Journal of Statistical Research *https://doi.org/10.3329/jsr.v59i2.88073*
2025, Vol. 59, No. 2, pp. 325-342 *ISSN 0256 - 422 X*

A COMPREHENSION OF THE BURR TYPE-II DISTRIBUTION BASED ON RECORDS: PROPERTIES, ESTIMATION AND PREDICTION

ALIREZA PAKGOHAR*
Department of Statistics, Payame Noor University, Tehran, Iran.
Email: a_pakgohar@pnu.ac.ir

PARVIZ NASIRI
Department of Statistics, Payame Noor University, Tehran, Iran.
Email: pnasiri@pnu.ac.ir

SUMMARY

This article introduces the Burr Type-II distribution and derives various point estimates for unknown parameters based on record statistics. It also presents predictions for future record values from a Bayesian perspective, supported by a numerical study to illustrate the results. Furthermore, the article covers the properties of the Burr Type-II distribution and includes the calculation of entropy and extropy functions, as well as an investigation of stochastic orders such as usual stochastic order, likelihood ratio order, and hazard rate order.

Keywords and phrases: Bayes estimation, Maximum likelihood estimation, Mean square error, Prediction, Record value, Stochastic order, Uncertainty measures.

1 Introduction

Investigating the statistical distribution, with a focus on its properties, stochastic ordering, and various information measures, is essential for enhancing the understanding and application of this distribution in statistical modeling and analysis. For example, the following articles have focused on examining the properties of stochastic ordering and information measures: Pakgohar (2024), Alghamdi et al. (2023) and Sanusi et al. (2020). In order to make predictions using the Burr Type-II distribution, statistical techniques are utilized to accurately determine its parameters. By analyzing data and fitting the distribution model to available information, the optimal shape, scale, and threshold parameters can be identified. These parameter estimates are essential for forecasting future occurrences that display a similar distribution pattern.

A common method for estimating parameters in the Burr Type-II distribution is maximum likelihood estimation. This involves identifying the parameter values that maximize the likelihood of observing the given data. Once the parameters have been determined, they can be used to predict future occurrences or events. This predictive capability can be valuable in various applications,

* Corresponding author

such as forecasting extreme weather conditions, fluctuations in financial markets, or failure rates in engineering systems.

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distribution (iid) random variables with cumulative distribution function (cdf), $F(x;\theta)$ and probability density function (pdf) $f(x;\theta)$. Define,

$$Y_m = \max(\min)\{X_1, X_2, \ldots, X_n\}, m \geq 1.$$

Then, X_j is an upper (lower) record value of this sequence if $X_j > (<)Y_{j-1}, j > 1$. If $\{L(m), m \geq 1\}$ is defined by

$$L(1) = 1, L(m) = \min\{j : j > L(m-1), X_j < X_{L(m-1)}\};$$

for $m \geq 2$ then the sequence $X_{L(m)} \geq 1$ provides a sequence of lower record statistics.

The sequence $L(m), m \geq 1$ represents the record times. Chandler (1952) defined the model of record statistics as a model for successive extremes. In a sequence of independent and identically distributed random variables, record statistics can be viewed as order statistics from a sample whose size is determined by the values and the order of occurrence of the observations. These statistics are of interest and important in several real-life problems involving weather, economics, and sports data. For more details and applications in the record values, see Ahsanullah (1995), Arnold et al. (2011), and Nevzorov (2000). In 1942, Burr developed a set of twelve cumulative distribution functions to model lifetime data, and this Burr family includes distributions like Burr Type X, which is useful for analyzing strength and general lifetime data (See Mahto et al. (2018) and Burr (1942)). The family has since expanded to include new categories such as the Inverse Burr-Generalized Family, which provides statistical features like density functions, survival rates, and hazard functions (Osagie et al., 2023). Techniques such as maximum likelihood estimation have been employed to estimate parameters for Burr distributions, with research studies comparing various estimation methods (Ismail (2014) and Hassan et al. (2021)). Moreover, outlier detection techniques like ECOD make use of empirical cumulative distribution functions to pinpoint anomalies within datasets, demonstrating promising outcomes in terms of precision and efficiency (Li et al., 2022).

The two important members of the family are Burr types II and XII. The two important distributions, Burr Type-II and Burr Type-XII, are interrelated through simple transformation. Burr Type-II distribution allows for a wider region for skewness and kurtosis plane, which covers several distributions including the log-logistics, and the Weibull and Burr type XII distributions. For more details see Arabi Belaghi et al. (2014), Laslan and Nadar (2017) and Kumar and Rani (2018).

In this article, we explore several characteristic properties of the Burr Type-II distribution and address the problem of estimation and prediction using lower record values. Specifically, in Section 2, we investigate various properties, including characteristics related to estimation, uncertainty measures, and stochastic orders, providing a comprehensive overview of the distribution's theoretical foundations. In Section 3, we derive both maximum likelihood and Bayes estimates for the unknown parameters based on lower record statistics. The Bayes estimates are computed under two loss functions: squared error and LINEX, offering flexibility in handling different types of estimation errors and biases. In Section 4, we focus on predictive inference by constructing Bayes prediction intervals for future lower records. This section highlights the practical utility of the proposed methods in forecasting and decision-making processes. Finally, in Section 5, we present a numerical example and conduct a Monte Carlo simulation study to demonstrate the effectiveness and applicability of the derived results. These analyses provide empirical evidence supporting the robustness and accuracy of the estimation and prediction techniques discussed in this article.

2 Properties of Burr Type-II and Record Data

Let us consider the cumulative distribution function, probability density function of the Burr Type-II distribution is given, respectively, by

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

$$f(x;\theta,c) = \theta c x^{(c-1)}(1+x^c)^{-(\theta+1)}; \tag{2.2}$$

where the parameters $c>0$ and $\theta>0$ are the shape parameters of the distribution.

2.1 Expectation

The computation of the expected value of X^k plays a fundamental role in statistical estimation, with the moment generating function (MGF) serving as a classical tool for its evaluation. In this work, we propose a direct method to estimate $E[X^k]$, bypassing reliance on the MGF.

By using equation (2.2) the the extended k^{th} raw moment becomes

$$E(X^k) = \int\limits_0^\infty x^k f(x)dx = \theta c\int_0^\infty x^{k+c-1}(1+x^c)^{-(\theta+1)}dx. \tag{2.3}$$

Substituting the term $1+x^c$ into $\frac{1}{1-z}$ and letting $k=pc$ $(p\leq\theta)$ yields

$$E(X^k) = \theta\int_0^1 (z)^p(1-z)^{\theta-p-1}dz = \theta B(p+1,\theta-p) = \theta\frac{\Gamma(p+1)\Gamma(\theta-p)}{\Gamma(\theta+1)} = \frac{1}{\binom{\theta-1}{p}}; \tag{2.4}$$

Furthermore, based on equation (2.3), we obtain

$$\frac{E(X^k)}{\theta c} = \int_0^\infty x^{k+c-1}(1+x^c)^{-(\theta+1)}dx. \tag{2.5}$$

Typically, by substituting $k=c-1$, we obtain

$$\frac{E(X^{c-1})}{\theta c} = \int_0^\infty (1+x^c)^{-(\theta+1)}dx. \tag{2.6}$$

Then, for a non-negative lifetime random variable X distributed by Burr Type-II, we can establish a relationship for the expected value of the random variable (X) raised to the (k^{th}) power, which is given by

$$E(X) = \int_0^\infty \bar{F}(x)dx = \int_0^\infty (1+x^c)^{-\theta}dx = \frac{E(X^{c-1})}{c(\theta-1)}. \tag{2.7}$$

That yields

$$E(X^{c-1}) = c(\theta - 1)E(X).$$

The integral of the survival distribution function raised to the power of k for a random variable X following a Burr Type-II distribution can be expressed as

$$\int_0^\infty (1 + x^c)^{-k\theta} dx = \frac{E(X^{c-1})}{c(k\theta - 1)}.$$

2.2 Uncertainty measures

In the context of the Burr Type-II distribution, simplifying calculations for entropy and extropy can be achieved by setting certain parameters, particularly by setting $c = 1$. This simplification can lead to a clearer understanding of the distribution's properties. By experimenting with different values of the parameter theta and observing its effects on entropy and extropy, researchers have the opportunity to discover valuable insights and relationships within the distribution.

Proposition 2.1. *The Shannon entropy $H(X)$ for the random variable X with the Burr Type-II distribution is given by*

$$\begin{aligned} H(X) &= -\int_0^\infty f(x)ln(f(x))dx \\ &= -\int_0^\infty \theta c x^{(c-1)}(1 + x^c)^{-(\theta+1)} \ln\left(\theta c x^{(c-1)}(1 + x^c)^{-(\theta+1)}\right) dx. \end{aligned} \tag{2.8}$$

Remark 1. Assume that $c = 1$ implies equation (2.8) into

$$\begin{aligned} H(X) &= -\int_0^\infty \theta(1 + x)^{-(\theta+1)} \ln\left(\theta(1 + x)^{-(\theta+1)}\right) dx \\ &= -\ln(\theta) - E(\ln((1 + x)^{-(\theta+1)})) \\ &\overset{Jensen\ inequality}{\geq} -\ln(\theta) - \ln\int_0^\infty \theta(1 + x)^{-2(\theta+1)})dx \\ &= -\ln(\theta) - \ln(\frac{\theta}{2\theta + 1}) = \ln(\frac{2\theta + 1}{\theta^2}). \end{aligned} \tag{2.9}$$

Proposition 2.2. *The extropy measure $J(X)$ for a non-negative random variable X distributed by the Burr Type-II distribution is given by (Lad et al., 2015)*

$$J(X) = -\frac{1}{2}\int_0^\infty (\theta c)^2 x^{2(c-1)}(1 + x^c)^{-2(\theta+1)} dx. \tag{2.10}$$

Remark 2. Let us $c = 1$ then equation (2.10) yields to

$$J(X) = -\frac{1}{2}\frac{\theta^2}{2\theta + 1}. \tag{2.11}$$

Proposition 2.3. *There is a connection between cumulative residual extropy, $CRJ(X)$, and the expected value of X^{c-1}, scaled by the parameters c and θ (Jahanshahi et al., 2020).*

$$CRJ(X) = -\frac{1}{2}\int_0^\infty \bar{F}^2(x)dx = -\frac{E(X^{c-1})}{2c(2\theta-1)}.$$

Remark 3. By setting $c = 1$ in the above integral equation, we obtain

$$CRJ(X) = -\frac{1}{2(2\theta-1)}.$$

Proposition 2.4. *The Rényi entropy for a continuous random variable X with the Burr Type-II distribution is given by the formula (Renner and Wolf, 2004)*

$$H_r(X) = \frac{1}{1-r}\ln\int_0^\infty f^r(x)dx = \frac{1}{1-r}\ln\int_0^\infty (\theta c)^r x^{r(c-1)}(1+x^c)^{-r(\theta+1)}. \tag{2.12}$$

Remark 4. Suppose that $c = 1$ then the equation (2.12) yields to

$$H_r(X) = \frac{1}{1-r}\ln\int_0^\infty \theta^r(1+x)^{-r(\theta+1)} = \frac{1}{1-r}\ln\frac{\theta^r}{r(\theta+1)-1}.$$

Proposition 2.5. *The Tsallis entropy S_r for a continuous random variable X with the Burr Type-II distribution is (Anastasiadis, 2012)*

$$S_r(X) = \frac{1-\int f^r(x)\,dx}{r-1} = \frac{1-\int_0^\infty (\theta c)^r x^{r(c-1)}(1+x^c)^{-r(\theta+1)}\,dx}{r-1}. \tag{2.13}$$

Remark 5. Let us set (c = 1). In this case, equation (2.13) simplifies to the following

$$S_r(X) = \frac{1-\int_0^\infty \theta^r(1+x)^{-r(\theta+1)}\,dx}{r-1} = \frac{1-\frac{\theta^r}{r(\theta+1)-1}}{r-1}. \tag{2.14}$$

2.3 Stochastic orders

Proposition 2.6. *Let X and Y be two continuous lifetime random variables distributed by Burr Type-II with the parameters c, θ_1 and c, θ_2 respectively and $\theta_1 \leq (\geq)\theta_2$. Then $X \overset{st}{\leq (\geq)} Y$.*

Proof. Clearly $\theta_1 \leq (\geq)\theta_2$ yields $(1+x^c)^{-\theta_1} \geq (\leq)(1+x^c)^{-\theta_2}$ that implies $F_{c,\theta_1}(x) \leq (\geq)F_{c,\theta_2}(y)$. Hence from the definition of stochastic order (st) the proof is completed.

Proposition 2.7. *Let X and Y be two continuous lifetime random variables distributed by Burr Type-II with the parameters c_1, θ and c_2, θ respectively and $c_1 \leq (\geq)c_2$. Then $X \overset{st}{\leq (\geq)} Y$.*

Proof. Similar to Proposition (2.6), letting $c_1 \leq (\geq)c_2$ obtained $F_{c_1,\theta}(x) \leq (\geq)F_{c_2,\theta}(y)$.

Proposition 2.8. *Let X and Y are two random variables with the condition in Proposition* (2.6). *Then X is less (greater) than Y in likelihood ratio order $X \overset{lr}{\leq} (\overset{lr}{\geq})Y$ if $\theta_1 \leq (\geq)\theta_2$.*

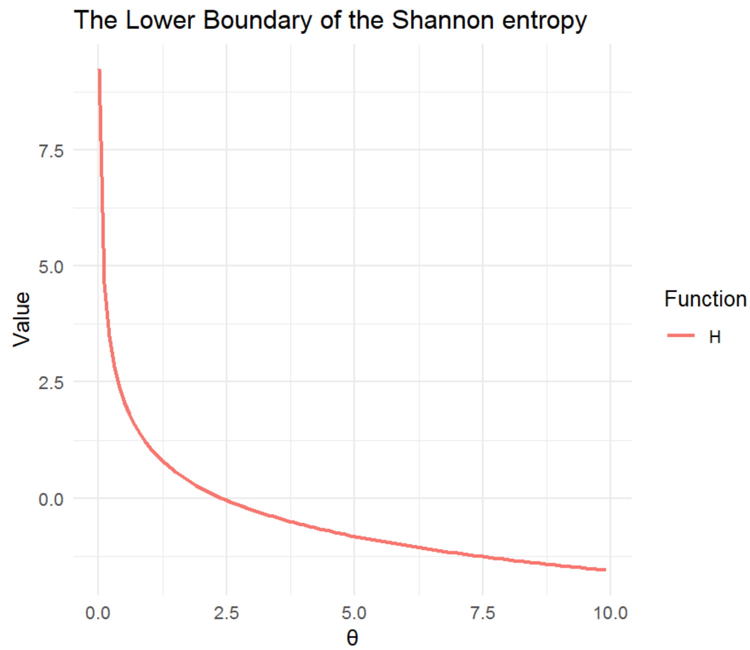

Figure 1: Examine the behavior of Shannon entropy in relation to the parameter θ in the Burr-Type II distribution, while keeping the parameter c fixed at 1, as outlined in proposition (2.1).

Proof. The likelihood ratio order (denoted by $X \overset{lr}{\leq} Y$) holds if the ratio $\frac{g_X(t)}{g_Y(t)}$ is decreasing in t (Di Crescenzo et al., 2001). Letting $f(t) = \frac{g_X(t)}{g_Y(t)}$ implies

$$f(t) = \frac{\theta_1}{\theta_2} \left[1 + t^c\right]^{\theta_2 - \theta_1}.$$

Taking the derivative of f(t) with respect to t, we get

$$f^{'}(t) = \frac{d}{dt} f(t) = \frac{\theta_1}{\theta_2} (\theta_2 - \theta_1) c t^{c-1} \left[1 + t^c\right]^{\theta_2 - \theta_1 - 1}.$$

Hence, $f^{'}(t) \leq (\geq) 0$ if $\theta_1 \geq (\leq) \theta_2$. Therefore $\frac{f_X(t)}{f_Y(t)}$ is decreasing (increasing) with the condition $\theta_1 \leq (\geq) \theta_2$ and the proof is completed.

Proposition 2.9. *Let X and Y are two random variables with the condition in Proposition* (2.7)*. Then X is greater (less) than Y in likelihood ratio order* $X \overset{lr}{\geq} Y (X \overset{lr}{\leq} Y)$ *if* $c_1 \geq (\leq) c_2$.

Proof. Let us $f(t) = \frac{c_1}{c_2} t^{c_1 - c_2} \left[\frac{1+t^{c_2}}{1+t^{c_1}}\right]^{\theta}$. Assume that $c_2 = k c_1$. Then, $f^{'}(t) \geq (\leq) 0$ if $k \geq (\leq) 1$ and the proof is completed.

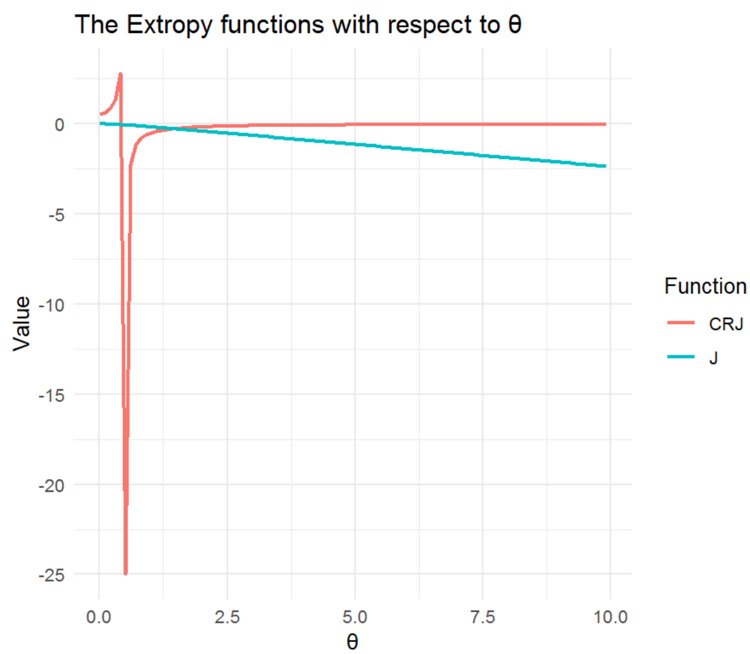

Figure 2: Examine the behavior of the Lad extropy and cumulative residual extropy measures in relation to the parameter θ in the Burr-Type II distribution, with the parameter c fixed at 1.

Proposition 2.10. *Let X and Y are two random variables with the condition in Proposition* (2.6)*. Then X is less (greater) than Y in hazard rate order* $X \overset{hr}{\leq} Y$ *if* $\theta_1 \geq (\leq)\theta_2$.

Proof. The hazard rate order (denoted by $X \overset{hr}{\leq} Y$) holds if the ratio $\frac{\bar{G}_X(t)}{\bar{G}_Y(t)}$ is decreasing in t (Di Crescenzo et al., 2001). Then, we solve $f^{'}(t) = \frac{d}{dt}\left(\frac{\bar{G}_X(t)}{\bar{G}_Y(t)}\right) = c(\theta_2 - \theta_1)x^{c-1}(1+x^c)^{\theta_2-\theta_1-1}$. Then taking $\theta_1 \leq (\geq)\theta_2$ yields $f(t)$ is increasing (decreasing).

Proposition 2.11. *Let X and Y are two random variables with the condition in Proposition* (2.7)*. Then X is less (greater) than Y in hazard rate order* $X \overset{hr}{\leq} Y$ *if* $\theta_1 \leq (\geq)\theta_2$.

Proof. Let $c_2 = kc_1$. We define the function

$$f(t) = \left[\frac{1+t^{c_2}}{1+t^{c_1}}\right]^{\theta}.$$

Then, it follows that$f^{'}(t) \geq (\leq)0$ if $k \geq (\leq)1$. Thus, the proof is complete.

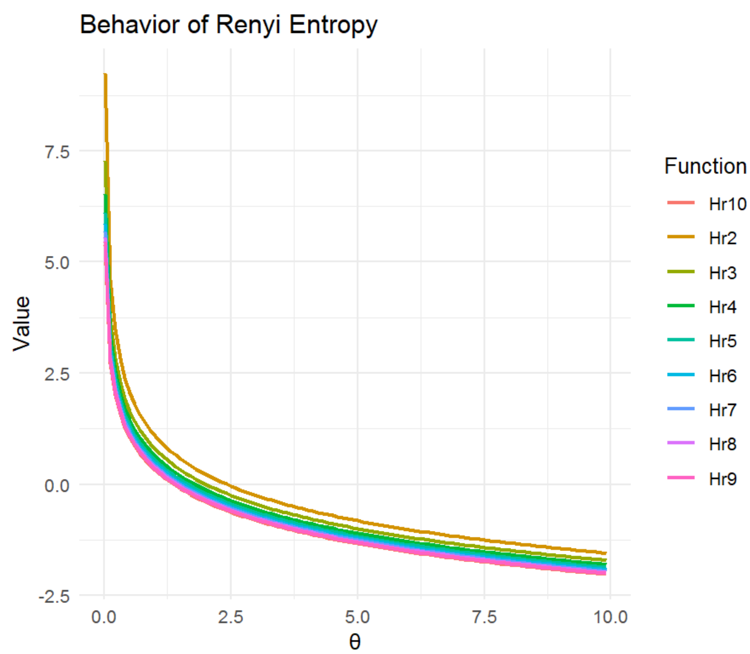

Figure 3: Investigate the impact of varying the constant "r" in Renyi entropy and the parameter "θ" in the Burr-Type II distribution, while keeping the parameter "c" fixed at 1.

2.4 Burr type-II record data

Record values play a crucial role in statistical analysis across a variety of fields, such as weather patterns, sports performance, economic trends, and life expectancy. For example, the Guinness World Records showcase impressive feats like rapid recitation of the periodic table, quick tennis matches, and indoor marathon victories. While numerous attempts are made to break records, only successful endeavors result in the establishment of new achievements. Despite the widespread use of record values in statistical analysis, complete data on record-breaking attempts is not always available, with only the final records being documented. As a result, ordered record data has been the subject of extensive research due to its importance in various disciplines. Understanding the characteristics of record values is therefore essential for accurate analysis and inference. Furthermore, recurrence relations play a significant role in the study of record values, adding another layer of depth to their analysis and interpretation.

Within the framework of information theory and uncertainty analysis in data distributions, the Kullback-Leibler divergence emerges as a pivotal methodological tool, particularly when confronting heavy-tailed distributions or conducting comparative analyses of extreme values. By concentrating on the informational structure inherent to distributional tail behavior, this approach rigorously quantifies the uncertainty associated with rare-event phenomena. It plays an instrumental role in

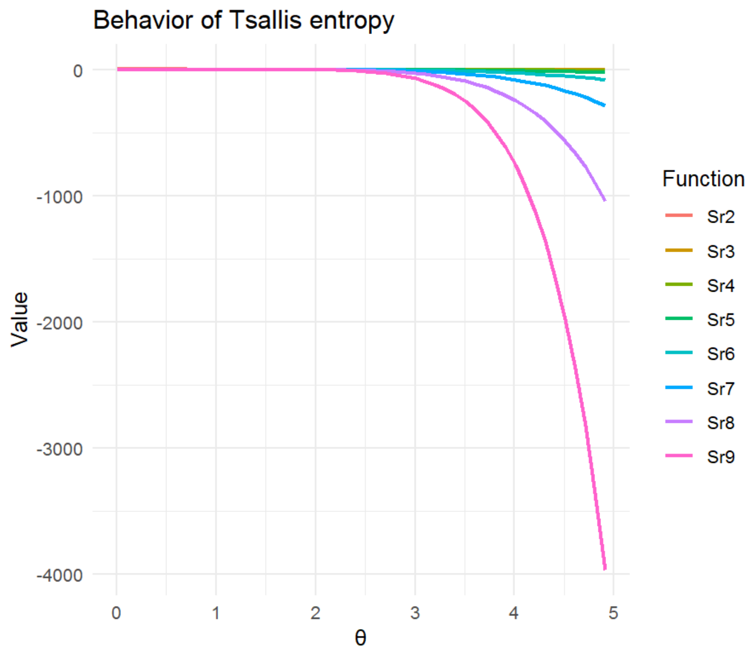

Figure 4: Investigate the impact of varying the constant "r" in Tsallis entropy and the parameter "θ" in the Burr-Type II distribution, while keeping the parameter "c" fixed at 1.

applications such as financial risk modeling, natural disaster forecasting, and communication network optimization, where the precision of tail-centric distributional analysis is critical to ensuring robustness and reliability. For more details see Pakgohar and Yousefzadeh (2025).

Based on the joint probability density function for the k^{th} upper record values as proposed by Pawlas and Szynal (1999) and Pawlas and Szynal (2001) the marginal probability density function for the k^{th} upper record value can be determined as

$$f_X^k(x) = \frac{c\theta^k}{(k-1)!} x^{c-1}(1+x^c)^{-(\theta+1)} \big(\ln(1+x^c)\big)^{k-1}. \tag{2.15}$$

3 Estimation

Suppose we have observed the first m lower record values, denoted as $X_{L(1)} = x_1, X_{L(2)} = x_2, \ldots, X_{L(m)} = x_m$ derived from the equation (2.1).

3.1 Maximum likelihood estimation

The likelihood function for these observations is expressed as follows see (Arnold et al., 2011)

$$L(\theta, c|\underset{-}{x}) = f(x_m;\theta,c)\Pi_{i=1}^{m-1}\frac{f(x_i;\theta,c)}{F(x_m;\theta,c)}; \tag{3.1}$$

where, $\underset{-}{x} = (x_1, x_2, \ldots, x_m)$. Also, the likelihood function $L(\theta, c|\underset{-}{x})$ quantifies the probability of observing the given record values under the specified model parameters, thereby facilitating the estimation of these parameters through maximum likelihood estimation.

Substituting equations (2.1) and (2.2) in equation (3.1), the likelihood function is

$$\begin{aligned} L(\theta, c|\underset{-}{x}) &= \theta^m c^m x_m^{-(c+1)}(1+x_m^{-c})^{-(\theta+1)}\Pi_{i=1}^{m-1}\frac{x_i^{-(c+1)}}{1+x_i^{-c}} \\ &= \theta^m c^m (1+x_m^{-c})^{-\theta}\Pi_{i=1}^{m}\frac{x_i^{-(c+1)}}{1+x_i^{-c}}. \end{aligned} \tag{3.2}$$

The log-likelihood function is

$$\begin{aligned} \mathcal{L} &= \ln L(\theta, c|\underset{-}{x}) \\ &= m\ln(\theta) + m\ln(c) - \theta\ln(1+x_m^{-c}) + \sum\nolimits_{i=1}^{m}\ln\left(\frac{x_i^{-(c+1)}}{1+x_i^{-c}}\right). \end{aligned} \tag{3.3}$$

From equation (3.3), we obtain the likelihood equations as

$$\frac{\partial \mathcal{L}}{\partial \theta} = \frac{m}{\theta} - \ln(1+x_m^{-c}) = 0; \tag{3.4}$$

where

$$\frac{m}{\theta} = \ln(1+x_m^{-c}). \tag{3.5}$$

From equation (3.5), the maximum likelihood estimate (MLE) of the parameter θ is given by

$$\hat{\theta} = \frac{m}{\ln(1+x_m^{-\hat{c}})}; \tag{3.6}$$

where $\hat{c}$ is the MLE of the parameter c which can be obtained as a solution of the following nonlinear equation.

$$\frac{m}{\hat{c}} + \frac{m}{\ln(1+x_m^{-\hat{c}})}\frac{x_m^{-\hat{c}}\ln(x_m)}{1+x_m^{-\hat{c}}} - \sum_{i=1}^{m}\ln(x_i) + \sum\nolimits_{i=1}^{m}\frac{x_i^{-\hat{c}}\ln(x_i)}{1+x_i^{-c}} = 0. \tag{3.7}$$

3.2 Bayes estimation

In the Bayesian inference procedures, we specify a loss function $L(\hat{\theta}, \theta)$ that describes the loss incurred by making an estimate $\hat{\theta}$ when the true value of the parameter is θ. In the literature, the most commonly used loss function is squared error loss (SEL) function. The symmetric nature of this function gives equal weight to overestimation as well as underestimation, while in the estimation of parameters of life time model; overestimation may be more serious than underestimation or vice-versa. For example, in the estimation of reliability and failure rate functions, an overestimation is usually much more serious than underestimation; in this case the use of symmetric loss function may be inappropriate as has been recognized by Basu and Ebrahimi (1991). This leads us to thinking that an asymmetrical loss function may be more appropriate.

One of the most popular asymmetric loss functions is the liner-exponential loss function. This loss function was introduced by Varrian (1975) cited by Zellner (1986) and was extensively discussed by Zellner (1986) and Azimi and Yaghmaei (2013). Under the assumption that the minimal loss occurs at $\theta^* = \theta$, the LINEX loss function for θ can be expressed as

$$L(\Delta) = e^{a\Delta} - a\Delta - 1, a \neq 0; \tag{3.8}$$

where $\Delta - \hat{\theta} = \theta$, and $\hat{\theta}$ is an estimate of θ.

The sign and magnitude of the shape parameter a represents the direction and degree of symmetry, respectively. For a close to zero, LINEX the loss is approximately SEL and therefore almost symmetric. The posterior expectation of the LINEX loss equation (3.8) is

$$E_\theta[L(\hat{\theta} - \theta)] = E_\theta(e^{a\Delta} - a\Delta - 1) = e^{a\hat{\theta}} E_\theta(e^{-a\theta}) + aE_\theta(\theta) - (a\hat{\theta} + 1); \tag{3.9}$$

where $E_\theta(.)$ denotes the posterior expectation with respect to the posterior density of θ. The Bayes estimator of θ, denoted by $\hat{\theta}_{BL}$ under LINEX loss function is the value $\hat{\theta}$ which minimizes the equation (3.9). It is

$$\hat{\theta}_{BL} = -\frac{1}{a} \ln(E_\theta(e^{-a\theta})); \tag{3.10}$$

provided that the expectation $E_\theta(e^{-a\theta})$ exists and is finite.

It is assumed that the parameter θ has a gamma prior distribution with shape and scale parameters as α and β , respectively and it has pdf

$$\pi(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} \theta^{\alpha-1} e^{-\beta\theta}. \tag{3.11}$$

Combining the likelihood function (equation (3.2)) and prior pdf (equation 3.11), we obtain the conditional posterior pdf of θ given c as

$$\begin{aligned}
\pi(\theta|c;x) &= \frac{L(x|c,\theta)\pi(\theta)}{\int\limits_0^\infty L(x|c,\theta)\pi(\theta)d\theta} = \frac{\theta^{m+\alpha-1} e^{-\beta\theta} (1 + x_m^{-c})^{-\theta}}{\int\limits_0^\infty \theta^{m+\alpha-1} e^{-\beta\theta} (1 + x_m^{-c})^{-\theta} d\theta} \\
&= \frac{\theta^{m+\alpha-1} e^{-\beta\theta} (1 + x_m^{-c})^{-\theta}}{\int\limits_0^\infty \theta^{m+\alpha-1} (\frac{e^{-\beta}}{1+x_m^{-c}})^{-\theta} d\theta} = \frac{\theta^{m+\alpha-1} e^{-\beta\theta} (1 + x_m^{-c})^{-\theta}}{\int\limits_0^\infty \theta^{m+\alpha-1} e^{-A\theta} d\theta};
\end{aligned} \tag{3.12}$$

where $A = \beta + \ln(1 + x_m^{-c})$. Then

$$\begin{aligned}\pi(\theta|c;x) &= \frac{A^{m+\alpha}\theta^{m+\alpha-1}e^{-\beta\theta}(1+x_m^{-c})^{-\theta}}{\Gamma(m+\alpha)} \\ &= \frac{\left(\beta+\ln(1+x_m^{-c})\right)^{m+\alpha}\theta^{m+\alpha-1}e^{-\beta\theta}(1+x_m^{-c})^{-\theta}}{\Gamma(m+\alpha)}. \end{aligned} \tag{3.13}$$

Under the squared error loss function, the usual estimate of a parameter is the posterior mean. Thus, Bayes estimates of the parameter θ obtained by using equation (3.13).

$$\begin{aligned}\hat{\theta}_{Bs} &= E(\theta|c;x) = \int_0^\infty \theta\pi(\theta|c;x)\, d\theta \\ &= \frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}}{\Gamma(m+\alpha)}\int_0^\infty \theta^{m+\alpha}e^{-\beta\theta}(1+x_m^{-c})^{-\theta}d\theta \\ &= \frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}}{\Gamma(m+\alpha)}\frac{\Gamma(m+\alpha+1)}{(\beta+\ln(1+x_m^{-c}))^{m+\alpha+1}} = \frac{m+\alpha}{\beta+\ln(1+x_m^{-c})}. \end{aligned} \tag{3.14}$$

Under the LINEX loss function, the Bays estimate $\hat{\theta}_{BL}$ is obtained by using equations (3.10) and 3.13, and is given by

$$\begin{aligned}\hat{\theta}_{BL} &= \frac{-1}{a}\ln(E_\theta(e^{-a\theta})) \\ &= \frac{-1}{a}\ln[\int_0^\infty e^{-a\theta}\frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}}{\Gamma(m+\alpha)}\theta^{m+\alpha-1}e^{-\beta\theta}(1+x_m^{-c})^{-\theta}d\theta] \\ &= \frac{-1}{a}\ln[\frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}}{\Gamma(m+\alpha)}\int_0^\infty \theta^{m+\alpha-1}e^{-(\beta+a)\theta}(1+x_m^{-c})^{-\theta}d\theta] \\ &= \frac{-1}{a}\ln[\frac{\beta+\ln(1+x_m^{-c})}{(\beta+a)+\ln(1+x_m^{-c})}]^{m+\alpha} = \frac{-(m+\alpha)}{a}\ln(\frac{\beta+\ln(1+x_m^{-c})}{(\beta+a)+\ln(1+x_m^{-c})}). \end{aligned} \tag{3.15}$$

4 Prediction

Suppose that we observe only the first m lower record observations, $X_{L(1)} = x_1, X_{L(2)} = x_2, \dots, X_{L(m)} = x_m$ and the goal is to predict the lower record value, $1 < m < s$. Let $Y \equiv X_{L(s)}$ be the lower record value, $1 < m < s$, the conditional distribution of Y given $x = (x_1, x_2, \dots, x_m)$ is just the distribution of Y given $X_{L(s)} \equiv X_m$ due to the well-known Markovian property of record statistics. It follows that (see Arnold et al. (2011))

$$f(y \mid x_m;\theta,c) = \frac{[H(y)-H(x_m)]^{s-m-1}}{\Gamma(s-m)}\frac{f(y|;\theta,c)}{F(x_m;\theta,c)}, 0 < y < x_m < \infty; \tag{4.1}$$

where, $H(.) = -\ln F(.)$. For the Burr Type-II distribution, with pdf given by 2.2, the function $f(y|x_m)$ is obtained as

$$f(y|x_m;\theta,c) = \frac{[\theta\ln(1+y^{-c}) - \theta\ln(1+x_m^{-c})]^{s-m-1}}{\Gamma(s-m)} \frac{\theta c y^{-(c+1)}(1+y^{-c})^{-(\theta+1)}}{(1+x_m^{-c})^{-\theta}}. \tag{4.2}$$

The Bayes predictive density function of $Y = X_{L(s)}$ given the observed record x_m is given by Nasiri and Pazira (2010)

$$f^*(y|x_m) = \int\limits_{\theta} f(y|x_m;\theta,c)\Pi(\theta|c,x_m)d\theta. \tag{4.3}$$

Substituting from equations (3.13) and (4.2) in equation (4.3), we get

$$\begin{aligned} f^*(y|x_m) &= \int\limits_0^\infty \frac{[\theta\ln(1+y^{-c}) - \theta\ln(1+x_m^{-c})]^{s-m-1}}{\Gamma(s-m)} \frac{\theta c y^{-(c+1)}(1+y^{-c})^{-(\theta+1)}}{(1+x_m^{-c})^{-\theta}} \\ &\times \frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}\theta^{m+\alpha+1}e^{-\beta\theta}}{\Gamma(m+\alpha)}(1+x_m^{-c})^{-\theta}d\theta \\ &= \frac{c y^{-(c+1)}[\ln(1+y^{-c}) - \ln(1+x_m^{-c})]^{s-m-1}}{\Gamma(s-m)} \\ &\times \frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}}{\Gamma(m+\alpha)} \int\limits_0^\infty \theta^{s+\alpha-1}e^{-\beta\theta}(1+y^{-c})^{-(\theta+1)}d\theta \\ &= B\frac{(\beta+\ln(1+x_m^{-c}))^{m+\alpha}}{\Gamma(m+\alpha)} \int\limits_0^\infty \theta^{m+\alpha+1}e^{-\beta\theta}(1+x_m^{-c})^{-\theta}d\theta; \end{aligned} \tag{4.4}$$

where, $B = \frac{c y^{-(c+1)}(1+y^{-c})^{-1}\left[\ln(1+y^{-c}) - \ln(1+x_m^{-c})\right]^{s-m-1}}{\Gamma(s-m)}$. So,

$$\begin{aligned} f^*(y|x_m) &= B\frac{\left(\beta+\ln(1+x_m^{-c})\right)^{m+\alpha}}{\Gamma(m+\alpha)}(1+y^{-c}) \int\limits_0^\infty \theta^{s+\alpha-1}e^{-\beta\theta}(1+y^{-c})^{-\theta}d\theta \\ &= (m+\alpha)c y^{-(c+1)}(1+y^{-c})^{-1} \frac{\left(\beta+\ln(1+x_m^{-c})\right)^{m+\alpha}}{\left(\beta+\ln(1+y^{-c})\right)^{m+\alpha+1}}; \quad 0 < y < x_m. \end{aligned} \tag{4.5}$$

For the special case, when $s = m+1$, which is practically of special interest, the $f^*(y|x_m)$ is given by

$$\begin{aligned} f^*(y|x_m) &= \frac{\Gamma(s+\alpha)c^{m+1}y^{-(c+1)}(1+y^{-c})\beta^\alpha}{\Gamma(\alpha)} \Pi_{i=1}^m \frac{x_i^{-(c+1)}}{1+x_i^{-c}} \left(\frac{1}{\beta+\ln(1+y^{-c})}\right)^{m+\alpha+1} \\ &= \frac{\Gamma(m+\alpha)c^{m+1}\beta^\alpha}{\Gamma(\alpha)} \Pi_{i=1}^m \frac{x_i^{-(c+1)}}{1+x_i^{-c}} \frac{y^{-(c+1)}(1+y^{-c})^{-1}}{(\beta+\ln(1+y^{-c}))^{m+\alpha+1}}; \quad 0 < y < x_m. \end{aligned} \tag{4.6}$$

Now, the lower and upper $100(1-\gamma)\%$ predication bounds for $Y \equiv X_{l(s)}$ can be obtained by evaluating the predictive survival function $P_r(Y \geq d|x_m)$, for some d. It follows, from 4.6, that

$$P_r(Y \geq d|x_m) = \int_d^{x_m} f^*(y|x_m)dy = A_1 \int_d^{x_m} \frac{y^{-(c+1)}(1+y^{-c})^{-1}}{(\beta+\ln(1+y^{-c}))^{m+\alpha+1}}dy; \tag{4.7}$$

where, $A_1 = (m+\alpha)c(\beta+\ln(1+x_m^{-c}))^{m+\alpha}$. Iterative numerical methods are required to obtain the lower and upper $100\tau\%$ prediction bounds for $Y = X_{L(s)}$ by finding d from using

$$P_r(L(x_m) \leq Y \leq U(x_m)) = \tau; \tag{4.8}$$

where $L(x_m)$ and $U(x_m)$ are the lower and upper limits, respectively, satisfying

$$P_r(Y \geq L(x_m)|x_m) = \frac{1+\tau}{2} \quad and \quad P_r(Y \geq U(x_m)|x_m) = \frac{1-\tau}{2}. \tag{4.9}$$

Now consider the case when the c is known. Without loss of generality, we can take $c = 1$. In this case, the predictive survival function $P_r(X_{L(m+1)} \geq d|x_m)$ can be written as

$$P_r(X_{L(m+1)} \geq d|x_m) = A_1 \int_d^{x_m} \frac{1}{(y+y^2)(\beta+\ln(1+\frac{1}{y}))^{m+\alpha+1}}dy; \tag{4.10}$$

where, $A_1 = (m+\alpha)(\beta+\ln(1+x_m^{-1}))^{m+\alpha}$. Assume that $\beta+\ln(1+\frac{1}{y}) = u, dy = -(y+y^2)du$, also, for $y = x_m \to u_2 = \beta+\ln(1+\frac{1}{x_m})$, besides, $y = d \to u_1 = \beta+\ln(1+\frac{1}{d})$. Then

$$\begin{aligned} P_r(X_{L(m+1)} \geq d|x_m) &= -A_1 \int_{u_1}^{u_2} \frac{-(y+y^2)}{(y+y^2)u^{m+\alpha+1}}du \\ &= \frac{A_1}{m+\alpha}u^{-(m+\alpha)}|_{u_1}^{u_2} = \frac{A_1}{m+\alpha}[u_2^{-(m+\alpha)} - u_1^{-(m+\alpha)}] \\ &= 1 - (\frac{\beta+\ln(1+x_m^{-1})}{\beta+\ln(1+\frac{1}{d})})^{m+\alpha}. \end{aligned} \tag{4.11}$$

From 4.9 and 4.11, the lower and upper limits are given, respectively, by

$$1 - (\frac{\beta+\ln(1+x_m^{-1})}{\beta+\ln(1+\frac{1}{d})})^{m+\alpha} = \frac{1+\tau}{2}.$$

It is easy to show that,

$$d = \frac{1}{\exp\left[(\frac{1-\tau}{2})^{\frac{1}{m+\alpha}}\left(\beta+\ln(1+x_m^{-1})\right) - \beta - 1\right]}.$$

Hence,

$$L_{(x_m)} = \frac{1}{\exp\{(\frac{1-\tau}{2})^{\frac{-1}{m+\alpha}}(\beta+\ln(1+x_m^{-1})) - \beta\} - 1}.$$

Similarly,

$$U_{(x_m)} = \frac{1}{\exp\{(\frac{1+\tau}{2})^{\frac{-1}{m+\alpha}}(\beta+\ln(1+x_m^{-1})) - \beta\} - 1}.$$

5 Numerical Computation

In this section, the behavior of the Burr Type-II distribution parameter is investigated according to estimator methods. The estimators are compared with the help of a Monte Carlo simulation study and using the mean square error criterion. Our main aim is to compare the MLE and Bayes estimator for parameter θ of Burr Type-II distribution under squared error and LINEX loss function. The results work for sample size $n = 10(10)60$, relative to scale parameters $\theta = 2(3)5$ and $c = 1$ respectively. We chose prior distribution parameters as $\alpha = 1, 1.25,$ and 2.5; also $\beta = 1.5$. The bias and mean squared errors (MSE) are computed for different sample size. By using the prediction procedure, the 95% prediction intervals for the next lower record are computed. In Section 5, the estimators are compared with the help of a Monte Carlo simulation study and using the mean square error criterion.

According to the results of Tables 1, 2 and 3, it can be said that:

A - By increasing the sample size, the mean square error of the Bayesian estimators decreases under the loss of square error and LINEX.

B - With the increase of the value of α , the mean square error of the Bayesian estimator under the LINEX loss function is always lower than the mean squared error under loss function.

C - As the value of θ increases, the mean square error of the estimators decreases.

D - The simulation results show that the prediction interval for $Y = X_{l(s)}$ based on Bayesian estimator under the loss function of Linux is better.

Table 1: Simulated values of bias and MSE of $\hat{\theta}_{Bs}$, $\hat{\theta}_{BI}$ and $L_1(x_m), U_1(x_m)$ for $\theta = 2$.

	$\hat{\theta}_{Bs}$		$\hat{\theta}_{BI}, \alpha = 1.00$		$\hat{\theta}_{BI}, \alpha = 1.25$		$\hat{\theta}_{BI}, \alpha = 2.50$		
n	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	$L_1(x_m), U_1(x_m)$
10	1.6554	0.1034	2.9984	0.1010	2.00783	0.0633	0.0065	0.0154	(0.0661,2.4813)
20	1.1241	0.0832	2.4588	0.0939	1.5693	0.0541	-0.3224	0.0147	(0.0026,1.0447)
30	0.8764	0.0601	2.1935	0.0733	1.3387	0.0476	-0.4338	0.0119	(0.0003,0.7069)
40	0.7716	0.0541	2.0218	0.0688	1.2105	0.0390	-0.5012	0.0103	(0.0001,0.5467)
50	0.6002	0.0382	1.8852	0.0504	1.1029	0.0380	-0.5441	0.0088	(0.0000,0.4729)
60	0.5161	0.0400	1.7878	0.0545	1.0308	0.0336	-0.5842	0.0084	(0.0000,0.4143)

6 Conclusion

The Burr Type-II distribution based on records provides a valuable framework for statistical modeling and prediction across various fields such as weather, finance, and engineering systems. This

Table 2: Simulated values of bias and MSE of $\hat{\theta}_{Bs}$, $\hat{\theta}_{BI}$ and $L_1(x_m), U_1(x_m)$ for $\theta = 3$.

	$\hat{\theta}_{Bs}$		$\hat{\theta}_{BI}, \alpha = 1.00$		$\hat{\theta}_{BI}, \alpha = 1.25$		$\hat{\theta}_{BI}, \alpha = 2.50$		
n	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	$L_1(x_m), U_1(x_m)$
10	0.9268	0.0629	2.2851	0.0596	1.2097	0.0364	-0.9817	0.0088	(0.1952,3.8569)
20	0.5005	0.0585	1.8587	0.0541	0.8787	0.0379	-1.0577	0.0086	(0.0213,1.7359)
30	0.2921	0.0496	1.6337	0.0529	0.7096	0.0341	-1.2442	0.008	(0.0059,1.2571)
40	0.1607	0.0401	1.4924	0.0455	0.6008	0.0287	-1.2511	0.008	(0.0023,1.0174)
50	0.0512	0.0389	1.3867	0.0404	0.5079	0.0286	-1.3485	0.007	(0.0011,0.8790)
60	-0.035	0.0348	1.2983	0.0395	0.4334	0.0263	-1.2811	0.0064	(0.0006,0.7879)

Table 3: Simulated values of bias and MSE of $\hat{\theta}_{Bs}$, $\hat{\theta}_{BI}$ and $L_1(x_m), U_1(x_m)$ for $\theta = 5$.

	$\hat{\theta}_{Bs}$		$\hat{\theta}_{BI}, \alpha = 1.00$		$\hat{\theta}_{BI}, \alpha = 1.25$		$\hat{\theta}_{BI}, \alpha = 2.50$		
n	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE	$L_1(x_m), U_1(x_m)$
10	0.8035	0.0317	0.514	0.0266	-0.5848	0.0153	2.7914	0.0039	(0.5777,6.9487)
20	1.1100	0.0288	0.2296	0.0261	-0.8157	0.0183	2.9098	0.0045	(0.1250,3.1696)
30	1.2751	0.0281	0.0713	0.026	-0.9471	0.0172	2.9700	0.0044	(0.0587,2.3805)
40	1.3795	0.0261	-0.0303	0.0254	-1.0341	0.0163	3.0147	0.0041	(0.0325,1.9642)
50	1.4650	0.0245	-0.1148	0.0244	-1.0944	0.0163	3.0483	0.0041	(0.0208,1.7245)
60	1.5236	0.0218	-0.133	0.0221	-1.1387	0.1436	3.0703	0.0038	(0.0152,1.5843)

study explores its properties, estimation techniques, and predictive capabilities, emphasizing the importance of record statistics in understanding extremes and sequential data. The investigation covers estimation methods including maximum likelihood and Bayesian estimation, demonstrating their application in deriving parameter estimates and making predictions based on record values. Additionally, the study delves into entropy, extropy, and other uncertainty measures to shed light on the distribution's behavior and insights gained from manipulating its parameters. Furthermore, the exploration of properties, uncertainty measures, and stochastic orders associated with the Burr Type-II distribution offers a comprehensive overview of its characteristics and applications. The study combines theoretical analysis and simulations to enhance understanding of this distribution and its relevance in statistical modeling and prediction scenarios.

References

Ahsanullah, M. (1995), *Record Statistics*, Nova Science Publishers.

Alghamdi, A. S., Alghamdi, H., and Fayomi, A. (2023), "The New Dagum-X Family of Distributions: Properties and Applications," *Int. J. Anal. Appl.*, 21, 45.

Anastasiadis, A. (2012), "Tsallis entropy," *Entropy*, 14, 174–176.

Arabi Belaghi, R., Arashi, M., and Tabatabaey, S. M. M. (2014), "Improved Confidence Intervals for the Scale Parameter of Burr XII Model Based on Record Values," *Computational Statistics*, 29, 1153–1173.

Arnold, B. C., Balakrishnan, N., and Nagaraja, H. N. (2011), *Records*, John Wiley & Sons.

Azimi, R. and Yaghmaei, F. (2013), "Bayesian Estimation for the Burr Type II Distribution under Type II Doubly Censored Data," *International Journal of Advanced Statistics and Probability*, 1, 1–3.

Basu, A. P. and Ebrahimi, N. (1991), "Bayesian Approach to Life Testing and Reliability Estimation Using Asymmetric Loss Function," *J. Statist. Plann. Infer.*, 29, 21–31.

Burr, I. W. (1942), "Cumulative Frequency Distribution," *Annals of Mathematical Statistics*, 13, 214–232.

Chandler, K. N. (1952), "The Distribution and Frequency of Record Values," *Journal of the Royal Statistical Society: Series B (Methodological)*, 14, 220–228.

Di Crescenzo, A., Longobardi, M., et al. (2001), "The Up Reversed Hazard Rate Stochastic Order," *Scientiae Mathematicae Japonicae*, 54, 575–581.

Hassan, A. S., Assar, S., Ali, K. A., and Nagy, H. F. (2021), "Estimation of the Density and Cumulative Distribution Functions of the Exponentiated Burr XII Distribution," *Statistics in Transition New Series*, 22, 171–189.

Ismail, N. H. (2014), "Parameter Estimations and Copula Methods for Burr Type II and Type XII Distributions," .

Jahanshahi, S., Zarei, H., and Khammar, A. (2020), "On cumulative residual extropy," *Probability in the Engineering and Informational Sciences*, 34, 605–625.

Kumar, V. and Rani, R. (2018), "Quantile approach of dynamic generalized entropy (divergence) measure," *Statistica*, 78, 105–126.

Lad, F., Sanfilippo, G., and Agro, G. (2015), "Extropy: Complementary dual of entropy," .

Laslan, F. K. I. and Nadar, M. (2017), "Statistical Inference of P(X ¡ Y) for the Burr Type XII Distribution Based on Records," *Hacettepe Journal of Mathematics and Statistics*, 46, 713–742.

Li, Z., Zhao, Y., Hu, X., Botta, N., Ionescu, C., and Chen, G. (2022), "Ecod: Unsupervised Outlier Detection Using Empirical Cumulative Distribution Functions," *IEEE Transactions on Knowledge and Data Engineering*.

Mahto, A. K., Tripath, Y., and Dey, S. (2018), "Estimation of Density and Distribution Functions of a Burr X Distribution," *Journal of Statistical Research*.

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

Nevzorov, V. B. (2000), *Records: Mathematical Theory*, American Mathematical Society.

Osagie, S. A., Uyi, S., and Osemwenkhae, J. E. (2023), "The Inverse Burr-Generalized Family of Distributions: Theory and Applications," *Earthline Journal of Mathematical Sciences*.

Pakgohar, A. (2024), "Exploring the Burr type III distribution: Properties, stochastic orders, and information measures," *Journal of Statistical Modelling: Theory and Applications*, 5, 165–180.

Pakgohar, A. and Yousefzadeh, F. (2025), "A Quantile Approach to the Kullback-Leibler Divergence on Type-I Censored Variable and Record Values," *Journal of Mathematical Extension*, 19, 1–17.

Pawlas, P. and Szynal, D. (1999), "Recurrence Relations for Single and Product Moments of K-th Record Values from Pareto, Generalized Pareto, and Burr Distributions," *Communications in Statistics-Theory and Methods*, 28, 1699–1709.

— (2001), "Recurrence Relations for Single and Product Moments of Generalized Order Statistics from Pareto, Generalized Pareto, and Burr Distributions," *Communications in Statistics-Theory and Methods*, 30, 739–746.

Renner, R. and Wolf, S. (2004), *Smooth Renyi entropy and applications*.

Sanusi, A. A., Doguwa, S. I. S., Audu, I., and Baraya, Y. M. (2020), "Burr X Exponential–G Family of Distributions: Properties and Application," *Asian Journal of Probability and Statistics*, 7, 58–75.

Zellner, A. (1986), "Bayesian Estimation and Prediction Using Asymmetric Loss Functions," *Journal of the American Statistical Association*, 81, 446–451.

Received: September 11, 2024

Accepted: May 16, 2025