds $T_1, T_2, \ldots$ and censoring numbers $m_1, m_2, \ldots$ (for details see Section 2). One of the most complex hybrid censoring schemes discussed so far with three thresholds and three censoring numbers can be found in Górny and Cramer (2018a). However, a closer look at the probabilistic and inferential results reveals a huge amount of similarities and analogies. As pointed out first in Górny and Cramer (2018b), this observation can be explained by a fundamental decomposition of the joint distribution of the hybrid censored data in so-called *module types* which can be interpreted as basic components of the distribution (see Section 2). This discussion has been refined in Cramer (2025) who showed in detail that, in addition to this modularization approach, the model can be decomposed w.r.t. the censoring design and the data (modeled by the joint distribution of the corresponding random variables). This breakthrough in the analysis of hybrid censoring models thus allows for an application of hybrid censoring schemes for any kind of ordered data. In addition, the present paper presents a general construction design for hybrid censoring schemes called $(k, \zeta, \mathbf{M})$-hybrid censoring schemes in Section 2 which covers the hybrid censoring schemes discussed so far in the literature.

In addition to the design of hybrid censoring schemes, a variety of data situations were also considered, where most hybrid censoring schemes have been designed for order statistics' data (also called without replacement samples). Most of them have been extended to progressively Type-II censored data (see Chapter 7 in Balakrishnan et al. (2023) and, for detailed information on progressively censored data, Balakrishnan and Cramer (2014)). 'With replacement sampling' has been discussed in Epstein (1954) and afterwards in Ebrahimi (1986,

1992), Fairbanks et al. (1982), Draper and Guttman (1987), and recently in Berzborn and Cramer (2023). In the latter paper, it is pointed out that this data situation can alternatively be considered as record or minimal repair data. As a consequence, all the hybrid censoring designs can be applied to such data. An extension of Type-I hybrid censoring to sequential order statistics has been discussed in Burkschat et al. (2016). For extensions to generalized order statistics, we refer to van Bentum and Cramer (2019) and Cramer (2025).

The usefulness of the modularization approach has been stressed in Cramer (2025) by presenting estimation results particularly for exponentially distributed lifetimes. Supplementing this approach, we consider in the present paper an application to point prediction, i.e., maximum likelihood prediction and maximum observed likelihood prediction. Compared to the vast literature on point and interval estimation based on hybrid censored data, prediction problems have only been discussed in the literature for a few hybrid censoring schemes and distributions, including the exponential and Weibull distributions. Some relevant references are Ebrahimi (1992), Asgharzadeh et al. (2015), Valiollahi et al. (2017), and Ateya and Mohammed (2018). On the other hand, there exists a vast literature on prediction based on Type-II censored data, which, as we will show, can be directly utilized in prediction based on hybrid censored data. For more information on prediction for ordered data, we refer to Kaminsky and Nelson (1998), Balakrishnan and Cramer (2014), Balakrishnan et al. (2023), Volovskiy (2018), Volovskiy and Kamps (2020a), Volovskiy and Kamps (2023a), Volovskiy and Kamps (2023b), Empacher et al. (2023), Empacher et al. (2025) and the references cited therein.

In the following, we consider a sequence (or finite set) of strictly increasing random variables $X_{(1)}, X_{(2)}, \ldots$ subject to hybrid censoring. We consider a sample size of n, where in case of an infinite sequence $n = \infty$ is assumed. Moreover, for convenience we define $X_{(0)} = -\infty$ and use the convention $X_{(\infty)} = \infty$. As pointed out above, we discuss two concepts of point prediction in the present manuscript: Maximum likelihood prediction and maximum observed likelihood prediction. For random vectors $\mathbf{X}$ and $\mathbf{Y}$ with joint density $f_{\boldsymbol{\theta}}^{\mathbf{X},\mathbf{Y}}$ parametrized by $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m$, $m \in \mathbb{N}$, maximum likelihood prediction as introduced by Kaminsky and Rhodin (1985) considers the predictive likelihood function (PLF) defined by

$$\mathscr{L}_{\mathsf{PLF}}(\mathbf{y}, \boldsymbol{\theta} \mid \mathbf{x}) = f_{\boldsymbol{\theta}}^{\mathbf{X},\mathbf{Y}}(\mathbf{x}, \mathbf{y}),$$

as a function of $\mathbf{y}$ and $\boldsymbol{\theta}$ (in case of unknown distributional parameters). $\mathbf{Y}$ corresponds to unobservable quantities to be predicted. Then, statistics $\widehat{\mathbf{Y}}_{\mathsf{PLF}} = \widehat{\mathbf{Y}}_{\mathsf{PLF}}(\mathbf{X})$ and $\widehat{\boldsymbol{\theta}}_{\mathsf{PL}} = \widehat{\boldsymbol{\theta}}_{\mathsf{PL}}(\mathbf{X})$ satisfying

$$\mathscr{L}_{\mathsf{PLF}}(\widehat{\mathbf{y}}_{\mathsf{PLF}}(\mathbf{x}), \widehat{\boldsymbol{\theta}}_{\mathsf{PL}}(\mathbf{x}) \mid \mathbf{x}) = \sup_{\mathbf{y},\boldsymbol{\theta}} \mathscr{L}_{\mathsf{PLF}}(\mathbf{y}, \boldsymbol{\theta} \mid \mathbf{x})$$

are called maximum likelihood predictor (MLP) of $\mathbf{Y}$ and predictive maximum likelihood estimator (PMLE) of $\boldsymbol{\theta}$, respectively. In maximum observed likelihood prediction, the joint density function $f_{\boldsymbol{\theta}}^{\mathbf{X},\mathbf{Y}}$ is replaced by the (canonical) conditional density function $f_{\boldsymbol{\theta}}^{\mathbf{X}|\mathbf{Y}}$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^m$, $m \in \mathbb{N}$, of the random vectors $\mathbf{X}$ and $\mathbf{Y}$. This principle has been proposed in Volovskiy and Kamps (2020a) (see also Volovskiy (2018)) as an alternative to maximum likelihood prediction. For details on the motivation and justification of this approach, we refer to these references. Mathematically, maximum observed likelihood prediction is based on the maximization of the observed predictive likelihood function (OPLF) defined by

$$\mathscr{L}_{\mathsf{OPL}}(\mathbf{y}, \boldsymbol{\theta} \mid \mathbf{x}) = f_{\boldsymbol{\theta}}(\mathbf{x} \mid \mathbf{y}).$$

Then, statistics $\widehat{\mathbf{Y}}_{\mathsf{MOLP}} = \widehat{\mathbf{Y}}_{\mathsf{MOLP}}(\mathbf{X})$ and $\widehat{\boldsymbol{\theta}}_{\mathsf{MOL}} = \widehat{\boldsymbol{\theta}}_{\mathsf{MOL}}(\mathbf{X})$ so that

$$\mathscr{L}_{\mathsf{OPL}}(\widehat{\mathbf{y}}_{\mathsf{MOLP}}(\mathbf{x}), \widehat{\boldsymbol{\theta}}_{\mathsf{MOL}}(\mathbf{x}) \mid \mathbf{x}) = \sup_{\mathbf{y},\boldsymbol{\theta}} \mathscr{L}_{\mathsf{OPL}}(\mathbf{y}, \boldsymbol{\theta} \mid \mathbf{x})$$

are called maximum observed likelihood predictor (MOLP) of $\mathbf{Y}$ and predictive maximum observed likelihood estimator (PMOLE) of $\boldsymbol{\theta}$, respectively.

The proposed methods will be illustrated by an application to strength data which was originally obtained by Badar and Priest (1982) and reported in Kundu and Gupta (2006) (see Table 1.1 and Section 4.1). Asgharzadeh et al. (2015) utilized the data to predict larger measurements under Type-I hybrid censoring of the data. The lifetimes were assumed to be Weibull distributed. Additionally, we will apply the prediction approach to hybrid censored

1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397
2.445	2.454	2.474	2.518	2.522	2.525	2.532	2.575	2.614
2.616	2.618	2.624	2.659	2.675	2.738	2.740	2.856	2.917
2.928	2.937	2.937	2.977	2.996	3.030	3.125	3.139	3.145
3.220	3.223	3.235	3.243	3.264	3.272	3.294	3.332	3.346
3.377	3.408	3.435	3.493	3.501	3.537	3.554	3.562	3.628
3.852	3.871	3.886	3.971	4.024	4.027	4.225	4.395	5.020

Table 1.1: Strength of 63 single fibers tested under tension at gauge lengths of 10 mm (measured in GigaPascal (GPA)). The data is reported increasingly ordered, that is, as order statistics.

minimal repair data reported in Kumar and Klefsjö (1992) (see Section 4.1).

The paper is organized as follows. In Section 2, we introduce a general approach to design hybrid censoring schemes which are then called $(k, \zeta, \mathbf{M})$-hybrid censoring schemes. We explain the construction in detail and show that the schemes can be categorized in the group of Type-I hybrid censoring schemes (with bounded test duration but possibly empty sample) and the group of Type-II hybrid censoring schemes (ensuring a minimal sample size but possibly leading to an unbounded test duration). A general decision tree showing the resulting data situations is depicted in Figure 2.1 illustrating the construction principle clearly. Moreover, we show that the hybrid censoring schemes discussed in the literature fit in this scheme (see Table 2.2). Then, we study the distributions involved in the probabilistic analysis of such a scheme and present an expression for the joint distribution in (2.3). Assuming generalized order statistics as the underlying sample, we obtain a representation of the joint density function of hybrid censored generalized order statistics in Theorem 2.1. This leads to the predictive likelihood function given in Theorem 2.2 which is used to obtain point prediction results in Section 3. First, maximum likelihood prediction is studied in Section 3.1, while Section 3.2 deals with maximum observed likelihood prediction. For illustration, we present the results for exponential and Weibull distributed lifetimes and different data situations, that is, order statistics (without replacement sampling), records (minimal repair times, with replacement sampling), progressively Type-II censored data, and generalized order statistics. However, it should be noted that the same approach can be applied to any other lifetime model with ordered measurements. In Sections 4.1 and 4.2, we illustrate the prediction results for hybrid censored order statistics' and minimal repair data, respectively. Finally, we summarize the findings and sketch some future directions of research in Section 5.

2. GENERAL HYBRID CENSORING SCHEMES

2.1. Design of general hybrid censoring schemes

We start the discussion of general hybrid censoring schemes with a detailed design plan. In order to define $(k, \zeta, \mathbf{M})$-hybrid censoring schemes with $k \in \mathbb{N}_0$, $\zeta \in \{\mathsf{I}, \mathsf{II}\}$ and $\mathbf{M} \in \{\mathbf{O}, \mathbf{A}\}$ consider

- increasing thresholds $-\infty = T_0 < T_1 < T_2 < \cdots < T_k < T_{k+1} = \infty$ and
- decreasing censoring numbers $\infty \geq n \geq m_1 > m_2 > \cdots > m_k > m_{k+1} \geq 0$.

For each threshold T_i, we define the random counter $D_i = D(T_i)$ by

$$D_i = \sum_{j=1}^{n} \mathbb{1}_{(-\infty, T_i]}(X_{(j)}), \quad i \in \{1, \ldots, k\}. \tag{2.1}$$

To avoid trivialities, we assume that the conditions (MIN) and (MAX) hold:

(MIN) either $m_{k+1} > 0$ or, if $m_{k+1} = 0$, then $\Pr(D_k > 0) > 0$;

(MAX) either $m_1 < \infty$ or, if $m_1 = n = \infty$, then $\Pr(D_1 < \infty) = 1$.

Essentially, these conditions ensure that the observed sample size for the hybrid censoring scheme under consideration is almost surely neither zero nor infinity, respectively (see flow chart in Figure 2.1 for details). Note that if $m_{k+1} = 0$, then a sufficient condition for (MIN) to hold is that T_k is an interior point of the support of the distribution of $X_{(j)}$, $j \in \{1, \ldots, n\}$. Similarly, in case $m_1 = n = \infty$, then a sufficient condition for (MAX) to hold is that T_1 is an interior point of the support of the distribution of $X_{(j)}$, $j \in \mathbb{N}$. For instance, for $k \in \mathbb{N}$, the condition $0 < P(X_{(j)} \leq T_i) < 1$, $i \in \{1, \ldots, k\}$, $j \in \{1, \ldots, n\}$, and, for $k = 0$, the condition $0 < m_{k+1} = m_1 < \infty$ ensure that (MIN) and (MAX) are satisfied, respectively, with the latter case corresponding to ordinary Type-II censoring (see Table 2.2). Note that this is also the reason why T_{k+1} is a dummy variable, while m_{k+1} is not.

The random counters D_i will play an important role in the probabilistic analysis of hybrid censoring schemes. In fact, the equivalence

$$D_i \geq d \iff X_{(d)} \leq T_i \tag{2.2}$$

for $d \in \{1, \ldots, n\}$ and $i \in \{1, \ldots, k\}$ will be very useful in the following derivations. Throughout, the effectively observed sample size and test duration under a hybrid censoring scheme will be denoted by D_{HCS} and W_{HCS} with corresponding realizations denoted by ℓ and w, respectively.

We distinguish Type I and Type II hybrid censoring schemes recalling a similarity to standard Type-I and Type-II censoring. For Type I hybrid censoring scheme, the sample may be empty. For a Type II hybrid censoring scheme, a *minimal* sample size is guaranteed and given by m_{k+1}. Hence, in the Type I case, we have the minimal sample size $m_{k+1} = 0$ so that the *final* condition $T_k < X_{(m_{k+1})} (= T_k < -\infty)$ in the flow chart in Figure 2.1 is an event of probability zero. In the Type II setting, we have a positive minimal sample size $m_{k+1} > 0$.

		Type I		Type II	
		O	**A**	**O**	**A**
sample size	minimal maximal	random random	random fixed	fixed random	fixed fixed
special parameters	m_{k+1} m_1	$= 0$ $= \infty$	$= 0$ $< \infty$	> 0 $= \infty$	> 0 $< \infty$

Table 2.1: Data scenarios for different hybrid censoring schemes.

Hybrid censoring scheme			censoring numbers m_i				thresholds T_j		
		k	1	2	3	4	1	2	3
Type I HCS	Type-I	1	n	0			T_1		
	Type-I hybrid	1	m_1	0			T_1		
	gen. Type-II hybrid	2	n	m_2	0		T_1	T_2	
	uni. Type-I hybrid	2	m_1	m_2	0		T_1	T_2	
	uni. Type-III hybrid	3	n	m_2	m_3	0	T_1	T_2	T_3
Type II HCS	Type-II	0	m_1						
	Type-II hybrid	1	n	m_2			T_1		
	gen. Type-I hybrid	1	m_1	m_2			T_1		
	uni. Type-II hybrid	2	n	m_2	m_3		T_1	T_2	
	uni. Type-IV hybrid	2	m_1	m_2	m_3		T_1	T_2	
	gen. uni. Type hybrid	3	n	m_2	m_3	m_4	T_1	T_2	T_3

Table 2.2: Hybrid censoring schemes with types and parameters.

These two situations have to be handled separately in the analysis of the hybrid censored data (and, of course, for statistical inference based on the data).

Furthermore, m_1 is considered as the *maximal* sample size of the hybrid censoring scheme, which leads to a classification of the hybrid censoring schemes w.r.t. the parameter **M**. Then, a $(k, \zeta, \mathbf{A})$-hybrid censoring scheme has the property that the maximal sample size is fixed, while for a $(k, \zeta, \mathbf{O})$-hybrid censoring scheme the maximal sample size is random. Hence, in the former case $m_1 < \infty$, while in the latter case $n = m_1 = \infty$ so that the *first* condition $X_{(m_1)} \leq T_1$ or equivalently $D_1 = \infty$ in the flow chart in Figure 2.1 is an event of probability zero (see condition (MAX)). This naming convention is inspired by the module type corresponding to the maximal sample size. Details on the corresponding module types are given below. The data scenarios corresponding to the censoring schemes are summarized in Table 2.1.

It is worth noting that both $(k, \mathsf{I}, \mathbf{O})$- and $(k, \mathsf{II}, \mathbf{O})$-hybrid censoring schemes are only relevant when $n = \infty$, i.e., the initial sample size is unbounded (as in the case of, e.g., record and minimal repair data; see Berzborn and Cramer (2023)). In contrast, the initial sample size is usually assumed to be finite for generalized order statistics' data so that we only need to consider $(k, \mathsf{I}, \mathbf{A})$- and $(k, \mathsf{II}, \mathbf{A})$-hybrid censoring schemes. Note that this also applies to the situation of generalized order statistics with $\gamma_1 = \cdots = \gamma_n = 1$ corresponding to record data with $n \in \mathbb{N}$ finite. Here, it is assumed that only the first n (upper) records are available. In conclusion, we consider the (complete) sequence of upper records in the first setting while the data is restricted to the first n upper records in the second scenario.

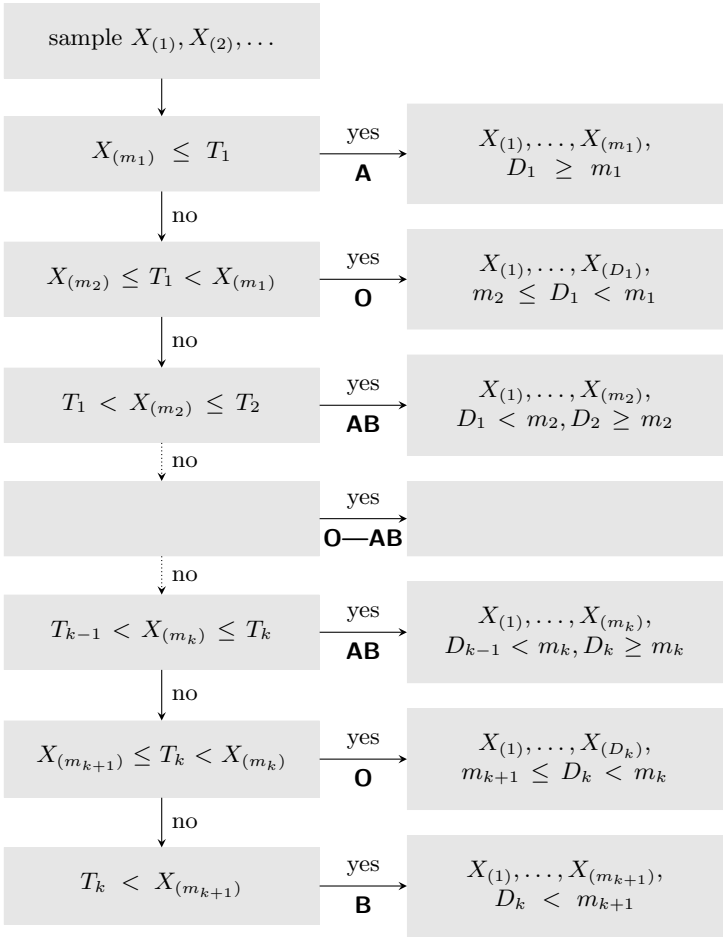


Figure 2.1: Decision tree and rules for a $(k, \zeta, \mathbf{M})$-hybrid censoring scheme.

2.2. Distributions of general hybrid censored data

As pointed out in Górny and Cramer (2018b), Cramer (2025) and illustrated in Figure 2.1, the resulting data scenarios under hybrid censoring can be characterized in terms of four module types **O**, **A**, **B**, **AB**. These module types exhibit a particular structure of the corresponding distribution of the data, that is,

$$\textbf{O:} \quad \Pr(X_{(j)} \leq t_j, 1 \leq j \leq d, D_i = d) = F^{X_{(1)},\ldots,X_{(d)},D_i=d}(\boldsymbol{t}_d),$$

$$\textbf{A:} \quad \Pr(X_{(j)} \leq t_j, 1 \leq j \leq m_\star, D_i \geq m^\star) = F^{X_{(1)},\ldots,X_{(m^\star)},D_i\geq m^\star}(\boldsymbol{t}_{m^\star}),$$

$$\textbf{B:} \quad \Pr(X_{(j)} \leq t_j, 1 \leq j \leq m_\star, D_i < m_\star) = F^{X_{(1)},\ldots,X_{(m_\star)},D_i< m^\star}(\boldsymbol{t}_{m_\star}),$$

$$\textbf{AB:} \quad \Pr(X_{(j)} \leq t_j, 1 \leq j \leq m^\star_\star, D_i < m^\star_\star, D_{i+1} \geq m^\star_\star) = F^{X_{(1)},\ldots,X_{(m^\star_\star)},D_i<m^\star_\star,D_{i+1}\geq m^\star_\star}(\boldsymbol{t}_{m^\star_\star})$$

with appropriately chosen values for $d, m^\star_\star, m_\star, m^\star$ (see Cramer (2025)). Note that these types are very general in the sense that only the joint cumulative distribution function of the ordered data is involved. Furthermore, notice that, due to the ordering $D_i \leq D_{i+1}$, we

obtain the identity

$$\begin{aligned}&\Pr(X_{(j)} \le t_j, 1 \le j \le m_\star, D_i < m_\star^\star, D_{i+1} \ge m_\star^\star)\\&\quad= \Pr(X_{(j)} \le t_j, 1 \le j \le m_\star, D_{i+1} \ge m_\star^\star) - \Pr(X_{(j)} \le t_j, 1 \le j \le m_\star, D_i \ge m_\star^\star)\\&\quad= F^{X_{(1)},\ldots,X_{(m_\star^\star)}, D_{i+1} \ge m_\star^\star}(\boldsymbol{t}_{m_\star^\star}) - F^{X_{(1)},\ldots,X_{(m_\star^\star)}, D_i \ge m_\star^\star}(\boldsymbol{t}_{m_\star^\star})\end{aligned}$$

so that the distribution for module type **AB** can be evaluated using the distribution for module type **A** with different random counters.

Proceeding as in Cramer (2025), we decompose the data situations using the decomposition into disjoint sets as follows:

$$\underbrace{\{D_k < m_{k+1}\}}_{W_{\mathsf{HCS}}=X_{(m_{k+1})}} \cup \Bigg(\bigcup_{i=1}^{k} \underbrace{\{m_{i+1} \le D_i < m_i\}}_{W_{\mathsf{HCS}}=T_i} \Bigg) \cup \Bigg(\bigcup_{i=1}^{k-1} \underbrace{\{D_i < m_{i+1}, D_{i+1} \ge m_{i+1}\}}_{W_{\mathsf{HCS}}=X_{(m_{i+1})}} \Bigg) \cup \underbrace{\{D_1 \ge m_1\}}_{W_{\mathsf{HCS}}=X_{(m_1)}}.$$

The decomposition is constructed by considering the stopping time W_{HCS} of the data. Note that the final term $\{D_1 \ge m_1\}$ is only present when $m_1 < \infty$. This is not true, e.g., for minimal repair and record data (where $m_1 = n = \infty$). Moreover, note that, for $m_1 = n < \infty$, we have $\{D_1 \ge n\} = \{D_1 = n\}$. Then, we get the following decomposition of the joint distribution for a $(k, \zeta, \mathbf{M})$-hybrid censoring scheme

$$\begin{aligned}\Pr(X_{(j)} \le t_j, 1 \le j \le D_{\mathsf{HCS}}) &= \Pr(X_{(j)} \le t_j, 1 \le j \le m_{k+1}, D_k < m_{k+1})\\&\quad+ \sum_{i=1}^{k} \sum_{d=m_{i+1}}^{m_i-1} \Pr(X_{(j)} \le t_j, 1 \le j \le d, D_i = d)\\&\quad+ \sum_{i=1}^{k-1} \Pr(X_{(j)} \le t_j, 1 \le j \le m_{i+1}, D_i < m_{i+1}, D_{i+1} \ge m_{i+1})\\&\quad+ \Pr(X_{(j)} \le t_j, 1 \le j \le m_1, D_1 \ge m_1).\end{aligned}$$

Proceeding as in Cramer (2025), we arrive at

$$\begin{aligned}&= F_{1\ldots m_{k+1}}(\boldsymbol{t}_{m_{k+1}}) - F_{1\ldots m_{k+1}}(\boldsymbol{t}_{m_{k+1}} \wedge T_k)\\&\quad+ \sum_{i=1}^{k} \sum_{d=m_{i+1}}^{m_i-1} \mathbb{1}_{[T_i,\infty)}\Big(\min_{d+1 \le j \le m_i-1} t_j\Big) \Big\{ F_{1\ldots d}(\boldsymbol{t}_d \wedge T_i) - F_{1\ldots d+1}(\boldsymbol{t}_d \wedge T_i, T_i) \Big\}\\&\quad+ \sum_{i=1}^{k-1} \Big(F_{1\ldots m_{i+1}}(\boldsymbol{t}_{m_{i+1}} \wedge T_{i+1}) - F_{1\ldots m_{i+1}}(\boldsymbol{t}_{m_{i+1}} \wedge T_i) \Big) + F_{1\ldots m_1}(\boldsymbol{t}_{m_1} \wedge T_1),\end{aligned} \tag{2.3}$$

where, for convenience, we define $F_{1\ldots 0} \equiv 0$ when $m_{k+1} = 0$ (i.e., in case of a general $(k, \mathsf{I}, \mathbf{M})$-censoring scheme).

As a consequence, we get a closed-form expression when the data is given by generalized order statistics. Note that, according to Kamps (1995) and Cramer and Kamps (2003), the marginal density function of the first $\ell \in \{1, \ldots, n\}$ generalized order statistics is given by

$$f_{(1\ldots\ell)}(\mathbf{x}_\ell) = f^{X_{(j)}, 1 \le j \le \ell}(\mathbf{x}_\ell) = \Big(\prod_{j=1}^{\ell} \gamma_j \Big) \Big(\prod_{j=1}^{\ell-1} f(x_j) \overline{F}^{\gamma_j - \gamma_{j+1} - 1}(x_j) \Big) \overline{F}^{\gamma_\ell - 1}(x_\ell), \quad x_1 \le \cdots \le x_\ell. \tag{2.4}$$

Theorem 2.1 follows along the lines of the corresponding result in Cramer (2025) so that its proof is omitted.

Theorem 2.1. *Consider a $(k, \zeta, \mathbf{M})$-hybrid censoring scheme. Then, for the different module types, the joint density function of the hybrid censored generalized order statistics $X_{(j)}, 1 \le j \le D_{\mathsf{HCS}}$, and the (random) sample size D_{HCS} is given by*

$$f^{X_{(j)},1\le j\le D_{\mathsf{HCS}},D_{\mathsf{HCS}}}(\mathbf{x}_\ell,\ell) =$$

$$\begin{aligned}
\mathbf{O}&: \frac{\overline{F}(T)}{\gamma_{\ell+1} f(T)} f_{(1,\dots,\ell+1)}(\mathbf{x}_\ell, T_i), \quad x_\ell \le T_i < x_{\ell+1},\ \ell \in \{m_{i+1}, \dots, m_i - 1\},\ i \in \{1,\dots,k\}\\
\mathbf{A}&: f_{(1,\dots,\ell)}(\boldsymbol{t}_\ell), \quad x_\ell \le T_1,\ \ell = m_1\\
\mathbf{B}&: f_{(1,\dots,\ell)}(\boldsymbol{t}_\ell), \quad T_k < x_\ell,\ \ell = m_{k+1}\\
\mathbf{AB}&: f_{(1,\dots,\ell)}(\boldsymbol{t}_\ell), \quad T_i < x_\ell \le T_{i+1},\ \ell = m_{i+1},\ i \in \{1,\dots,k-1\}.
\end{aligned}$$

Hence, writing $W_{\mathsf{HCS}} = W(\boldsymbol{X}_{D_{\mathsf{HCS}}}, D_{\mathsf{HCS}})$ for the (random) test duration and $w = w(\mathbf{x}_\ell, \ell)$ for the corresponding realization, the density function can be written in compact form as

$$f^{X_{(j)},1\le j\le D_{\mathsf{HCS}},D_{\mathsf{HCS}}}(\mathbf{x}_\ell,\ell) = \left(\prod_{j=1}^{\ell}\gamma_j\right)\left(\prod_{j=1}^{\ell} f(x_j)\overline{F}^{\gamma_j-\gamma_{j+1}-1}(x_j)\right)\overline{F}^{\gamma_{\ell+1}}(w),$$
$$x_1 \le \dots \le x_\ell \le w, \ell \in \{0,\dots,m_1\}.$$

As a result, the likelihood function for hybrid censored data $x_1, \dots, x_\ell$ with $D_{\mathsf{HCS}} = \ell$ is given by

$$\mathscr{L}_{\mathsf{HCS}}(\boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell) = \left(\prod_{j=1}^{\ell}\gamma_j\right)\left(\prod_{j=1}^{\ell} f_{\boldsymbol{\theta}}(x_j)\overline{F}_{\boldsymbol{\theta}}^{\gamma_j-\gamma_{j+1}-1}(x_j)\right)\overline{F}_{\boldsymbol{\theta}}^{\gamma_{\ell+1}}(w), \tag{2.5}$$

with $\mathbf{x}_\ell = (x_1, \dots, x_\ell)$ and $w \in \{x_\ell, T_i\}$ again denoting the realization of the stopping time of the hybrid censored experiment. Note that the particular hybrid censoring scheme is involved in the definition of ℓ and w only so that it does not affect the functional form of the likelihood.

To derive the predictive likelihood function, we first recall some preliminary distributional results for generalized order statistics. We denote by $\boldsymbol{a}_{1\dots r} = (a_1, \dots, a_r)$ and by

$$\mathbf{G}_{r,r}^{r,0}[\,\cdot \mid \boldsymbol{a}_{1\dots r}] = \mathbf{G}_{r,r}^{r,0}[\,\cdot \mid a_1,\dots,a_r] = \mathbf{G}_{r,r}^{r,0}\left[\,\cdot \,\middle|\, \begin{matrix} a_1,\dots,a_r \\ a_1-1,\dots,a_r-1 \end{matrix}\right]$$

a particular Meijer's G-function defined on the interval $[0,1)$ (see Mathai (1993)). Some useful properties are summarized in Lemma 2.1 (see Cramer et al. (2004)).

Lemma 2.1. *Let $r \ge 2$, $z \in [0,1)$, and $a_1, \dots, a_r > 0$ be positive values with $a_{1:r} \le \dots \le a_{r:r}$ denoting their increasingly ordered values. Furthermore, let $\boldsymbol{a}^{*r} = (a, \dots, a) \in \mathbb{R}^r$ for $a \in \mathbb{R}$. Then,*

(i) $\boldsymbol{G}_{1,1}^{1,0}[z|a_1] = z^{a_1-1}$,

(ii) $\boldsymbol{G}_{2,2}^{2,0}[z|a_1,a_2] = \frac{1}{a_2-a_1}\left(z^{a_1-1} - z^{a_2-1}\right)$, *for* $a_1 \ne a_2$,

(iii) $(a_r - a_1)\boldsymbol{G}_{r,r}^{r,0}[z|\boldsymbol{a}_{1\dots r}] = \boldsymbol{G}_{r-1,r-1}^{r-1,0}[z|\boldsymbol{a}_{1\dots r-1}] - \boldsymbol{G}_{r-1,r-1}^{r-1,0}[z|\boldsymbol{a}_{2\dots r}]$,

(iv) $\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{G}_{r,r}^{r,0}[z|\boldsymbol{a}_{1\dots r}] = \frac{1}{z}\left((a_r-1)\boldsymbol{G}_{r,r}^{r,0}[z|\boldsymbol{a}_{1\dots r}] - \boldsymbol{G}_{r-1,r-1}^{r-1,0}[z|\boldsymbol{a}_{1\dots r-1}]\right)$,

(v) $\boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}] = \frac{1}{z^a}\boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}+\boldsymbol{a}^{*r}]$, $a \in \mathbb{R}$.

(vi) *If* $a_{1:r}=1$ *and* $a_{2:r}>1$, *then* $\lim_{z\to 0^+}\boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}] = \prod_{j=2}^{r}\frac{1}{a_j-1}$.

(vii) *If* $a_{1:r}>1$, *then* $\lim_{z\to 0^+}\boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}] = 0$.

(viii) $\lim_{z\to 1^-}\boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}] = 0$.

(ix) *If* $a=a_1=\cdots=a_r$, *then* $\boldsymbol{G}^{r,0}_{r,r}[z\mid \mathbf{a}_{1\ldots r}] = \frac{1}{(r-1)!}z^{a-1}(-\ln(z))^{r-1}$.

(x) *If* $a_{1:r}\le 1$ *then* $z\mapsto \boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}]$ *is strictly increasing. If* $a_{1:r}>1$ *then* $z\mapsto \boldsymbol{G}^{r,0}_{r,r}[z|\boldsymbol{a}_{1\ldots r}]$ *is strictly unimodal.*

According to Cramer (2006), denoting $1\le j_1<\cdots<j_r\le n$ and $\delta_\rho = j_\rho - j_{\rho-1}$, $1\le\rho\le r$, $j_0=0$, the joint density function of an arbitrary selection of generalized order statistics $X_{(j_1)},\ldots X_{(j_r)}$ can be represented as

$$f_{j_1,j_2,\ldots,j_r}(\mathbf{x}_r) = \left(\prod_{v=1}^{j_r}\gamma_v\right)\left(\prod_{\rho=1}^{r-1}\frac{f(x_\rho)}{\overline{F}(x_\rho)}\right)\left(\prod_{\rho=1}^{r}\mathbf{G}^{\delta_\rho,0}_{\delta_\rho,\delta_\rho}\left[\frac{\overline{F}(x_\rho)}{\overline{F}(x_{\rho-1})}\middle|\boldsymbol{\gamma}_{j_{\rho-1}+1\ldots j_\rho}\right]\right)f(x_r), \tag{2.6}$$

where $-\infty = x_0 < x_1 \le \cdots \le x_r < \boldsymbol{\omega}(F)$, with $\boldsymbol{\omega}(F)$ denoting the right endpoint of the support. In particular, for $\delta_\rho = 1$, $\rho = 1,\ldots,r$, this expression simplifies by Lemma 2.1 (i) to the joint density function of the first r generalized order statistics given in (2.4).

Furthermore, the conditional density function of $X_{(s)}$ given $X_{(r)}$ can be written as

$$f_{s|r}(y|x) = \left(\prod_{v=r+1}^{s}\gamma_v\right)\frac{f(y)}{\overline{F}(x)}\mathbf{G}^{\delta_\rho,0}_{\delta_\rho,\delta_\rho}\left[\frac{\overline{F}(y)}{\overline{F}(x)}\middle|\boldsymbol{\gamma}_{r+1\ldots s}\right], \quad x\le y<\boldsymbol{\omega}(F), \tag{2.7}$$

see Cramer (2003), Theorem 3.3.2. These results enable us to establish the identity (2.8) in Lemma 2.2.

Lemma 2.2. *Let* $s\ge 2$ *and* $a_1,\ldots,a_s>0$. *Then, for any* $-\infty<b<c\le\boldsymbol{\omega}(F)$, *we have the identity*

$$\int_b^c f(x)\,\overline{F}^{a_1-2}(x)\,\boldsymbol{G}^{s-1,0}_{s-1,s-1}\left[\frac{\overline{F}(c)}{\overline{F}(x)}\middle|\boldsymbol{a}_{2\ldots s}\right]\mathrm{d}x = \overline{F}(b)^{a_1-1}\,\boldsymbol{G}^{s,0}_{s,s}\left[\frac{\overline{F}(c)}{\overline{F}(b)}\middle|\boldsymbol{a}_{1\ldots s}\right]. \tag{2.8}$$

Proof: First, by (2.6), (2.4) and Lemma 2.1 (i) we get for $-\infty=x_0<b<x<c<\boldsymbol{\omega}(F)$ the identities

$$\begin{aligned} f_{1,s+1}(b,c) &= \left(\prod_{v=1}^{s+1}\gamma_v\right)\left(\frac{f(b)}{\overline{F}(b)}\right)\left(\frac{\overline{F}(b)}{\overline{F}(x_0)}\right)^{\gamma_1-1}\mathbf{G}^{s,0}_{s,s}\left[\frac{\overline{F}(c)}{\overline{F}(b)}\middle|\boldsymbol{\gamma}_{2\ldots s+1}\right]f(c)\\ &= \left(\prod_{v=1}^{s+1}\gamma_v\right)f(b)\overline{F}^{\gamma_1-2}(b)f(c)\mathbf{G}^{s,0}_{s,s}\left[\frac{\overline{F}(c)}{\overline{F}(b)}\middle|\boldsymbol{\gamma}_{2\ldots s+1}\right] \end{aligned} \tag{2.9}$$

for the joint density function of $X_{(1)}$ and $X_{(s+1)}$. Similarly, we find from (2.6) and (2.4) an expression for the joint density function of $X_{(1)}, X_{(2)}$ and $X_{(s+1)}$, that is,

$$f_{1,2,s+1}(b,x,c) = \left(\prod_{v=1}^{s+1}\gamma_v\right) f(b)\overline{F}^{\gamma_1-2}(b) f(c)\overline{F}^{-(\gamma_2-1)}(b) f(x)\overline{F}^{\gamma_2-2}(x)\, \mathbf{G}_{s-1,s-1}^{s-1,0}\left[\frac{\overline{F}(c)}{\overline{F}(x)}\,\middle|\,\boldsymbol{\gamma}_{3\ldots s+1}\right]. \tag{2.10}$$

However, by marginalizing, we must also have

$$f_{1,s+1}(b,c) = \int_b^c f_{1,2,s+1}(b,x,c)\ \mathrm{d}x.$$

Comparing the equations (2.9) and (2.10) yields the result. □

We are now able to give a closed-form representation of the predictive likelihood function under a general hybrid censoring scheme.

Theorem 2.2. *Denote by* $(\boldsymbol{X}_{(\ell)}, D_{HCS}, W_{HCS}) = (\mathbf{x}_\ell, \ell, w)$ *with* $w \in \{x_\ell, T_i\}$ *appropriately chosen data obtained from the hybrid censoring experiment. Let* $x_{\ell+s} \geq w$, $\ell+s \leq n$. *Then, the predictive likelihood function is given by*

$$\begin{aligned}\mathscr{L}_{PLF}(x_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell) &= \mathscr{L}_{HCS}(\boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell) \cdot \mathscr{L}_{PLF}(x_{\ell+s} \mid \boldsymbol{\theta}, \mathbf{x}_\ell, \ell) \\ &= \left(\prod_{j=1}^{\ell}\gamma_j\right)\left(\prod_{j=1}^{\ell} f_{\boldsymbol{\theta}}(x_j)\overline{F}_{\boldsymbol{\theta}}^{\gamma_j-\gamma_{j+1}-1}(x_j)\right)\overline{F}_{\boldsymbol{\theta}}^{\gamma_{\ell+1}}(w) \\ &\quad \times \left(\prod_{i=\ell+1}^{\ell+s}\gamma_i\right)\frac{f_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\boldsymbol{G}_{s,s}^{s,0}\left[\frac{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right],\end{aligned} \tag{2.11}$$

where

$$\mathscr{L}_{PLF}(x_{\ell+s} \mid \boldsymbol{\theta}, \mathbf{x}_\ell, \ell) = \left(\prod_{i=\ell+1}^{\ell+s}\gamma_i\right)\frac{f_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\boldsymbol{G}_{s,s}^{s,0}\left[\frac{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right]. \tag{2.12}$$

Proof: Let $s = 1$. Then, by (2.4)

$$\begin{aligned}&f_{1\ldots\ell+1,D_{\mathrm{HCS}}}(\mathbf{x}_{\ell+1}, \ell) = f_{1\ldots\ell+1}(\mathbf{x}_{\ell+1}) \\ &= \left(\prod_{i=1}^{\ell}\gamma_i\right)\left(\prod_{i=1}^{\ell} f(x_i)\overline{F}(x_i)^{\gamma_i-\gamma_{i+1}-1}\right) f(x_{\ell+1})\overline{F}^{\gamma_{\ell+1}-1}(x_{\ell+1}), \quad x_1 \leq \ldots x_\ell \leq w \leq x_{\ell+1} < \boldsymbol{\omega}(F).\end{aligned} \tag{2.13}$$

Now, let $s \geq 2$. By analogy with the proof of Lemma 2.2, we have by (2.6) and Lemma 2.1 (i)

$$\begin{aligned}&f_{1\ldots\ell+1,\ell+s}(\mathbf{x}_{\ell+1}, x_{\ell+s}) \\ &= C\left(\prod_{i=1}^{\ell+1}\frac{f(x_i)}{\overline{F}(x_i)}\right)\left(\prod_{i=1}^{\ell+1}\left(\frac{\overline{F}(x_i)}{\overline{F}(x_{i-1})}\right)^{\gamma_i-1}\right)\cdot \mathbf{G}_{s-1,s-1}^{s-1,0}\left[\frac{\overline{F}(x_{\ell+s})}{\overline{F}(x_{\ell+1})}\,\middle|\,\boldsymbol{\gamma}_{\ell+2\ldots\ell+s}\right] f(x_{\ell+s}) \\ &= C\left(\prod_{i=1}^{\ell} f(x_i)\overline{F}(x_i)^{\gamma_i-\gamma_{i+1}-1}\right) f(x_{\ell+1})\overline{F}^{\gamma_{\ell+1}-2}(x_{\ell+1})\mathbf{G}_{s-1,s-1}^{s-1,0}\left[\frac{\overline{F}(x_{\ell+s})}{\overline{F}(x_{\ell+1})}\,\middle|\,\boldsymbol{\gamma}_{\ell+2\ldots\ell+s}\right] f(x_{\ell+s}), \\ &\qquad\qquad x_1 \leq \cdots \leq x_{\ell+1} \leq x_{\ell+s} < \boldsymbol{\omega}(F),\end{aligned}$$

with $C = \prod_{i=1}^{\ell+s} \gamma_i$. Then, marginalization yields

$$
\begin{aligned}
&f_{1\ldots\ell,\ell+s,D_{\mathsf{HCS}}}(\mathbf{x}_\ell, x_{\ell+s}, \ell) = \int_w^{x_{\ell+s}} f_{1\ldots\ell+1,\ell+s}(\mathbf{x}_\ell, x, x_{\ell+s})\,\mathrm{d}x \\
=&C\left(\prod_{i=1}^{\ell} f(x_i)\overline{F}(x_i)^{\gamma_i-\gamma_{i+1}-1}\right)\overline{F}^{\gamma_{\ell+1}-1}(w)\mathbf{G}_{s,s}^{s,0}\left[\frac{\overline{F}(x_{\ell+s})}{\overline{F}(w)}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right] f(x_{\ell+s}) \\
=&\left(\prod_{i=1}^{\ell}\gamma_i\right)\left(\prod_{i=1}^{\ell} f(x_i)\overline{F}(x_i)^{\gamma_i-\gamma_{i+1}-1}\right)\overline{F}^{\gamma_{\ell+1}}(w)\cdot\left(\prod_{i=\ell+1}^{\ell+s}\gamma_i\right)\frac{f(x_{\ell+s})}{\overline{F}(w)}\mathbf{G}_{s,s}^{s,0}\left[\frac{\overline{F}(x_{\ell+s})}{\overline{F}(w)}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right] \\
=&f_{1\ldots\ell,D_{\mathsf{HCS}}}(\mathbf{x}_\ell,\ell)\cdot f_{\ell+s|\ell}(x_{\ell+s}|w), \quad x_1 \le \cdots \le x_\ell \le w \le x_{\ell+s} < \boldsymbol{\omega}(F)
\end{aligned}
$$

by Lemma 2.2. Comparing the result with (2.5) and (2.7) yields the corresponding decomposition. Finally, note that the expression reduces for $s = 1$ to (2.13) by Lemma 2.1(i). □

Remark 2.1. 1. It is worth noting that, for $w = x_\ell$, that is, the observed sample is of Type-II, it is easily verified that $f_{1\ldots\ell,\ell+s,D_{\mathsf{HCS}}}$ as given above reduces to $f_{1\ldots\ell,\ell+s}$, i.e., simply the joint density function of $\mathbf{X}_{(\ell)}$ and $X_{(\ell+s)}$.

Furthermore, the result is not surprising in the sense that it is an extension of the well-known Markov property of generalized order statistics: Given $(\mathbf{X}_{(\ell)}, D_{\mathsf{HCS}}, W_{\mathsf{HCS}}) = (\mathbf{x}_\ell, \ell, w)$, the result shows that $X_{\ell+s}$ is distributed like the s-th generalized order statistics based on parameters $\gamma_{\ell+1}, \ldots, \gamma_n$ and parent distribution truncated on the left at w.

2. The decomposition of the predictive likelihood in (2.11) shows that, for known distribution parameter $\boldsymbol{\theta}$, the likelihood prediction is based only on (2.12).

3. In the present discussion, we assume that the prediction is based on the first D_{HCS} observed hybrid censored random variables $X_{(1)}, \ldots, X_{(D_{\mathsf{HCS}})}$. However, the prediction can also be conducted based on, e.g., $X_{(D_{\mathsf{HCS}})}$ or any selection using the density function given in (2.6). For a single measurement $X_{(D_{\mathsf{HCS}})}$, one can proceed using results established in Buono et al. (2024) for generalized order statistics.

3. APPLICATIONS TO POINT PREDICTION

The above results hold for arbitrary hybrid censoring schemes. Hence, prediction can be treated in a unifying manner. In particular, we recall known results from the literature that so far have been derived for particular censoring schemes or particular models of ordered data. The various models of ordered data can be obtained for appropriate choices for γ_i, $i \in \{1, \ldots, n\}$.

Throughout we use the parametrizations

$$
f(z;\theta) = \begin{cases} \frac{1}{\theta}e^{-\frac{z}{\theta}}, & z \ge 0 \\ 0, & z < 0 \end{cases}, \qquad F(z;\theta) = \begin{cases} 1 - e^{-\frac{z}{\theta}}, & z \ge 0 \\ 0, & z < 0 \end{cases}, \quad \theta > 0,
$$

for the exponential distribution and

$$f(y;\alpha,\lambda)=\begin{cases}\alpha\lambda y^{\alpha-1}e^{-\lambda y^{\alpha}}, & y>0\\ 0, & y\le 0\end{cases},\qquad F(y;\alpha,\lambda)=\begin{cases}1-e^{-\lambda y^{\alpha}}, & y>0\\ 0, & y\le 0\end{cases},\quad \alpha,\lambda>0,$$

for the Weibull distribution with $\boldsymbol{\theta}=(\lambda,\alpha)$, respectively.

3.1. Maximum likelihood prediction

In the current context, the PLF for $X_{D_{\mathsf{HCS}}+s}$ based on hybrid censored ordered data is given in Theorem 2.2. When the distributional parameters $\boldsymbol{\theta}$ are known, the MLP of $X_{D_{\mathsf{HCS}}+s}$ is obtained by maximizing (2.12) w.r.t. $x_{\ell+s}$, that is,

$$\mathscr{L}_{\mathsf{PLF}}(x_{\ell+s}\mid\boldsymbol{\theta},\mathbf{x}_{\ell},\ell)\propto\frac{f_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\mathbf{G}_{s,s}^{s,0}\left[\frac{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right].\tag{3.1}$$

Hence, applying Lemma 2.1 (iv), the normal equation obtained from the log-PLF is given by

$$0=\frac{f'_{\boldsymbol{\theta}}(x_{\ell+s})}{f_{\boldsymbol{\theta}}(x_{\ell+s})}-\frac{f_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}\left((\gamma_{\ell+s}-1)-\frac{\mathbf{G}_{s-1,s-1}^{s-1,0}\left[\frac{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\,\middle|\,\gamma_{\ell+1\ldots\ell+s-1}\right]}{\mathbf{G}_{s,s}^{s,0}\left[\frac{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\,\middle|\,\gamma_{\ell+1\ldots\ell+s}\right]}\right).\tag{3.2}$$

A solution for $x_{\ell+s}$ yields a stationary point and, thus, a candidate for the prediction.

For exponentially distributed lifetimes with mean $\theta>0$ and data $(\mathbf{z}_{\ell},\ell,w)$ we get especially by Lemma 2.1 (v)

$$\mathscr{L}_{\mathsf{PLF}}(z_{\ell+s}\mid\theta,\mathbf{z}_{\ell},\ell)\propto\mathbf{G}_{s,s}^{s,0}\left[\exp\left\{-\left(\frac{z_{\ell+s}-w}{\theta}\right)\right\}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}+\mathbf{1}^{*s}\right]\tag{3.3}$$

(see also Volovskiy (2018), p. 32) where terms independent of $z_{\ell+s}$ in equation (3.3) have been dropped. Note that, for $s\ge 2$, the G-function in (3.3) is strictly unimodal according to Lemma 2.1 (x), since the parameters of the G-function exceed 1. Thus, the maximum is attained at an inner point of the interval (w,∞). For $s=1$, it is decreasing so that the maximum is attained at the left boundary, that is, for $\widehat{z}_{\ell+1}=w$.

By analogy, we arrive for an underlying Weibull distribution at

$$\mathscr{L}_{\mathsf{PLF}}(y_{\ell+s}\mid\theta,\mathbf{y}_{\ell},\ell)\propto y_{\ell+s}^{\alpha-1}\mathbf{G}_{s,s}^{s,0}\left[\exp\left\{-\lambda(y_{\ell+s}^{\alpha}-w^{\alpha})\right\}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}+\mathbf{1}^{*s}\right]\tag{3.4}$$

so that we get the normal equation

$$0=\frac{\alpha-1}{y_{\ell+s}}-\alpha\lambda y_{\ell+s}^{\alpha-1}\left(\gamma_{\ell+s}-\frac{\mathbf{G}_{s-1,s-1}^{s-1,0}\left[\exp\left\{-\lambda(y_{\ell+s}^{\alpha}-w^{\alpha})\right\}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s-1}\right]}{\mathbf{G}_{s,s}^{s,0}\left[\exp\left\{-\lambda(y_{\ell+s}^{\alpha}-w^{\alpha})\right\}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right]}\right)$$

(3.5)
$$\iff \frac{1-1/\alpha}{-\ln t+\lambda w^{\alpha}}=\gamma_{\ell+s}-\frac{\mathbf{G}_{s-1,s-1}^{s-1,0}\left[t\mid\boldsymbol{\gamma}_{\ell+1\ldots\ell+s-1}\right]}{\mathbf{G}_{s,s}^{s,0}\left[t\mid\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}\right]}\quad\text{with } t=\exp\left\{-\lambda(y_{\ell+s}^{\alpha}-w^{\alpha})\right\}.$$

Comments on the solutions of equation (3.5) can be found in Remark 3.3.

In case of unknown distributional parameters $\boldsymbol{\theta}$, the MLP can then be substituted into (2.11), if possible, to obtain the reduced predictive likelihood function and subsequently maximizing w.r.t. $\boldsymbol{\theta}$.

3.1.1. Progressively Type-II censored order statistics

Exponential distribution As already mentioned above, we deduce from Lemma 2.1 (x) that $\mathscr{L}_{\mathsf{PLF}}(z_{\ell+s}\mid\theta,\mathbf{z}_{\ell},\ell)$ is strictly unimodal (see also Cramer, 2004). Furthermore, by Lemma 2.1 (i), the PLF $\mathscr{L}_{\mathsf{PLF}}(z_{\ell+s}\mid\theta,\mathbf{z}_{\ell},\ell)$ attains its maximum on the boundary w for $s=1$. Hence, denoting by $\tilde{z}[\mathbf{a}_{1\ldots r}]$ the uniquely existing mode of $z\mapsto\mathbf{G}_{r,r}^{r,0}[z|\mathbf{a}_{1\ldots r}]$ we immediately get that

$$\widehat{Z}_{(D_{\mathsf{HCS}}+s)}=\begin{cases}W_{\mathsf{HCS}}, & s=1,\\ W_{\mathsf{HCS}}-\theta\ln(\tilde{z}[\boldsymbol{\gamma}_{D_{\mathsf{HCS}+1}\ldots D_{\mathsf{HCS}+s}}+\mathbf{1}^{*}]), & s\geq 2,\end{cases}\tag{3.6}$$

yields the corresponding MLP in case of a known parameter $\theta>0$.

If $\theta>0$ is supposed unknown, plugging in the estimate and ignoring terms independent of $\theta>0$ we obtain (see also Volovskiy (2018), p. 34)

$$\mathscr{L}_{\mathsf{PLF}}(\widehat{z}_{\ell+s}\mid\theta,\mathbf{z}_{\ell},\ell)\propto\frac{1}{\theta}$$

and hence

$$\begin{aligned}\mathscr{L}_{\mathsf{black}}(\theta\mid\mathbf{z}_{\ell+s},\ell)&=\mathscr{L}_{\mathsf{PLF}}(\widehat{z}_{\ell+s},\theta\mid\mathbf{z}_{\ell},\ell)\propto\frac{1}{\theta}\,\mathscr{L}_{\mathsf{HCS}}(\theta\mid\mathbf{z}_{\ell},\ell)\\&\propto\frac{1}{\theta^{\ell+1}}\exp\left\{-\frac{1}{\theta}\left(\sum_{i=1}^{\ell}\gamma_i(z_i-z_{i-1})+\gamma_{\ell+1}(z_{\ell}-w)\right)\right\}=\frac{1}{\theta^{\ell+1}}\exp\left\{-\frac{1}{\theta}\mathsf{TTT}_{\ell}\right\},\end{aligned}$$

where we denote by $\mathsf{TTT}_{\ell}=\sum_{i=1}^{\ell}\gamma_i(z_i-z_{i-1})+\gamma_{\ell+1}(w-z_{\ell})$ the total time on test statistic. Hence, we get

$$\widehat{\theta}_{\mathsf{PL}}=\frac{\mathsf{TTT}_{D_{\mathsf{HCS}}}}{D_{\mathsf{HCS}}+1}.\tag{3.7}$$

Remark 3.1. Note that maximum likelihood prediction for the exponential distribution in the quite general setting of an observable selection $(Z_{(j_1)},\ldots,Z_{(j_r)})$ of general order statistics with additional Type-II censoring has already been considered in Volovskiy (2018), Chapter 3. We can see that the estimators are functionally invariant if we substitute W_{HCS} for the largest observation and the random counter D_{HCS} for the corresponding index. Similar comments apply to subsequently considered submodels resulting from particular choices for the parameters γ_i, $i\in\{1,\ldots,n\}$, which are contained in results presented in Volovskiy (2018), Corollary 3.3.

Remark 3.2. While it does not seem possible to give a more explicit representation of the mode of corresponding Meijer's G-function for general parameters, there exists an explicit sum representation, which can be used for numerical computations (see Cramer and Kamps (2003)). Since parameters are pairwise different for progressively Type-II censored generalized order statistics, (3.3) simplifies to the finite sum (see Kamps and Cramer (2001))

(3.8)
$$\mathbf{G}_{s,s}^{s,0}\left[\exp\left\{-\left(\frac{z_{\ell+s}-w}{\theta}\right)\right\}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}+\mathbf{1}^{*s}\right]=\sum_{k=\ell+1}^{\ell+s}a_{k,\ell+s}^{(\ell)}\exp\left\{-\gamma_k\left(\frac{z_{\ell+s}-w}{\theta}\right)\right\}$$

with $a_{k,\ell+s}^{(\ell)}=\prod_{j=\ell+1,j\neq k}^{\ell+s}(\gamma_j-\gamma_k)^{-1}$.

The models considered in the following differ mainly in that they are based on special choices for the parameters γ_i, $i\in\{1,\ldots,n\}$, which in some cases lead to explicit representations of the mode.

Weibull distribution The existence of a solution follows directly from the PLF in (3.4): First, the PLF is a continuous and non-negative function on the interval (w,∞) with limit $\lim_{y_{\ell+s}\to w+}\mathscr{L}_{\mathsf{PLF}}(y_{\ell+s}\mid\theta,\mathbf{y}_\ell,\ell)=0$ by Lemma 2.1 (viii). For $y_{\ell+s}\to\infty$, we get by Lemma 2.1 (viii) that $\lim_{y_{\ell+s}\to\infty}\mathscr{L}_{\mathsf{PLF}}(y_{\ell+s}\mid\theta,\mathbf{y}_\ell,\ell)=0$ when $\alpha<1$. For $\alpha>1$, we apply Lemma 2.1 (v), (vi) to get

$$\begin{aligned}&\lim_{y_{\ell+s}\to\infty}\mathscr{L}_{\mathsf{PLF}}(y_{\ell+s}\mid\theta,\mathbf{y}_\ell,\ell)\\&=\lim_{y_{\ell+s}\to\infty}y_{\ell+s}^{\alpha-1}\exp\left\{-\lambda\varepsilon(y_{\ell+s}^{\alpha}-w^{\alpha})\right\}\mathbf{G}_{s,s}^{s,0}\left[\exp\left\{-\lambda(y_{\ell+s}^{\alpha}-w^{\alpha})\right\}\,\middle|\,\boldsymbol{\gamma}_{\ell+1\ldots\ell+s}+(1-\varepsilon)^{*s}\right]=0,\end{aligned}$$

where $\varepsilon=\min_{1\le k\le s}\gamma_{\ell+k}/2$. Therefore, for given α,λ, the global maximum of the PLF is attained at an inner point in the interval (w,∞).

So far, we can prove uniqueness only when $\alpha\ge1$. In this case, the Weibull distribution has a log-concave density (see, e.g., Marshall and Olkin (2007, p. 324)). It is easy to verify that the density of a truncated distribution with log-concave density is also log-concave. In particular, the truncated Weibull has a log-concave density in this case. Furthermore, by Remark 2.1, given a realization $(\mathbf{Y}_{(\ell)},D_{\mathsf{HCS}},W_{\mathsf{HCS}})=(\mathbf{y}_\ell,\ell,w)$, $Y_{(\ell+s)}$ is distributed as the s-th generalized order statistics with parameters $\gamma_{\ell+1},\ldots\gamma_n$ and a parent distribution truncated on the left at w. Hence, according to Theorem 2.6 in Cramer (2004), $\mathscr{L}_{\mathsf{PLF}}(y_{\ell+s}\mid\mathbf{y}_\ell,\ell,\alpha,\lambda)$ is a log-concave function of $y_{\ell+s}$, implying that there exists a unique maximizer.

The case $0<\alpha<1$ remains an open problem so far, since the preceding arguments can not be applied. However, extensive numerical studies suggest that the solution is also unique for any $\alpha\in(0,1)$.

Remark 3.3. Alternatively, one can argue that, for $\alpha>1$, the left hand side of (3.5) is an increasing function in t. Using results of Korwar (2003) (see also Belzunce et al. (2005), Izadi and Khaledi (2007)), uniform generalized order statistics are likelihood ratio ordered so that the ratio on the right is a strictly increasing function of t. Hence, the function on the right hand side is strictly decreasing so that the graphs cross at most once and equation (3.5) has at most one solution.

In case of unknown distributional parameters, we get taking partial derivatives w.r.t. λ and α in addition to the normal equation in (3.5) the reduced normal equations

$$\frac{\ell+1}{\lambda} - \left(\sum_{i=1}^{\ell}(\gamma_i-\gamma_{i+1})y_i^{\alpha}+\gamma_{\ell+1}w^{\alpha}\right) - (y_{\ell+s}^{\alpha}-w^{\alpha})\left(\frac{1-1/\alpha}{\lambda y_{\ell+s}^{\alpha}}\right) = 0 \tag{3.9}$$

for $\lambda > 0$ and

$$\frac{\ell+1}{\alpha} + \sum_{i=1}^{\ell}\ln(y_i) + y_{\ell+1} - \lambda\left(\sum_{i=1}^{\ell}(\gamma_i-\gamma_{i+1})y_i^{\alpha}\ln(y_i)+\gamma_{\ell+1}w^{\alpha}\ln(w)\right) - \left(y_{\ell+s}^{\alpha}\ln(y_{\ell+s}) - w^{\alpha}\ln(w)\right)\frac{1-1/\alpha}{y_{\ell+s}^{\alpha}} = 0$$

for $\alpha > 0$, respectively. Solving equation (3.9) for λ leads to

$$\lambda = \frac{1}{\sum_{i=1}^{\ell}(\gamma_i-\gamma_{i+1})y_i^{\alpha}+\gamma_{\ell+1}w^{\alpha}}\left(\ell+1-\left(1-\frac{1}{\alpha}\right)\left(1-\frac{w^{\alpha}}{y_{\ell+s}^{\alpha}}\right)\right)$$

as a function of α and $y_{\ell+s}$. Note that this expression reduces to the reciprocal of (3.7) in the case $\alpha = 1$. Solutions for $\alpha > 0$ and $y_{\ell+s} > w$ have to be found numerically.

3.1.2. Order statistics

For a sample of order statistics, the parameters γ_i, $i \in \{1,\dots,n\}$, satisfy $\gamma_i = n-i+1$, $\gamma_i - \gamma_{i+1} = 1$, $i \in \{1,\dots,n\}$, $\gamma_{n+1} = 0$. Hence, (cf. Volovskiy (2018), Lemma A.4 and Corollary 3.3.)

$$\mathbf{G}_{s,s}^{s,0}[z \mid \gamma_{\ell+1\dots\ell+s} + \mathbf{1}^{*s}] = \frac{1}{(s-1)!}z^{n-\ell-s+1}(1-z)^{s-1}, \quad z \in (0,1), \tag{3.10}$$

with unique mode

$$\tilde{z}[\gamma_{\ell+1\dots\ell+s} + \mathbf{1}^{*s}] = \frac{\gamma_{\ell+s}}{\gamma_{\ell+1}} = \frac{n-\ell-s+1}{n-\ell}. \tag{3.11}$$

The mode can also be obtained from well-known results for the mode of the beta distribution (see also Buono et al. (2024), Section 4.2.1).

Exponential distribution Using (3.11), the MLP is given by

$$\widehat{Z}_{(D_{\mathsf{HCS}}+s)} = W_{\mathsf{HCS}} - \theta\ln\left(\frac{n-D_{\mathsf{HCS}}-s+1}{n-D_{\mathsf{HCS}}}\right) = \begin{cases} W_{\mathsf{HCS}}, & s=1, \\ W_{\mathsf{HCS}} - \theta\ln\left(\dfrac{n-D_{\mathsf{HCS}}-s+1}{n-D_{\mathsf{HCS}}}\right), & s \geq 2. \end{cases} \tag{3.12}$$

Furthermore, the functional form of $\widehat{\theta}_{\mathsf{PL}}$ is equivalent, with the total time on test statistic simplifying to

$$\mathsf{TTT}_{D_{\mathsf{HCS}}} = \sum_{i=1}^{D_{\mathsf{HCS}}} Z_{(i)} + (n-D_{\mathsf{HCS}})\,W_{\mathsf{HCS}}. \tag{3.13}$$

Remark 3.4. Using identity (3.10), the joint PLF for order statistics with $\gamma_i - \gamma_{i+1} - 1 = 0$, $i \in \{1, \dots, n\}$ in (2.11) simplifies to

$$\mathscr{L}_{\mathsf{PLF}}(x_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell) = \frac{n!}{(s-1)!(n-\ell-s)!} \left(\prod_{i=1}^{\ell} f_{\boldsymbol{\theta}}(x_i) \right) (F_{\boldsymbol{\theta}}(x_{\ell+s}) - F_{\boldsymbol{\theta}}(w))^{s-1} \times (1 - F_{\boldsymbol{\theta}}(x_{\ell+s}))^{n-\ell-s} f(x_{\ell+s}).$$

This representation can be found for Type-II censored data in the original paper by Kaminsky and Rhodin (1985) for, e.g., Type-I hybrid censoring in Asgharzadeh et al. (2015) and more generally in Balakrishnan et al. (2023), Chapter 12. Furthermore, the functional form of the estimators agrees with results for the exponential distribution presented in, e.g., Kaminsky and Rhodin (1985) for Type-II censoring and Ebrahimi (1992) for Type-I hybrid censoring.

Weibull distribution Using (3.5) and (3.10), we arrive at the simplified normal equation

$$0 = \frac{\alpha - 1}{y_{\ell+s}} - \alpha\lambda y_{\ell+s}^{\alpha-1} \left((n - \ell - s + 1) - (s-1) \frac{\exp\{-\lambda(y_{\ell+s}^{\alpha} - w^{\alpha})\}}{1 - \exp\{-\lambda(y_{\ell+s}^{\alpha} - w^{\alpha})\}} \right),$$

which has the same functional representation as the one reported in Asgharzadeh et al. (2015) for Type-I hybrid censoring. Hence, we immediately find from their results that

$$\widehat{Y}_{(D_{\mathsf{HCS}}+s)} = \left(W_{\mathsf{HCS}}^{\widehat{\alpha}_{\mathsf{PL}}} - \frac{1}{\widehat{\lambda}_{\mathsf{PL}}} \ln \left(1 - \frac{(s-1)\widehat{\lambda}_{\mathsf{PL}} W^{\widehat{\alpha}_{\mathsf{PL}}}}{\widehat{\alpha}_{\mathsf{PL}}(D_{\mathsf{HCS}} - \widehat{\lambda}_{\mathsf{PL}} \sum_{i=1}^{D_{\mathsf{HCS}}} Y_i^{\widehat{\alpha}_{\mathsf{PL}}}) + 1} \right) \right)^{1/\widehat{\alpha}_{\mathsf{PL}}},$$

where $\widehat{\alpha}_{\mathsf{PL}}$ and $\widehat{\lambda}_{\mathsf{PL}}$ are obtained from the corresponding normal equations.

Remark 3.5. Clearly, the preceding approach can be applied to other lifetimes in the same manner. To our knowledge, maximum likelihood prediction has been discussed for Weibull distributed lifetimes under Type-I hybrid censoring in Asgharzadeh et al. (2015). For generalized exponential distributions, a similar study can be found in Valiollahi et al. (2017). For the log-normal distribution, Sen et al. (2018) discussed the problem briefly. A generalized Type-II progressive hybrid censoring scheme has been studied in Ateya and Mohammed (2018) for likelihood prediction with Burr-XII distributions.

3.1.3. Minimal repair and record data

Exponential distribution Minimal repair data of $N \in \mathbb{N}$ items is distributionally equivalent to N-record data (see Kamps (1995), Section 1.5) or a superposition of non-homogeneous Poisson processes (see Gupta and Kirmani (1988), Berzborn and Cramer (2023)).

Hence, in this case $n = \infty$ for the formulation of the hybrid censoring scheme and for the parameters we have $\gamma_i = N$, $i \in \mathbb{N}$, and by Lemma 2.1 (ix)

$$\mathbf{G}_{s,s}^{s,0}[z \mid (\boldsymbol{N}+\mathbf{1})^{*s}] = \frac{1}{(s-1)!} z^N (-\ln(z))^{s-1}, \quad z \in (0,1), \tag{3.14}$$

with unique mode (see Volovskiy (2018), Corollary 3.3 and Corollary 3.12(i))

$$\tilde{z}[(\boldsymbol{N}+\mathbf{1})^{*s}] = \exp\left\{-\frac{s-1}{N}\right\}.$$

Consequently, the MLP is given by

$$\widehat{Z}_{(D_{\mathsf{HCS}}+s)} = W_{\mathsf{HCS}} + \theta\frac{s-1}{N}. \tag{3.15}$$

Furthermore, the functional form of $\widehat{\theta}_{\mathsf{PL}}$ is equivalent, with the total time on test statistic simplifying to

$$\mathsf{TTT}_{D_{\mathsf{HCS}}} = N\, W_{\mathsf{HCS}}.$$

We note that the plug-in estimator for unknown parameter $\theta > 0$ has a particularly simple form in this case, given by

$$\widehat{Z}_{(D_{\mathsf{HCS}}+s)}(\widehat{\theta}_{\mathsf{PL}}) = \frac{D_{\mathsf{HCS}}+s}{D_{\mathsf{HCS}}+1} W_{\mathsf{HCS}}.$$

Remark 3.6. Similarly to the comments in the preceding section, we can use the representation in (3.14) and the fact that $\gamma_i - \gamma_{i+1} - 1 = -1$ in this setting, so that (2.11) simplifies to

$$\mathscr{L}_{\mathsf{PLF}}(x_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell) = \frac{N^{\ell+s}}{(s-1)!}\left(\prod_{i=1}^{\ell}\frac{f_{\boldsymbol{\theta}}(x_i)}{\overline{F}_{\boldsymbol{\theta}}(x_i)}\right)\overline{F}_{\boldsymbol{\theta}}^{N-1}(x_{\ell+s}) f_{\boldsymbol{\theta}}(x_{\ell+s})\left(-\ln\left(\frac{\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})}{\overline{F}_{\boldsymbol{\theta}}(w)}\right)\right)^{s-1}.$$

Again, results for the exponential distribution under Type-II censoring and $N = 1$ can be found in Kaminsky and Rhodin (1985) and for Type-I hybrid censoring in Ebrahimi (1992) using the terminology 'with replacement sampling'.

Weibull distribution From Remark 3.6 we conclude that the PLF in case of an underlying Weibull distribution is given by

$$\begin{aligned}\mathscr{L}_{\mathsf{PL}}(y_{\ell+s}, \alpha, \lambda \mid \mathbf{y}_\ell, \ell) &= \frac{N^{\ell+s}}{(s-1)!}\left(\prod_{i=1}^{\ell}\lambda\alpha y_i^{\alpha-1}\right) e^{-\lambda N y_{\ell+s}^{\alpha}}\, \alpha\lambda y_{\ell+s}^{\alpha-1}\left(\lambda\left(y_{\ell+s}^{\alpha} - w^{\alpha}\right)\right)^{s-1} \\ &\propto (\lambda N)^{\ell+s}\,\alpha^{\ell+1} e^{-\lambda N y_{\ell+s}^{\alpha}}\left(\prod_{i=1}^{\ell} y_i\right)^{\alpha-1} y_{\ell+s}^{\alpha-1}\left(y_{\ell+s}^{\alpha} - w^{\alpha}\right)^{s-1}.\end{aligned} \tag{3.16}$$

Remark 3.7.

1. Note that with the reparametrization $\beta = N\lambda$, say, this is equivalent to the predictive likelihood for record values (i.e. $N = 1$) based on the Weibull distribution with parameter vector $(\alpha, \beta) \in \mathbb{R}^2_{>0}$.
2. While hybrid censoring does not seem to have been considered in the literature so far, prediction based on Type-II censored data has already been considered for the three-parameter Weibull distribution (with additional shift parameter θ) in Raqab et al. (2019) as well as in Volovskiy (2018), Section 5.3.5, and Volovskiy and Kamps (2023b) with a different parametrization.

3. Letting the shift parameter be equal to zero, we see that the functional form in (3.16) is very similar: In case $w = y_\ell$, this is (unsurprisingly, see Remark 2.1) equivalent to the predictive likelihood based on a Type-II censored sample of size ℓ. In case $w > y_\ell$, up to the missing factor $\alpha w^{\alpha-1}$ this is equivalent to the predictive likelihood for a Type-II censored sample of size $\ell+1$ with largest observation w. Hence, subsequent arguments along the same lines yield results very similar to those given in Raqab et al. (2019) and Volovskiy (2018).

Equation (3.16) as a function of $\lambda > 0$ is proportional to $G(\lambda) = \lambda^{\ell+s}\, e^{-\lambda N y_{\ell+s}^\alpha}$, which is easily seen to have a unique maximum at

$$\widehat{\lambda} = \frac{\ell+s}{N y_{\ell+s}^\alpha}. \tag{3.17}$$

Plugging this into (3.16) yields

$$\begin{aligned}\mathscr{L}_{\mathsf{PL}}(y_{\ell+s}, \alpha, \widehat{\lambda} \mid \mathbf{y}_\ell, \ell) &\propto y_{\ell+s}^{-\alpha(\ell+s)+\alpha-1} \alpha^{\ell+1} \left(\prod_{i=1}^{\ell} y_i\right)^{\alpha-1} \left(y_{\ell+s}^\alpha - w^\alpha\right)^{s-1} \\ &= \alpha^{\ell+1} \left(\prod_{i=1}^{\ell} y_i\right)^{\alpha-1} w^{-\alpha\ell-1} \left(\left(\frac{w}{y_{\ell+s}}\right)^\alpha\right)^{\ell+1/\alpha} \left(1 - \left(\frac{w}{y_{\ell+s}}\right)^\alpha\right)^{s-1}, \end{aligned} \tag{3.18}$$

which, as a function of $y_{\ell+s}$, is proportional to

$$H(y_{\ell+s}) \propto \left(\left(\frac{w}{y_{\ell+s}}\right)^\alpha\right)^{\ell+1/\alpha} \left(1 - \left(\frac{w}{y_{\ell+s}}\right)^\alpha\right)^{s-1}.$$

By properties of the beta distribution, the maximum is attained for

$$\frac{w^\alpha}{y_{\ell+s}^\alpha} = \frac{\ell+1/\alpha}{\ell+1/\alpha+s-1} \iff y_{\ell+s} = \left(\frac{\alpha(\ell+s-1)+1}{\alpha\ell+1}\right)^{1/\alpha} w. \tag{3.19}$$

Denoting the solution by $\widehat{y}$ and plugging this into (3.18), we arrive at

$$\mathscr{L}_{\mathsf{PL}}(\widehat{y}, \alpha, \widehat{\lambda} \mid \mathbf{y}_\ell, \ell) \propto \alpha^{\ell+1} \left(\prod_{i=1}^{\ell} y_i\right)^{\alpha-1} w^{-\alpha\ell-1} \left(\frac{\alpha\ell+1}{\alpha(\ell+s-1)+1}\right)^{\ell+1/\alpha} \left(\frac{\alpha(s-1)}{\alpha(\ell+s-1)+1}\right)^{s-1},$$

from which we see that the estimate $\widehat{\alpha}_{\mathsf{PL}}$ for α is defined by solutions of the equation

$$\alpha \left(\sum_{i=1}^{D_{\mathsf{HCS}}} \ln(Y_{(i)}) - D_{\mathsf{HCS}} \ln(W_{\mathsf{HCS}})\right) + (D_{\mathsf{HCS}}+1) + \frac{1}{\alpha} \ln\left(\frac{(D_{\mathsf{HCS}}+s-1)\alpha+1}{\alpha D_{\mathsf{HCS}}+1}\right) = 0.$$

Uniqueness for each realization follows by the same arguments as in Raqab et al. (2019) and Volovskiy (2018). Finally, we obtain

$$\widehat{Y}_{(D_{\mathsf{HCS}}+s)} = \left(\frac{\widehat{\alpha}_{\mathsf{PL}}(D_{\mathsf{HCS}}+s-1)+1}{\widehat{\alpha}_{\mathsf{PL}} D_{\mathsf{HCS}}+1}\right)^{1/\widehat{\alpha}_{\mathsf{PL}}} W_{\mathsf{HCS}}, \quad \widehat{\lambda}_{\mathsf{PL}} = \frac{(D_{\mathsf{HCS}}+s)(\widehat{\alpha}_{\mathsf{PL}} D_{\mathsf{HCS}}+1)}{N(\widehat{\alpha}_{\mathsf{PL}}(D_{\mathsf{HCS}}+s-1)+1)} W_{\mathsf{HCS}}^{-\widehat{\alpha}_{\mathsf{PL}}}.$$

Remark 3.8. We observe that the obtained estimators are again functionally equivalent to the ones obtained for Type-II censoring in Raqab et al. (2019) (and similarly to those reported in Volovskiy (2018) and Volovskiy and Kamps (2023a) after an appropriate reparametrization) if we substitute W_{HCS} for the largest observed record and the random counter D_{HCS} for the fixed number of observations. Furthermore, for $\alpha = 1$, we can easily see that the estimators reduce to the ones derived for the exponential distribution in the previous section with $\theta = 1/\lambda$, as expected. Finally, alternative representations are given by

$$\widehat{Y}_{(D_{\mathsf{HCS}}+s)} = \left(\frac{D_{\mathsf{HCS}}+s}{N\widehat{\lambda}_{\mathsf{PL}}}\right)^{1/\widehat{\alpha}_{\mathsf{PL}}}, \qquad \widehat{\lambda}_{\mathsf{PL}} = \frac{D_{\mathsf{HCS}}+s}{N\widehat{Y}_{(D_{\mathsf{HCS}}+s)}}.$$

3.1.4. Generalized order statistics

As for progressively Type-II censored order statistics, we get the same MLP for underlying exponential distributions since the density function of uniform generalized order statistics is log-concave and, thus, unimodal (see Cramer (2004), Theorem 2.2; see also Chen et al. (2009)). For Weibull distributions, the results can also be established provided that the ordered parameters $\gamma_{1:s;\ell}, \dots, \gamma_{s:s;\ell}$ of $\gamma_{\ell+1}, \dots, \gamma_{\ell+s}$ satisfy the constraints $\gamma_{1:s;\ell} \geq 1$, $\gamma_{k:s;\ell} - \gamma_{k-1:s;\ell} \geq 1$, $k \in \{2, \dots, s\}$ (see Cramer (2004), Corollary 4.2).

In case of so-called m-generalized order statistics, that is, $\gamma_i = k + (m+1)(n-i)$, $i = 1, \dots, n$, with $m > -1$ and $k > 0$, the parameters satisfy the identities $\gamma_i - \gamma_{i+1} = m+1$, $i \in \{1, \dots, n-1\}$. Hence, the unique mode of $z \mapsto \mathbf{G}^{s,0}_{s,s}[z \mid \gamma_{\ell+1\dots\ell+s} + \mathbf{1}^{*s}]$ is given by (see Volovskiy, 2018, Corollary 3.3, Corollary 3.12 and Lemma A.4)

$$\tilde{z}[\gamma_{\ell+1\dots\ell+s} + \mathbf{1}^{*s}] = \left(\frac{\gamma_{\ell+s}}{\gamma_{\ell+1}}\right)^{1/(m+1)} = \left(\frac{k+(m+1)(n-\ell-s)}{k+(m+1)(n-\ell-1)}\right)^{1/(m+1)}. \tag{3.20}$$

Therefore, we conclude directly from (3.20) that the MLP is given by

$$\begin{aligned}\widehat{Z}_{(D_{\mathsf{HCS}}+s)} &= W_{\mathsf{HCS}} - \frac{\theta}{m+1}\ln\left(\frac{k+(m+1)(n-D_{\mathsf{HCS}}-s)}{k+(m+1)(n-D_{\mathsf{HCS}}-1)}\right)\\ &= \begin{cases} W_{\mathsf{HCS}}, & s=1,\\ W_{\mathsf{HCS}} - \frac{\theta}{m+1}\ln\left(\frac{k+(m+1)(n-D_{\mathsf{HCS}}-s)}{k+(m+1)(n-D_{\mathsf{HCS}}-1)}\right), & s\geq 2.\end{cases}\end{aligned} \tag{3.21}$$

Furthermore, in case of an unknown parameter $\theta > 0$, the functional form of the PMLE $\widehat{\theta}_{\mathsf{PL}}$ is equivalent with the total time on test statistic simplifying to

$$\mathsf{TTT}_{D_{\mathsf{HCS}}} = (m+1)\sum_{i=1}^{D_{\mathsf{HCS}}} Z_{(i)} + (k+(m+1)(n-D_{\mathsf{HCS}}-1))\,W_{\mathsf{HCS}}. \tag{3.22}$$

In particular, note that, for $m = 0$ and $k = 1$, the model reduces to the case of order statistics and we find that (3.21) and (3.22) reduce to (3.12) and (3.13), respectively, as expected. For Weibull distributions, we can obtain expressions by analogy with the derivations for order statistics in Section 3.1.2.

Remark 3.9. For $m = -1$ and $k = N$, generalized order statistics lead to the submodel of N-record values (see Dziubdziela and Kopociński (1976), Kamps (1995), p. 33) which have been excluded in the above treatment by the assumption $m > -1$. Note that eq. (3.21) can not be utilized in this case since it would cause a division by zero. However, taking the limit from the right $m \searrow -1$ in (3.21), we can use l'Hôpital's theorem to obtain (3.15).

Remark 3.10. The preceding results can also be applied to other distributions using the same methodology although computational methods will be needed in most cases. In particular, results of Volovskiy and Kamps (2023a) who studied maximum likelihood prediction for record values obtained from Pareto and Lomax distributions can also be utilized under hybrid censoring. An application of point prediction based on Pareto records can be found in Empacher et al. (2025).

3.2. Maximum observed likelihood prediction

An introduction and the first comprehensive application of maximum observed likelihood prediction to models of ordered data is due to Volovskiy (2018) and Volovskiy and Kamps (2020a), wherein estimators based on Type-II censored record values are obtained for quite a collection of various distributions. In contrast, order statistics seem to have not been considered in the literature so far. We will apply procedures inspired by the above discussion to more general censoring schemes based on order statistics and records, respectively.

3.2.1. Order statistics

Note that for this submodel, according to Remark 3.4 and with

$$f^{X_{(\ell+s)}}(x_{\ell+s}) = \frac{n!}{(\ell+s-1)!(n-\ell-s)!} F^{\ell+s-1}(x_{\ell+s})(1-F(x_{\ell+s})^{n-\ell-s} f(x_{\ell+s})$$

we have

$$\frac{f^{X_{(j)},1\le j\le D_{\mathsf{HCS}},D_{\mathsf{HCS}},X_{(D_{\mathsf{HCS}}+s)}}(\mathbf{x}_\ell,\ell,x_{\ell+s})}{f^{X_{(\ell+s)}}(x_{\ell+s})} = \frac{(\ell+s-1)!}{(s-1)!}\left(\prod_{i=1}^{\ell}\frac{f(x_i)}{F(x_{\ell+s})}\right)\left(1-\frac{F(w)}{F(x_{\ell+s})}\right)^{s-1},$$
$$\boldsymbol{\alpha}(F) < x_1 \le \cdots \le x_\ell \le w \le x_{\ell+s} < \boldsymbol{\omega}(F),$$

whose functional form agrees with the density function of the first ℓ hybrid censored ordered statistics from a sample of size $\ell+s-1$, with parent distribution truncated on the right at $x_{\ell+s}$ and realized test duration w. Hence, we consider maximization of the predictive likelihood

$$\mathscr{L}_{\mathsf{OPL}}(x_{\ell+s},\boldsymbol{\theta}\mid \mathbf{x}_\ell,\ell) \propto \left(\prod_{i=1}^{\ell}\frac{f_{\boldsymbol{\theta}}(x_i)}{F_{\boldsymbol{\theta}}(x_{\ell+s})}\right)\left(1-\frac{F_{\boldsymbol{\theta}}(w)}{F_{\boldsymbol{\theta}}(x_{\ell+s})}\right)^{s-1}$$
$$= \left(\prod_{i=1}^{\ell}\frac{f_{\boldsymbol{\theta}}(x_i)}{F_{\boldsymbol{\theta}}(w)}\right)\left(\frac{F_{\boldsymbol{\theta}}(w)}{F_{\boldsymbol{\theta}}(x_{\ell+s})}\right)^{\ell}\left(1-\frac{F_{\boldsymbol{\theta}}(w)}{F_{\boldsymbol{\theta}}(x_{\ell+s})}\right)^{s-1}. \tag{3.23}$$

With properties of the beta distribution, the mode is attained for fixed $\boldsymbol{\theta} \in \Theta$ by

$$\frac{F_{\boldsymbol{\theta}}(w)}{F_{\boldsymbol{\theta}}(x_{\ell+s})} = \frac{\ell}{\ell+s-1} \iff x_{\ell+s} = F_{\boldsymbol{\theta}}^{-1}\left(\frac{\ell+s-1}{\ell} F_{\boldsymbol{\theta}}(w)\right).$$

Therefore, we get the MOLP

$$\widehat{x}_{\ell+s} = F_{\boldsymbol{\theta}}^{-1}\Big(\Big(1+\frac{s-1}{\ell}\Big)F_{\boldsymbol{\theta}}(w)\Big), \tag{3.24}$$

which can be considered as a regression type predictor. The expression in (3.24) reminds one of the quantile regression predictions discussed in Navarro (2021), Navarro and Buono (2023), and Buono et al. (2024) where the reliability function $\overline{F}$ is replaced by the cumulative distribution function F. An illustration of the prediction process is provided in Figure 3.1.

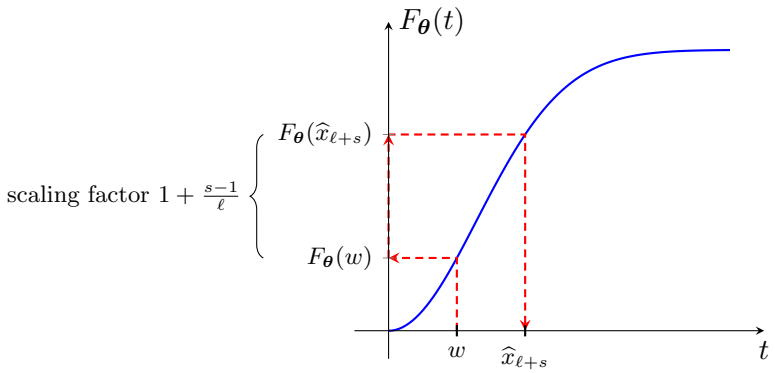


Figure 3.1: Generation of the regression type predictor in (3.24).

Furthermore, substituting the expression in (3.24) into $\mathscr{L}_{\mathsf{OPL}}(x_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell)$ as in (3.23), we get

$$\mathscr{L}_{\mathsf{OPL}}(\widehat{x}_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell) \propto \prod_{i=1}^{\ell} \frac{f_{\boldsymbol{\theta}}(x_i)}{F_{\boldsymbol{\theta}}(w)}. \tag{3.25}$$

Hence, finding the PMOLE of $\boldsymbol{\theta}$ amounts to finding the MLE based on a full sample of order statistics with sample size ℓ and parent distribution truncated on the right at w. In the following, we denote by $c(\ell, s) = 1 + \frac{s-1}{\ell}$ the scaling factor in (3.24). Note that, for $s = 1$, we have $c(\ell, s) = 1$ and hence $\widehat{X}_{(D_{\mathsf{HCS}}+1)} = W_{\mathsf{HCS}}$.

Exponential distribution Since the quantile function is given by $F^{-1}(p) = -\theta \ln(1-p)$, $p \in [0,1]$, we find

$$\widehat{Z}_{(D_{\mathsf{HCS}}+s)} = \begin{cases} W_{\mathsf{HCS}}, & s = 1, \\ -\theta \ln\Big\{1 - c(D_{\mathsf{HCS}}, s)\left(1 - \exp\left(-\frac{1}{\theta} W_{\mathsf{HCS}}\right)\right)\Big\}, & s \geq 2, \end{cases}$$

in case $\theta > 0$ is known. If $\theta > 0$ is unknown, we find from (3.25) that the corresponding normal equation is given by

$$\overline{z} - \theta + \frac{w}{e^{w/\theta} - 1} = 0 \quad \text{with} \quad \overline{z} = \frac{1}{\ell}\sum_{i=1}^{\ell} z_i. \tag{3.26}$$

From, e.g., Al-Athari (2008), we conclude that (3.26) has a unique solution if and only if $0 < \overline{z} < w/2$. Clearly, the maximum likelihood predictor for unknown $\theta > 0$ also does not exist for $\overline{z} \geq w/2$. However, we can use another estimator instead of the PMOLE in order to get a predictor.

Weibull distribution Since the quantile function is given by $F^{-1}(p) = \left(-\frac{1}{\lambda}\ln(1-p)\right)^{1/\alpha}$, $p \in [0,1]$, we find

$$\widehat{Y}_{(D_{\mathsf{HCS}}+s)} = \begin{cases} W_{\mathsf{HCS}}, & s = 1, \\ \left(-\frac{1}{\lambda} \ln\left\{1 - c(D_{\mathsf{HCS}}, s)\left(1 - \exp\left(-\lambda\, W_{\mathsf{HCS}}^{\alpha}\right)\right)\right\}\right)^{1/\alpha}, & s \geq 2, \end{cases}$$

in case $\alpha, \lambda > 0$ are known. In case the distributional parameters are unknown, we conclude from, e.g., Mittal and Dahiya (1989) that the problem of potential non-existence of estimators for α and λ persists.

Remark 3.11. To alleviate the problem of potential non-existence of PMOLEs one might consider using other estimators for the distributional parameters, with the obvious choice being the MLEs based on $\mathscr{L}_{\mathsf{HCS}}$. For the exponential distribution we have $\widehat{\theta} = \frac{\mathsf{TTT}_{D_{\mathsf{HCS}}}}{D_{\mathsf{HCS}}}$ provided $D_{\mathsf{HCS}} > 0$ and with $\mathsf{TTT}_{D_{\mathsf{HCS}}}$ as in (3.13). For the Weibull distribution, existence and uniqueness for Type-I and Type-II censoring was shown in Balakrishnan and Kateri (2008). In view of the decomposition into modules of Type-I and Type-II this readily extends to any hybrid censoring scheme.

3.2.2. Minimal repair and record data

Note that, according to Remark 3.6 and with

$$f^{X_{(\ell+s)}}(x_{\ell+s}) = \frac{N^{\ell+s}}{(\ell+s-1)!}\overline{F}^{N-1}(x_{\ell+s})(-\ln(\overline{F}(x_{\ell+s}))^{\ell+s-1} f(x_{\ell+s}),$$

we have for this submodel

$$\frac{f^{X_{(j)},1\leq j\leq D_{\mathsf{HCS}},D_{\mathsf{HCS}},X_{(D_{\mathsf{HCS}}+s)}}(\mathbf{x}_\ell,\ell,x_{\ell+s})}{f^{X_{(\ell+s)}}(x_{\ell+s})} = \frac{(\ell+s-1)!}{(s-1)!}\left(\prod_{i=1}^{\ell}\frac{f(x_i)}{\overline{F}(x_i)(-\ln(\overline{F}(x_{\ell+s})))}\right)\left(1-\frac{\ln(\overline{F}(w))}{\ln(\overline{F}(x_{\ell+s}))}\right)^{s-1},$$
$$\boldsymbol{\alpha}(F) < x_1 \leq \ldots \leq x_\ell \leq w \leq x_{\ell+s} < \boldsymbol{\omega}(F),$$

whose functional form agrees with the density function of the first ℓ order statistics from a sample of size $\ell+s-1$, baseline distribution

$$G(x) = \begin{cases} \dfrac{\ln(\overline{F}(x))}{\ln(\overline{F}(x_{\ell+s}))}, & x \leq x_{\ell+s}, \\ 1, & x > x_{\ell+s}, \end{cases}$$

realized test duration w, and corresponding likelihood function

$$\mathscr{L}_{\mathsf{OPL}}(x_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell) \propto \left(\prod_{i=1}^{\ell} \frac{f_{\boldsymbol{\theta}}(x_i)}{\overline{F}_{\boldsymbol{\theta}}(x_i)(-\ln(\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s})))} \right) \left(1 - \frac{\ln(\overline{F}_{\boldsymbol{\theta}}(w))}{\ln(\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s}))} \right)^{s-1}. \tag{3.27}$$

Hence, we immediately conclude from the preceding section that, for $\boldsymbol{\theta} \in \Theta$ fixed, the mode is attained for

$$\frac{\ln(\overline{F}_{\boldsymbol{\theta}}(w))}{\ln(\overline{F}_{\boldsymbol{\theta}}(x_{\ell+s}))} = \frac{\ell}{\ell+s-1} \iff x_{\ell+s} = F_{\boldsymbol{\theta}}^{-1}\big(1 - \overline{F}_{\boldsymbol{\theta}}^{c(\ell,s)}(w)\big).$$

Therefore, we get the MOLP

$$\widehat{x}_{\ell+s} = F_{\boldsymbol{\theta}}^{-1}(1 - \overline{F}_{\boldsymbol{\theta}}^{c(\ell,s)}(w)). \tag{3.28}$$

Furthermore, substituting the expression in (3.28) into $\mathscr{L}_{\mathsf{OPL}}(x_{\ell+s}, \boldsymbol{\theta} \mid \mathbf{x}_\ell, \ell)$ as in (3.27) we get analogously

$$\Psi(\boldsymbol{\theta} | \mathbf{x}_\ell, \ell) \propto \prod_{i=1}^{\ell} \frac{f_{\boldsymbol{\theta}}(x_i)}{\overline{F}_{\boldsymbol{\theta}}(x_i)(-\ln(\overline{F}_{\boldsymbol{\theta}}(w)))}$$

as the objective function for $\boldsymbol{\theta} \in \Theta$ in case of unknown distributional parameters.

Remark 3.12.

1. The distributional equivalence to order statistics based on G is known from Keseling (1999), Remark 1.17 (in the context of prediction, see also Volovskiy (2018), Remark 5.2) and it is independent of N.
2. The preceding results could therefore also have been obtained by analogy with Volovskiy (2018), Theorem 5.3. In fact, if $w = x_\ell$, the OPL, MOLP and the objective function Ψ agree with the corresponding ones therein, respectively, since in this case the realized sample again reduces to one of Type-II.
3. Similar comments hold for subsequent presentations of estimators for distributional parameters.
4. Comparing these results with the ones obtained for order statistics in the preceding section, we can see that the predictors for order statistics can be considered as quantile regression type predictors based on the arithmetic mean, while the corresponding ones for records can be considered as quantile regression type predictors based on the geometric mean.
5. The latter observation has been reported in Volovskiy (2018) to introduce the so-called maximum product of spacings prediction procedure, related to maximum product of spacings estimation independently introduced by Cheng and Amin (1983) and Ranneby (1984) and also mentioned in Raqab et al. (2019). Further resources include Volovskiy and Kamps (2020b), Empacher et al. (2023), Empacher et al. (2025).

Exponential distribution In this case, we conclude from (3.28) that

$$\widehat{Z}_{(D_{\mathsf{HCS}}+s)} = \begin{cases} W_{\mathsf{HCS}}, & s = 1, \\ c(D_{\mathsf{HCS}}, s)\, W_{\mathsf{HCS}}, & s \geq 2, \end{cases}$$

independent of $\theta > 0$. Similarly $\Psi(\theta | \mathbf{z}_\ell, \ell) = w^{-\ell}$ is independent of $\theta > 0$, so that no estimator can be obtained.

Weibull distribution In this case, we conclude from (3.28) that

$$\widehat{Y}_{(D_{\mathsf{HCS}}+s)} = \begin{cases} W_{\mathsf{HCS}}, & s = 1, \\ c(D_{\mathsf{HCS}}, s)^{1/\alpha} W_{\mathsf{HCS}}, & s \geq 2, \end{cases}$$

independent of $\lambda > 0$. Similarly

$$\Psi(\alpha, \lambda \mid \mathbf{y}_\ell, \ell) = \alpha^\ell \left(\prod_{i=1}^{\ell} \frac{y_i}{w} \right)^{\alpha-1} w^{-\ell} \propto \alpha^\ell \left(\prod_{i=1}^{\ell} \frac{y_i}{w} \right)^{\alpha-1} \tag{3.29}$$

independent of $\lambda > 0$ and with

$$\widehat{\alpha}_{MOL} = \frac{D_{\mathsf{HCS}}}{D_{\mathsf{HCS}} \ln(W_{\mathsf{HCS}}) - \sum_{i=1}^{D_{\mathsf{HCS}}} \ln(Y_{(i)})}.$$

3.3. Generalized order statistics

The results obtained for the order statistics' and record model are based on the property that the right-truncated distribution at w can be considered as a distribution of order statistics obtained from a right truncated distribution. This special property allows us to find explicit representations of the MOLP and the PMOLE. However, this is not possible for other models. For instance, as pointed out in Iliopoulos and Balakrishnan (2009), the distribution of progressively Type-II censored order statistics from a right-truncated distribution can be seen as a mixture of distributions of order statistics truncated on the right. An analogous result can be obtained for generalized order statistics using the same proof but bearing in mind that the parameters are arbitrary positive values. As a result, the corresponding distribution is a mixture distribution (not necessarily of order statistics' distributions as pointed out in Iliopoulos and Balakrishnan (2009), Section 5). Therefore, the maximization to obtain MOLP and PMOLE has been done computationally in this general setting.

4. ILLUSTRATIVE EXAMPLES

4.1. Order statistics: Gauge data

In order to illustrate the method, we apply the censoring and point prediction approaches to the data given in Table 1.1. To compute the predicted values, we consider a Weibull distribution after subtracting 1.75 from each measurement as has been done in Asgharzadeh et al. (2015), who pointed out that, after this shift transformation, the data can be considered as Weibull distributed. We consider all the hybrid censoring schemes given in Table 4.1 which lead to the hybrid censored data sets given in Table 4.2. Notice that, even though eleven different hybrid censoring schemes are applied, the prediction problem reduces to a discussion of four distinct data situations only. In two of these cases, the data is effectively Type-I or Type-II censored, respectively. This illustrates results from previous

sections that predictions obtained from methods presented here only depend on the effectively observed data and stopping time w (which is reported in the last column, respectively). The original design of the hybrid censoring scheme does not have an impact on the point prediction. For brevity, we restrict ourselves to prediction of the next four order statistics in the corresponding samples. Of course, the predictions for higher values of s can be obtained in the same manner. For better comparison, the predicted values based on the maximum likelihood approach are printed in the same table as an addition to the observed hybrid censored data. The predicted values are indicated by an asterix*. In the row directly below these predicted values, the point predictions based on the maximum observed likelihood approach are displayed. These values are indicated by a dagger†. Note that since the PMOLEs for α and λ do not exist for these data sets, respectively, we have instead substituted the MLEs (obtained in Asgharzadeh et al. (2015)) as suggested in Remark 3.11. For comparison, the original data (up to the 19th largest measurement) is given in the first row of Table 4.2. Details on the hybrid censoring schemes can be found in Cramer (2025) and Balakrishnan et al. (2023) as well as the references therein.

Hybrid censoring scheme			censoring numbers m_i				thresholds T_j		
		k	1	2	3	4	1	2	3
Type I HCS	Type-I	1	63	0			2.3		
	Type-I hybrid	1	15	0			2.3		
	gen. Type-II hybrid	2	63	15	0		2.3	2.5	
	uni. Type-I hybrid	2	15	10	0		2.3	2.5	
	uni. Type-III hybrid	3	63	15	10	0	2.3	2.5	2.8
Type II HCS	Type-II	0	15						
	Type-II hybrid	1	63	15			2.3		
	gen. Type-I hybrid	1	15	10			2.3		
	uni. Type-II hybrid	2	63	15	10		2.3	2.5	
	uni. Type-IV hybrid	2	15	10	8		2.3	2.5	
	gen. uni. Type hybrid	3	63	15	10	8	2.3	2.5	2.8

Table 4.1: Hybrid censoring schemes with parameter values used in the prediction based on the strength data.

The predicted values are in line with those given in Asgharzadeh et al. (2015, Table 2) except for those obtained for $s = 1$. Since the theoretical results yield the test duration as prediction in this case, the values for the MLP presented in Asgharzadeh et al. (2015, Table 2) are in error. Furthermore, Asgharzadeh et al. (2015) presented the PMLE as an overall PMLE, implicitly claiming that they hold for any s. However, due to the simultaneous optimization, the PMLEs depend on s so that they vary with the desired predicted value. In Table 4.3, the PMLEs are provided for any of the four data situations given in Table 4.2.

4.2. Minimal repair times: Load haul dump data

In order to illustrate the approach for minimal repair data, we consider data of hydraulic subsystems of load-haul dump machines used in mining taken from Kumar and Klefsjö (1992). In particular, we consider system LHD9 assuming Weibull distributed lifetimes (predicted values indicated by a dagger†), since this yields more reasonable predictions compared to

assuming exponential life times (see also predicted values in Table 4.4 which are provided for comparison; they are indicated by a double dagger‡). For the selection of the system, we refer to the remarks provided in Beutner and Cramer (2011). As in Berzborn and Cramer (2023), we apply the Type-I hybrid censoring scheme with $m_1 = 5$ and Type-II hybrid censoring scheme with $m_2 = 5$ as well as $T_1 = 1000$, respectively, leading to the censored data given in Table 4.4. The predicted values seem to be somewhat optimistic since they exceed the actually observed values for $s > 1$. This applies particularly to the Type-I hybrid censoring scheme. However, it should be noted that the point prediction is based on five and six observed values only, respectively. Taking this into account the predictions yield reasonable values.

5. CONCLUSION

In this paper, we have introduced a general method to construct hybrid censoring designs of any dimension by introducing censoring numbers and censoring thresholds. It is shown that the hybrid censoring schemes discussed in the literature can be subsumed in this approach. As a consequence, we have pointed out that point predictions based on such hybrid censoring schemes can be easily obtained. In particular, it is demonstrated that both the maximum likelihood prediction and the maximum observed likelihood prediction approach can be applied to data observed from any (known) hybrid censoring scheme using computational results already obtained for the Type-I and Type-II censoring case, respectively. The methods are especially applied to exponential and Weibull distributed lifetimes but, clearly, they can be directly applied to other lifetime distributions, too. The results are illustrated for order statistics' and minimal repair data by applying various hybrid censoring schemes to given data. Then, the predicted values are compared with the actually observed ones. This yields good predicted values even for small observed sample sizes.

However, in order to get a better impression of the quality of the predictions, the prediction outcomes have to be investigated with respect to distance measures. Clearly, this can be done by simulations. Another future project would be to apply the approach to Bayesian prediction as well as to interval prediction. This can be included in the comparison of the discussed prediction methods. Finally, it should be mentioned that the modularization approach will also be useful in other inferential approaches. This will also be subject of future research.

ACKNOWLEDGMENTS

The authors are grateful to an anonymous reviewer for comments and suggestions that improved the structure of the manuscript.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

Al-Athari, F. M. (2008). Estimation of the mean of truncated exponential distribution. *Journal of Mathematics and Statistics*, 4(4):284–288.

Asgharzadeh, A., Valiollahi, R., and Kundu, D. (2015). Prediction for future failures in Weibull distribution under hybrid censoring. *Journal of Statistical Computation and Simulation*, 85(4):824–838.

Ateya, S. and Mohammed, H. (2018). Prediction under Burr-XII distribution based on generalized Type-II progressive hybrid censoring scheme. *Journal of the Egyptian Mathematical Society*, 26(3):491–508.

Badar, M. and Priest, A. (1982). *Statistical aspects of fiber and bundle strength in hybrid composites*, pages 1129–1136. ICCM-IV, Tokyo.

Balakrishnan, N. and Cramer, E. (2014). *The Art of Progressive Censoring. Applications to Reliability and Quality*. Birkhäuser, New York.

Balakrishnan, N., Cramer, E., and Kundu, D. (2023). *Hybrid Censoring Know-How: Designs and Implementations*. Academic Press.

Balakrishnan, N. and Kateri, M. (2008). On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data. *Statistics & Probability Letters*, 78(17):2971–2975.

Balakrishnan, N. and Kundu, D. (2013). Hybrid censoring: Models, inferential results and applications (with discussions). *Comput. Statist. Data Anal.*, 57:166–209.

Belzunce, F., Mercader, J.-A., and Ruiz, J.-M. (2005). Stochastic comparisons of generalized order statistics. *Probab. Engrg. Inform. Sci.*, 19:99–120.

Berzborn, M. and Cramer, E. (2023). Inference for Type-I and Type-II hybrid censored minimal repair and record data. *Journal of Statistical Theory and Practice*, 17(53).

Beutner, E. and Cramer, E. (2011). Confidence intervals for quantiles in a minimal repair set-up. *Int. J. Appl. Math. Statist.*, 24:86–97.

Buono, F., Cramer, E., and Navarro, J. (2024). Predicting future failures in generalized order statistics and related models. *Statistics*, 58(4):918–942.

Burkschat, M., Cramer, E., and Górny, J. (2016). Type-I censored sequential k-out-of-n systems. *Appl. Math. Model.*, 40(19–20):8156–8174.

Chen, H., Xie, H., and Hu, T. (2009). Log-concavity of generalized order statistics. *Statist. Probab. Letters*, 79:396–399.

Cheng, R. C. H. and Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 45(3):394–403.

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. *Ann. Inst. Statist. Math.*, 55(2):319–330.

Cramer, E. (2003). Contributions to Generalized Order Statistics. Habilitation Thesis. University of Oldenburg, Oldenburg.

Cramer, E. (2004). Logconcavity and unimodality of progressively censored order statistics. *Statist. Probab. Letters*, 68:83–90.

Cramer, E. (2006). Dependence structure of generalized order statistics. *Statistics*, 40(5):409–413.

Cramer, E. (2025). Structure of hybrid censoring schemes and its implications. *Metrika*, 88:393–417.

Cramer, E. and Kamps, U. (2003). Marginal distributions of sequential and generalized order statistics. *Metrika*, 58(3):293–310.

Cramer, E., Kamps, U., and Rychlik, T. (2004). Unimodality of uniform generalized order statistics, with applications to mean bounds. *Annals of the Institute of Statistical Mathematics*, 56:183–192.

Draper, N. and Guttman, I. (1987). Bayesian analysis of hybrid life tests with exponential failure times. *Annals of the Institute of Statistical Mathematics*, 39(1):219–225.

Dziubdziela, W. and Kopociński, B. (1976). Limiting properties of the k-th record values. *Appl. Math.*, 15:187–190.

Ebrahimi, N. (1986). Estimating the parameters of an exponential distribution from a hybrid life test. *Journal of Statistical Planning and Inference*, 14(2):255 – 261.

Ebrahimi, N. (1992). Prediction intervals for future failures in the exponential distribution under hybrid censoring. *IEEE Transactions on Reliability*, 41(1):127–132.

Empacher, C., Kamps, U., and Schmiedt, A. B. (2025). Prediction intervals for future Pareto record claims. *European Actuarial Journal*, 15:163–197.

Empacher, C., Kamps, U., and Volovskiy, G. (2023). Statistical prediction of future sports records based on record values. *Stats*, 6(1):131–147.

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

Fairbanks, K., Madsen, R., and Dykstra, R. (1982). A confidence interval for an exponential parameter from a hybrid life test. *Journal of the American Statistical Association*, 77(377):137–140.

Gupta, R. C. and Kirmani, S. N. U. A. (1988). Closure and monotonicity properties of nonhomogeneous Poisson processes and record values. *Probability in the Engineering and Informational Sciences*, 2:475–484.

Górny, J. and Cramer, E. (2018a). Exact inference for a new flexible hybrid censoring scheme. *Journal of the Indian Society for Probability and Statistics*, 19(1):169–199.

Górny, J. and Cramer, E. (2018b). Modularization of hybrid censoring schemes and its application to unified progressive hybrid censoring. *Metrika*, 81:173–210.

Iliopoulos, G. and Balakrishnan, N. (2009). Conditional independence of blocked ordered data. *Statist. Probab. Letters*, 79:1008–1015.

Izadi, M. and Khaledi, B.-E. (2007). Progressive Type II censored order statistics and their concomitants: some stochastic comparisons results. *J. Iran. Stat. Soc. (JIRSS)*, 6:111–124.

Kaminsky, K. S. and Nelson, P. I. (1998). Prediction of order statistics. In Balakrishnan, N. and Rao, C. R., editors, *Handbook of Statistics: Applications, Vol. 17*, pages 431–450, Amsterdam. Elsevier.

Kaminsky, K. S. and Rhodin, L. S. (1985). Maximum likelihood prediction. *Annals of the Institute of Statistical Mathematics*, 37:507–517.

Kamps, U. (1995). *A Concept of Generalized Order Statistics*. Teubner.

Kamps, U. and Cramer, E. (2001). On distributions of generalized order statistics. *Statistics*, 35:269–280.

Keseling, C. (1999). *Charakterisierung von Wahrscheinlichkeitsverteilungen durch Verallgemeinerte Ordnungsstatistiken (in German)*. PhD thesis, RWTH Aachen University.

Korwar, R. (2003). On the likelihood ratio order for progressive type II censored order statistics. *Sankhyā A*, 65:793–798.

Kumar, U. and Klefsjö, B. (1992). Reliability analysis of hydraulic systems of LHD machines using the power law process model. *Reliab. Engrg. Syst. Safety*, 35(3):217 – 224.

Kundu, D. and Gupta, R. D. (2006). Estimation of $P[Y < X]$ for Weibull distributions. *IEEE Trans. Reliab.*, 55(2):270–280.

Ling, M. H., Cramer, E., and Bae, S. J. (2026). Statistical inference and optimal design for step-stress accelerated life testing with hybrid group censoring for non-destructive one-shot devices. *Reliability Engineering & System Safety*, 265:111506.

Marshall, A. W. and Olkin, I. (2007). *Life Distributions. Structure of Nonparametric, Semiparametric, and Parametric Families*. Springer, New York, NY.

Mathai, A. M. (1993). *A Handbook of Generalized Special Functions for Statistical and Physical Sciences*. Clarendon Press.

Mittal, M. M. and Dahiya, R. C. (1989). Estimating the parameters of a truncated Weibull distribution. *Communications in Statistics – Theory and Methods*, 18(6):2027–2042.

Navarro, J. (2021). Prediction of record values by using quantile regression curves and distortion functions. *Metrika*, 85(6):675–706.

Navarro, J. and Buono, F. (2023). Predicting future failure times by using quantile regression. *Metrika*, 86:543–576.

Ranneby, B. (1984). The maximum spacing method. An estimation method related to the maximum likelihood method. *Scandinavian Journal of Statistics*, 11(2):93–112.

Raqab, M. Z., Alkhalfan, L. A., Bdair, O. M., and Balakrishnan, N. (2019). Maximum likelihood prediction of records from 3-parameter Weibull distribution and some approximations. *Journal of Computational and Applied Mathematics*, 356:118–132.

Sen, T., Singh, S., and Tripathi, Y. M. (2018). Statistical inference for lognormal distribution with type-I progressive hybrid censored data. *American Journal of Mathematical and Management Sciences*, 38(1):70–95.

Valiollahi, R., Asgharzadeh, A., and Kundu, D. (2017). Prediction of future failures for generalized exponential distribution under Type-I or Type-II hybrid censoring. *Brazilian Journal of Probability and Statistics*, 31(1):41–61.

van Bentum, T. and Cramer, E. (2019). Stochastic monotonicity of MLEs of the mean for exponentially distributed lifetimes under sequential hybrid censoring. *Statist. Probab. Letters*, 148:1–8.

Volovskiy, G. (2018). *Likelihood-Based Prediction in Models of Ordered Data*. PhD thesis, RWTH Aachen University.

Volovskiy, G. and Kamps, U. (2020a). Maximum observed likelihood prediction of future record values. *TEST*, 29(4):1072–1097.

Volovskiy, G. and Kamps, U. (2020b). Maximum product of spacings prediction of future record values. *Metrika*, 83(7):853–868.

Volovskiy, G. and Kamps, U. (2023a). Comparison of likelihood-based predictors of future Pareto and Lomax record values in terms of Pitman closeness. *Communications in Statistics – Theory and Methods*, 52(6):1905–1922.

Volovskiy, G. and Kamps, U. (2023b). Likelihood-based prediction of future Weibull record values. *REVSTAT - Statistical Journal*, 21(3):425–445.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	w
complete	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.518	2.522	2.525	2.532	2.575	2.614	2.616	
Type I hybrid censoring schemes																				
Type-I	1.901	2.132	2.203	2.228	2.257	2.300*	2.334*	2.365*	2.393*											2.300
						2.300†	2.337†	2.371†	2.402†											2.300
Type-I H	1.901	2.132	2.203	2.228	2.257	2.300*	2.334*	2.365*	2.393*											2.300
						2.300†	2.337†	2.371†	2.402†											2.300
g Type-II H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.500*	2.522*	2.543*	2.564*				2.500
													2.5†	2.524†	2.546†	2.568†				2.5
u Type-I H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.445*	2.468*	2.49*	2.511*						2.445
											2.445†	2.470†	2.495†	2.518†						2.445
u Type-III H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.500*	2.522*	2.543*	2.564*				2.500
													2.500†	2.524†	2.546†	2.568†				2.500
Type II hybrid censoring schemes																				
Type-II	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.518	2.522	2.525	2.525*	2.542*	2.558*	2.574*	2.525
																2.525†	2.545†	2.564†	2.583†	2.525
Type-II H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.518	2.522	2.525	2.525*	2.542*	2.558*	2.574*	2.525
																2.525†	2.545†	2.564†	2.583†	2.525
g Type-I H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.445*	2.468*	2.49*	2.511*						2.445
											2.445†	2.47†	2.495†	2.518†						2.445
u Type-II H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.5*	2.522*	2.543*	2.564*				2.5
													2.5†	2.524†	2.546†	2.568†				2.5
u Type-IV H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.445*	2.468*	2.49*	2.511*						2.445
											2.445†	2.47†	2.495†	2.518†						2.445
g u Type H	1.901	2.132	2.203	2.228	2.257	2.350	2.361	2.396	2.397	2.445	2.454	2.474	2.5*	2.522*	2.543*	2.564*				2.5
													2.5†	2.524†	2.546†	2.568†				2.5

Table 4.2: Hybrid censored data sets obtained from hybrid censoring schemes with scheme parameters as in Table 4.1 (actually observed values are highlighted by a gray background). For comparison, the first row provides the first 19 order statistics of the observed data. Values indicated by * and † are predicted by the maximum likelihood and the maximum observed likelihood approach, respectively. For each hybrid censored data set, $s \in \{1, 2, 3, 4\}$ are considered.

PMLE $(\widehat{\alpha}, \widehat{\lambda})$ for (α, λ)	s			
Hybrid censoring schemes	1	2	3	4
Type-I, Type-I H	$(2.954, 0.581)$	$(2.975, 0.578)$	$(3.006, 0.581)$	$(3.027, 0.582)$
g Type-II, u Type-III H, u Type-II H, g u Type H	$(3.027, 0.547)$	$(3.033, 0.546)$	$(3.039, 0.545)$	$(3.044, 0.544)$
u Type-I H, g Type-I H, u Type-IV H	$(3.123, 0.591)$	$(3.131, 0.589)$	$(3.139, 0.588)$	$(3.146, 0.587)$
Type-II, Type-II H	$(3.437, 0.689)$	$(3.441, 0.688)$	$(3.445, 0.686)$	$(3.448, 0.685)$

Table 4.3: PMLE for (α, λ) and $s \in \{1, 2, 3, 4\}$ based on the data situations in Table 4.2 ('H' means hybrid censoring).

	Repair times										
	1	2	3	4	5	6	7	8	9	10	w
System LHD9	249	461	665	847	963	993	1017	1049	1087	1097	
Type-I hybrid	249	461	665	847	963	$963^{\dagger}$	$1058.299^{\dagger}$	$1146.195^{\dagger}$	$1228.213^{\dagger}$		963
						$963^{\ddagger}$	$1155.600^{\ddagger}$	$1348.200^{\ddagger}$	$1540.800^{\ddagger}$		
Type-II hybrid	249	461	665	847	963	993	$1000^{\dagger}$	$1074.13^{\dagger}$	$1142.771^{\dagger}$	$1206.949^{\dagger}$	1000
							$1000^{\ddagger}$	$1166.667^{\ddagger}$	$1333.333^{\ddagger}$	$1500^{\ddagger}$	

Table 4.4: Hybrid censored repair times of hydraulic subsystems of load-haul dump machines used in mining taken from Kumar and Klefsjö (1992) (actually observed values are highlighted by a gray background)). For comparison, the third row provides the first ten minimal repair times of the observed data. All the predicted values were obtained by maximum observed likelihood prediction. A dagger† indicates that the lifetime were assumed Weibull distributed whereas a double dagger‡ indicates that an exponential distribution has been assumed.

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.

Abstract and Indexing services

REVSTAT — Statistical Journal is covered *by Journal Citation Reports - JCR (Clarivate); Current Index to Statistics; Google Scholar; Mathematical Reviews® (MathSciNet®); Zentralblatt für Mathematic; Scimago Journal & Country Rank; Scopus*

Author guidelines

The whole submission and review processes for REVSTAT are conducted exclusively online on the journal's webpage https://revstat.ine.pt/ based in Open Journal System (OJS). Authors intending to submit any work must ***register***, ***login*** and follow the indications choosing ***Submissions***.

REVSTAT — Statistical Journal adopts the COPE guidelines on publication ethics.

Work presentation

- the only working language is English;
- the first page should include the name, ORCID iD (optional), Institution, country, and mail-address of the author(s);
- a summary of fewer than one hundred words, followed by a maximum of six keywords and the MSC 2020 subject classification should be included also in the first page;
- manuscripts should be typed only in black, in double-spacing, with a left margin of at least 3 cm, with numbered lines, and a maximum of 25 pages;
- the title should be with no more than 120 characters (with spaces);
- figures must be a minimum of 300dpi and will be reproduced online as in the original work, however, authors should take into account that the printed version is always in black and grey tones;
- authors are encouraged to submit articles using LaTeX which macros are available at *REVSTAT style*;
- citations in text should be included in the text by name and year in parentheses, as in the following examples: § article title in lowercase (Author 1980); § This theorem was proved later by AuthorB and AuthorC (1990); § This subject has been widely addressed (AuthorA 1990; AuthorB et al. 1995; AuthorA and AuthorB 1998);
- references should be listed in alphabetical order of the author's scientific surname at the end of the article;
- acknowledgments of people, grants or funds should be placed in a short section before the References title page. Note that religious beliefs, ethnic background, citizenship and political orientations of the author(s) are not allowed in the text;

- authors are welcome to suggest one of the Editors or Associate Editors or yet other reviewer expert on the subject providing a complete information, namely: name, affiliation, email, and personal URL or ORCID number in the Comments for the Editor (submission form).

Accepted papers

After final revision and acceptance of an article for publication, authors are requested to provide the corresponding LaTeX file, as in REVSTAT style.

Supplementary files may be included and submitted separately in .tiff, .gif, .jpg, .png .eps, .ps or .pdf format. These supplementary files may be published online along with an article, containing data, programming code, extra figures, or extra proofs, etc; however, REVSTAT is not responsible for any supporting information supplied by the author(s).

Copyright Notice

Upon acceptance of an article, the author(s) will be asked to transfer copyright of the article to the publisher, Statistics Portugal, 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.