STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **14**, September 2025, pp 1598–1610.
Published online in International Academic Press (www.IAPress.org)

On the Inference of Entropy Measures under Different Sampling Schemes

Hani Samawi [1,*], Amal Helu [2]

[1]*Department of Biostatistics, Epidemiology and Environmental Health Sciences Jiann-Ping Hsu College of Public Health, Georgia Southern University, USA*
[2]*Department of Mathematics, The University of Jordan, Amman-Jordan*

Abstract Entropy measures are fundamental measures for quantifying the uncertainty of random variables. In this study, we examine the maximum likelihood estimators (MLE) of five well-known entropy measures: Shannon, Rényi, Havrda-Charvát, Arimoto, and Tsallis, under both Simple Random Sampling (SRS) and Ranked Set Sampling (RSS). We derived the asymptotic bias and variance for these entropy estimators and conducted extensive simulations to assess the performance of SRS and RSS in estimating these entropy measures. The effectiveness of our estimators was demonstrated using breast cancer data.

Keywords Key Words: Burr XII; Shannon, Havrda-Charvát; Tsallis; Rényi Entropy; Arimoto; Ranked Set Sampling.

AMS 2010 subject classifications 62E10, 62N01, 62N02, 62G30

DOI: 10.19139/soic-2310-5070-2235

1. Introduction

Entropy, introduced by [31], is a key concept to measure uncertainty in random variables. It quantifies the average information content of a variable, with higher entropy indicating greater uncertainty and a broader probability distribution. In contrast, lower entropy reflects a more concentrated distribution with reduced uncertainty. Entropy has found broad applications across various scientific fields. For instance, in reliability studies, [28] and [34] emphasized entropy's relevance in evaluating the uncertainty of failure distributions, noting that higher entropy often correlates with less reliable outcomes. [8] examined its role in the insurance industry, particularly in risk assessment and the evaluation of extreme events, where higher entropy is linked to increased variability and potential losses. Additionally, entropy-based methods have been utilized in diverse areas such as neurobiology, statistics, cryptography, quantum computing, linguistics, and bioinformatics, as reported by [12], and [23]. These applications highlight the critical role of entropy in both theoretical and applied research.

1.1. The Burr XII Distribution and Its Applications

The Burr XII distribution is a highly versatile model, particularly effective in addressing non-monotonic failure rates, such as unimodal or bathtub-shaped rates, which are commonly encountered in reliability and biological research. While the Weibull distribution is often preferred for analyzing monotonic failure rates due to its ability to model both negatively and positively skewed distributions, the Burr XII distribution offers greater flexibility for modeling non-monotonic failure rates.

*Correspondence to: Hani Samawi (Email: hsamawi@georgiasouthern.edu) Department of Biostatistics, Epidemiology and Environmental Health Sciences Jiann-Ping Hsu College of Public Health, Georgia Southern University, USA.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

In reliability engineering, the Burr XII model is invaluable for predicting the lifespan of systems or components, especially in scenarios with incomplete data due to early test terminations. This distribution is also widely used in survival analysis to model various medical outcomes. In finance, the Burr XII distribution is instrumental in analyzing extreme events, such as financial crises, making it a crucial tool for risk management. Similarly, environmental scientists rely on the Burr XII distribution to study complex natural phenomena, such as rainfall patterns, which are vital for effective natural resource management. For further insights into the Burr XII distribution and its applications, refer to [2], and [35].

The cumulative distribution function (CDF) and reliability function of the Burr XII distribution are available in closed forms, making percentile and likelihood computations more straightforward, especially in varying sampling and censoring schemes. Let X be a random variable that follows a Burr XII (α, β) distribution. The probability density function (PDF) and cumulative distribution function (CDF) of X are expressed as follows:

$$f(x) = \alpha\beta x^{\beta-1}\left(1+x^{\beta}\right)^{-(\alpha+1)}, \quad x>0, \alpha, \beta>0, \tag{1}$$

$$F(x) = 1-\left(1+x^{\beta}\right)^{-\alpha}, \quad x>0, \alpha, \beta>0, \tag{2}$$

where, α and β are the shape parameters. The survival function, denoted as $S(x)$, and the hazard function, denoted as $h(x)$, are as follows:

$$S(x) = \left(1+x^{\beta}\right)^{-\alpha}, \quad x>0, \tag{3}$$

$$h(x) = \alpha\beta x^{\beta-1}\left(1+x^{\beta}\right)^{-1}, \quad x>0. \tag{4}$$

1.2. Entropy Measures of the Burr XII Distribution

Shannon Entropy: Let X be a random variable with the pdf given in Equation (1). The Shannon entropy of the Burr XII distribution is defined as:

$$SE = -E[\log f(x)] = -\int_0^{\infty} f(x)\log f(x)\,dx,$$

$$SE = -\left(\frac{1}{\beta}-1\right)[\psi(\alpha)+\lambda]+\frac{1}{\alpha}-\log(\alpha\beta)+1, \tag{5}$$

where,where $\psi(.)$ is the digamma function and $\lambda \approx 0.5772$ is the Euler-Mascheroni constant.

Shannon entropy is one of the earliest and most commonly used entropy measures. This measure has proven effective in the study of communication systems. However, one significant disadvantage of the Shannon measure, particularly in the continuous case, is that it may be negative for certain probability distributions, complicating its interpretation as a measure of uncertainty. Various generalizations have been proposed to address the limitations of Shannon entropy.

Rényi Entropy: [27] introduced a generalized entropy by extending the concepts of uncertainty and randomness. The Rényi entropy, which generalizes Shannon's entropy, is parameterized by a single parameter, p. As p approaches unity, it converges to the familiar Shannon entropy. A notable property of Rényi entropy is that in algorithms requiring entropy maximization, Rényi's entropy can be substituted directly for Shannon's, as both entropies reach their maximum under the same conditions (Principe2010). The Rényi entropy is calculated using the following formula:

$$RE = \frac{1}{1-p}\log\left[\int_0^{\infty}(f(x))^p\,dx\right] = \log\left[\frac{(\alpha\beta)^p}{\beta}\right]+\log\left(\frac{\Gamma\left[\frac{p(\beta-1)+1}{\beta}\right]\Gamma\left[\frac{p(\alpha\beta+1)-1}{\beta}\right]}{\Gamma\left[p(\alpha+1)\right]}\right), \quad p\geq 0, p\neq 1, \tag{6}$$

where, p is a parameter that leads to a positive entropy. The Rényi entropy is known as the quadratic entropy when $p = 2$, and $\Gamma(\cdot)$ is the complete gamma function. Eq. 6 exists if and only if $p(\beta - 1) + 1 > 0$, which is always satisfied if $\beta \geq 1$, a condition considered in the subsequent simulation study.

Havrda and Charvát Entropy: [15] proposed an extension of Rényi's entropy, known as HC entropy, which is defined as:

$$HCE = \frac{1}{2^{1-p}-1}\left[\int_0^\infty (f(x))^p\,dx - 1\right] = \frac{(\alpha\beta)^p\Gamma\left[\frac{p(\beta-1)+1}{\beta}\right]\Gamma\left[\frac{p(\alpha\beta+1)-1}{\beta}\right]}{\beta\Gamma\left[\alpha p + p\right]} - \frac{1}{2^{1-p}-1} \quad p \geq 0, p \neq 1, \quad (7)$$

HC entropy is often used in the context of fuzzy set theory and information retrieval, offering robustness in cases with incomplete or uncertain information.

Arimoto Entropy: [4] suggested another generalization of Shannon entropy, defined as:

$$AE = \frac{p}{1-p}\left\{\left[\int_0^\infty (f(x))^p\,dx\right]^{1/p} - 1\right\} = \frac{p}{1-p}\left\{\left[\frac{(\alpha\beta)^p\Gamma\left[\frac{p(\beta-1)+1}{\beta}\right]\Gamma\left[\frac{p(\alpha\beta+1)-1}{\beta}\right]}{\beta\Gamma\left[\alpha p + p\right]}\right]^{1/p} - 1\right\}, \quad p \geq 0, p \neq 1, \quad (8)$$

Tsallis Entropy: [37] generalized Shannon entropy and defined it as:

$$TE = \frac{1}{p-1}\left[1 - \int_0^\infty (f(x))^p\,dx\right] = \frac{1 - (\alpha\beta)^p\Gamma\left[\frac{p(\beta-1)+1}{\beta}\right]\Gamma\left[\frac{p(\alpha\beta+1)-1}{\beta}\right]}{(p-1)\,\beta\Gamma\left[\alpha p + p\right]}, \quad p \geq 0, p \neq 1, \quad (9)$$

Numerous researchers have explored entropy estimation for various life distributions. [9] analyzed entropy for the Weibull distribution under progressive censoring, while [7] examined entropy estimates for the Rayleigh distribution using doubly generalized Type II hybrid censoring. [5] studied entropy estimators for the inverse Lomax distribution within a multiple censored framework. [10] estimated Shannon entropy for the Lomax distribution using a non-informative prior. [26] compared the performance of maximum likelihood and Bayesian models with progressively censored data, and [38] applied Bayesian methods to estimate Shannon entropy for the Burr XII distribution using Type-II progressive censoring. [32] assessed the precision of entropy estimators for the Log-Logistic distribution, while [39] investigated Shannon entropy in the inverse Weibull distribution under progressive first-failure censoring, comparing credible and asymptotic intervals. Finally, [19] used Monte Carlo simulations to estimate Shannon entropy for progressively censored Maxwell distributions.

1.3. Ranked Set Sampling (RSS)

In agricultural, environmental studies, and more recently in human populations and reliability analysis, the cost of quantifying sampling units often exceeds that of physically acquiring them. As noted by [30] and [13], this challenge highlights the financial burden of extensive data collection in many research settings. Considerable cost savings can be realized by measuring only a subset of units while ensuring that all units contribute to the overall information content. Ranked set sampling (RSS) is an effective method to address this issue, introduced by [21]. It offers superior performance in hypothesis testing compared to simple random sampling (SRS) and is more efficient for estimating certain population parameters, as demonstrated by studies like [17] and [24].

In RSS, random samples are drawn and ranked based on the variable of interest, or a related variable, before measurement. The process repeats until a sufficient number of units are measured.

To implement this method, we select a random sample of size (r^2) elements from the population of interest, which is then partitioned into (r) sets of (r) units each. The classification of units within each set is based on

visual inspection or other variables associated with the variable under investigation. The measurement process begins with the first set of (r) units, where the unit with the lowest rank is selected for measurement of the variable of interest value. The second-ranked unit from the second set of (r) units is then measured for the biomarker of interest value, and the process continues until we measure the maximum-ranked unit in the (r)th set of size (r). The procedure can be reiterated (m) times to yield a sample with a size of $(n = mr)$. It is advisable to select the value of (r) within the range of 2 to 5 to reduced the ranking errors. The specific choice of (m) is contingent upon the desired sample size (n) ([29]). In case we ranked based on Concomitant Variables associated with the variable of interest that would introduce some ranking errors in the variable of interest, which will reduced the efficiency of estimation.

Variations such as extreme ranked set sampling (ERSS) by [29], median ranked set sampling (MRSS) by [22], and double ranked set sampling (DRSS) have further improved the efficiency of RSS ([3]).

Recent developments in RSS continue to expand its relevance across fields, including agriculture, public health, and reliability analysis. Notably, a 2023 study demonstrated the superior efficiency of RSS and its variations, such as Moving Extreme Ranked Set Sampling (MERSS) and MRSS, particularly when measurement costs are high ([11]). Recent research also emphasizes using RSS with concomitant variables and its potential for designing optimal sampling strategies in complex scenarios ([11]).

Advancements in information theory have further enhanced the understanding of RSS's benefits. [16] compared Shannon entropy, Rényi entropy, and Kullback-Leibler (KL) divergence under perfect and imperfect ranking conditions, showing that RSS consistently outperforms SRS in information content. Further studies, such as those by [36] and [41], have explored entropy estimators under RSS, demonstrating their efficacy in goodness-of-fit testing and Bayesian inference. Also, [1] provided estimation of different types of entropies for the Kumaraswamy distribution, while [18] provided bayesian nference for the entropy of the Rayleigh Model based on ordered RSS. [20] provided an extensive comparative review of estimation of Shannon differential entropy.

Overall, the evolution of RSS highlights its growing importance as a cost-effective, efficient, and flexible tool for statistical sampling across diverse disciplines. Its continued refinement and application will likely play a significant role in advancing research methodologies.

This paper focuses on the point and interval estimation of five entropies-Shannon, Rényi, Havrda-Charvát, Arimoto, and Tsallis, for the Burr XII distribution under RSS, and it provides a numerical comparison with estimators derived under SRS. The structure of the paper is as follows: Section 2 details the maximum likelihood estimation of the proposed entropies. Section 3 covers the construction $of 100(\alpha - 1)\%$ asymptotic confidence intervals for the entropy measures. A comprehensive simulation study is presented in Section 4. Section 5 applies the proposed estimators to real-world data, and the final remarks and discussions are provided in Section 6.

2. Maximum Likelihood Estimation

In this section, we concentrate on the maximum likelihood estimators (MLEs) for the parameters α and β of the Burr XII distribution under RSS. These estimators are vital for subsequent analysis, as they serve as the foundation for calculating various entropy measures and constructing their corresponding confidence intervals. The derivation process leverages the properties of the Burr XII distribution, in conjunction with the likelihood function specifically adapted for ranked set sampling.

2.1. Likelihood function and MLE using RSS

Suppose that $n = mr$ independent units are placed in a life-testing experiment whose lifetimes follow the Burr XII distribution with parameters α and β, with the pdf and cdf as shown in (1) and (2). The corresponding number of units measured using RSS is denoted by $(X_{(1)k}, \ldots, X_{(i)k}, \ldots, X_{(r)k})$ where (i) refers to the perfect ranking of the ith order. The likelihood function-based RSS is given by:

$$l(\alpha,\beta;\mathbf{x}) = \log L(\alpha,\beta;\mathbf{x}) \propto \sum_{k=1}^{m}\sum_{i=1}^{r}\Bigg[(i-1)\log\left(1-(1+X_{(i)k}^{\beta})^{-\alpha}\right) - (r-i)\alpha\log\left(1+X_{(i)k}^{\beta}\right) + \log(\alpha)+\log(\beta) + (\beta-1)\log(X_{(i)k}) + (\alpha+1)\log\left(1+X_{(i)k}^{\beta}\right)\Bigg]. \tag{10}$$

The MLEs of the parameters α and β can be obtained by using the partial derivative of the likelihood function Eq. 10 with respect to α and β, respectively, and equating the normal equations to 0 as follows:

$$\frac{\partial l(\alpha,\beta;\mathbf{x})}{\partial\alpha} = \sum_{k=1}^{m}\sum_{i=1}^{r}\Bigg[\frac{1}{\hat{\alpha}} - \log(1+X_{(i)k}^{\hat{\beta}}) - (r-i)\log(1+X_{(i)k}^{\hat{\beta}}) + \frac{(i-1)(1+X_{(i)k}^{\hat{\beta}})^{-\hat{\alpha}}\log(1+X_{(i)k}^{\hat{\beta}})}{1-(1+X_{(i)k}^{\hat{\beta}})^{-\hat{\alpha}}}\Bigg] = 0. \tag{11}$$

$$\frac{\partial l(\alpha,\beta;\mathbf{x})}{\partial\beta} = \sum_{k=1}^{m}\sum_{i=1}^{r}\Bigg[\frac{1}{\hat{\beta}} + \log(X_{(i)k}) - \frac{(r-i)\hat{\alpha}X_{(i)k}^{\hat{\beta}}\log(X_{(i)k})}{(1+X_{(i)k}^{\hat{\beta}})} - \frac{(\hat{\alpha}+1)X_{(i)k}^{\hat{\beta}}\log(X_{(i)k})}{(1+X_{(i)k}^{\hat{\beta}})} + \frac{(i-1)\hat{\alpha}(1+X_{(i)k}^{\hat{\beta}})^{-\hat{\alpha}-1}X_{(i)k}^{\hat{\beta}}\log(X_{(i)k})}{1-(1+X_{(i)k}^{\hat{\beta}})^{-\hat{\alpha}}}\Bigg] = 0. \tag{12}$$

We use the Newton-Raphson algorithm to solve for the MLE of α and β.

Due to the important property of MLE, the MLE of the discussed entropies in Eqs 5 - 9 can be obtained by plugging in the MLE of α and β, respectively, into equations 5 - 9 directly under SRS and RSS.

3. Point Interval estimation of Entropies

The asymptotic confidence intervals (CIs) for entropy measures are derived to assess the uncertainty associated with estimates from the Burr XII distribution under Simple Random Sampling (SRS) and Ranked Set Sampling (RSS). These intervals are constructed based on the normal approximation of the maximum likelihood estimators (MLEs), given certain regularity conditions.

To develop these confidence intervals, the delta method is employed. This method, in conjunction with the observed Fisher information matrix, facilitates the estimation of variances and covariances for the entropy measures, which are crucial for CI calculation. For further details on the delta method and its applications, refer to [6]. The next step involves determining the observed Fisher information matrix and the variance-covariance matrix as follows:

$$\mathbf{I}(\alpha,\beta) = \begin{bmatrix} -E\left(\frac{\partial^2 l(\alpha,\beta;\mathbf{x})}{\partial\alpha^2}\right) & -E\left(\frac{\partial^2 l(\alpha,\beta;\mathbf{x})}{\partial\alpha\partial\beta}\right) \\ -E\left(\frac{\partial^2 l(\alpha,\beta;\mathbf{x})}{\partial\alpha d\beta}\right) & -E\left(\frac{\partial^2 l(\alpha,\beta;\mathbf{x})}{\partial\beta^2}\right) \end{bmatrix}, \tag{13}$$

and the variance-covariance matrix is given by:

$$Cov(\hat{\alpha},\hat{\beta}) = \mathbf{I}^{-1}(\hat{\alpha},\hat{\beta}). \tag{14}$$

To apply the delta method, we must first derive the necessary partial derivatives for the proposed entropy measures outlined in equations 5 to 9. These derivatives are essential for calculating the asymptotic variance of the MLEs (Maximum Likelihood Estimators) for these estimated entropies, as follows:

Shannon Entropy

$$\nabla_{SE}^{T} = \left(\frac{\partial SE}{\partial \alpha} - \frac{\partial SE}{\partial \beta} \right),$$

$$\frac{\partial SE}{\partial \alpha} = \frac{(\beta-1)\psi'(\alpha)}{\beta} - \frac{(1+\alpha)}{\alpha^2},$$

$$\frac{\partial SE}{\partial \beta} = \frac{\lambda - \beta + \psi(\alpha)}{\beta^2},$$

where $\psi'(\alpha)$ is the derivative of the digamma.

Renyi Entropy

$$\nabla_{RE}^{T} = \left(\frac{\partial RE}{\partial \alpha} - \frac{\partial RE}{\partial \beta} \right),$$

$$\frac{\partial RE}{\partial \alpha} = \frac{p}{(1-p)\alpha} \left[\alpha \left(\psi(\alpha p + p) - \psi \left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \right) - 1 \right],$$

$$\frac{\partial RE}{\partial \beta} = -\frac{1}{\beta^2} \left[\beta + \psi \left(\frac{p(\beta-1)+1}{\beta} \right) - \psi \left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \right].$$

Havrda and Charvat Entropy

$$\nabla_{HCE}^{T} = \left(\frac{\partial HCE}{\partial \alpha} - \frac{\partial HCE}{\partial \beta} \right),$$

$$\frac{\partial HCE}{\partial \alpha} = \frac{2^p p(\alpha\beta)^{p-1} \Gamma\left(\frac{p(\alpha\beta+1)-1}{\beta}\right) \Gamma\left(\frac{p(\beta-1)+1}{\beta}\right)}{(2^p-2)\Gamma(\alpha p+p)} \alpha \left[\psi(\alpha p+p) - \psi\left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \right] - 1,$$

$$\frac{\partial HCE}{\partial \beta} = \frac{2^p (p-1)(\alpha\beta)^{p} \Gamma\left(\frac{p(\alpha\beta+1)-1}{\beta}\right) \Gamma\left(\frac{p(\beta-1)+1}{\beta}\right)}{(2^p-2)\beta^3\Gamma(\alpha p+p)} \beta + \psi\left(\frac{p(\beta-1)+1}{\beta} \right) - \psi\left(\frac{p(\alpha\beta+1)-1}{\beta} \right).$$

Arimoto Entropy

$$\nabla_{AE}^{T} = \left(\frac{\partial AE}{\partial \alpha} - \frac{\partial AE}{\partial \beta} \right),$$

$$\frac{\partial AE}{\partial \alpha} = \frac{p}{p}\alpha \left[\psi(\alpha p+p) - \psi\left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \right] - 1 \left[\alpha\beta\Gamma\left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \Gamma\left(\frac{p(\beta-1)+1}{\beta} \right) \right],$$

$$\frac{\partial AE}{\partial \beta} = \frac{p}{p}\alpha \left[\psi(\alpha p+p) - \psi\left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \right] - 1 \left[\alpha\beta\Gamma\left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \Gamma\left(\frac{p(\beta-1)+1}{\beta} \right) \right].$$

Tsallis Entropy

$$\nabla_{TE}^{T} = \left(\frac{\partial TE}{\partial \alpha} - \frac{\partial TE}{\partial \beta} \right),$$

$$\frac{\partial TE}{\partial \alpha} = \frac{p\alpha(\beta-1)(\alpha\beta)^{p-2} \Gamma\left(\frac{p(\alpha\beta+1)-1}{\beta}\right) \Gamma\left(\frac{(p-1)(\beta-1)}{\beta}\right)}{\Gamma(\alpha p+p)} \alpha \left[\psi(\alpha p+p) - \psi\left(\frac{p(\alpha\beta+1)-1}{\beta} \right) \right] - 1,$$

$$\frac{\partial TE}{\partial \beta} = \frac{p\alpha(\beta+1)(\alpha\beta)^{p-2} \Gamma\left(\frac{p(\alpha\beta+1)-1}{\beta}\right) \Gamma\left(\frac{(p-1)(\beta-1)}{\beta}\right)}{\beta^3\Gamma(\alpha p+p)} \beta + \psi\left(\frac{p(\beta-1)+1}{\beta} \right) - \psi\left(\frac{p(\alpha\beta+1)-1}{\beta} \right).$$

Therefore, the variance of any of the above estimated entropies (say $\hat{\theta}$) can be derived by delta methods as follows:

$$Var(\hat{\theta}) = \left(\frac{\partial\theta}{\partial\alpha} - \frac{\partial\theta}{\partial\beta}\right) T^{-1}(\hat{\alpha},\hat{\beta}) \begin{pmatrix} \frac{\partial\theta}{\partial\alpha} \\ \frac{\partial\theta}{\partial\beta} \end{pmatrix}.$$

Due to the complexity of the formulations, the full derivation of the variance is not provided here. However, Mathematica was utilized to simplify certain expressions. Therefore, the $100(1-\alpha)\%$ asymptotic confidence intervals for the entropy θ is given by $\hat{\theta} \pm Z_{\alpha/2}\sqrt{Var(\hat{\theta})}$, where $Z_{\alpha/2}$ denotes the upper $(\alpha/2)$-th percentile of the standard normal distribution.

4. Simulation Study

To gain a deeper understanding and compare the performance of the five entropy measures, Shannon, Rényi, Havrda-Charvát, Arimoto, and Tsallis, under both Simple Random Sampling (SRS) and Ranked Set Sampling (RSS), we conducted extensive simulations using various sample sizes. Specifically, we simulate different SRS and RSS scenarios with set sizes (r) of 3, 4, and 5, and a cycle size (m) of 30, resulting in total sample sizes (n) ranging from 90 to 150. We used multiple-parameter settings for the Burr XII distribution. $\alpha = 2, \beta = 2, p = 0.5, 1.5, 3$ to assess the robustness and variability of the entropy estimates. Our analysis uses 5,000 simulated samples for each scenario to ensure reliable and statistically significant comparisons. These simulations comprehensively evaluate the entropies' performance under varying sampling schemes and sample sizes.

Table 1. MLE, Bias, Variance, and Efficiency Estimators of RSS with Respect to SRS for $\alpha = 2$, $\beta = 2$, $p = 0.5$.

		SRS			RSS			
Entropy	Exact	MLE	Bias	Variance	MLE	Bias	Variance	Efficiency
				$\mathbf{r=3, m=30, n=90}$				
Shannon	0.6137	0.6059	-0.0078	0.0098	0.6065	-0.0073	0.0076	1.2896
Rényi	1.0547	1.0476	-0.0071	0.0238	1.0479	-0.0068	0.0196	1.2113
HCE	1.6765	1.6741	-0.0024	0.1006	1.6726	-0.0039	0.0832	1.2096
Arimoto	1.8711	1.8850	0.01391	0.2046	1.8799	0.00881	0.1688	1.2125
Tsallis	1.3889	1.4243	0.0354	0.0718	1.4101	0.0213	0.0383	1.8752
				$\mathbf{r=4, m=30, n=120}$				
Shannon	0.6137	0.6035	-0.0102	0.0075	0.6100	-0.0037	0.0050	1.4829
Rényi	1.0547	1.0432	-0.0115	0.0181	1.0521	-0.0026	0.0131	1.3763
HCE	1.6765	1.6622	-0.0143	0.0758	1.6779	0.0014	0.0553	1.3716
Arimoto	1.8711	1.8641	-0.0070	0.1518	1.8825	0.0114	0.1107	1.3720
Tsallis	1.3889	1.4041	0.0152	0.0535	1.4024	0.0135	0.0233	2.2984
				$\mathbf{r=5, m=30, n=150}$				
Shannon	0.6137	0.6064	-0.0073	0.0061	0.6107	-0.0030	0.0038	1.6043
Rényi	1.0547	1.0468	-0.0079	0.0148	1.0524	-0.0023	0.0101	1.4652
HCE	1.6765	1.6680	-0.0085	0.0622	1.6769	0.0005	0.0427	1.4566
Arimoto	1.8711	1.8698	-0.0013	0.1242	1.8790	0.0080	0.0856	1.4529
Tsallis	1.3889	1.4024	0.0136	0.0430	1.3987	0.0098	0.0154	2.7878

Tables 1-4 present the results of our simulation for various values of $p = 0.5, 1.5, 3$, as well as different set sizes $r = 3, 4,$ and 5. RSS proves to be more efficient than SRS for the same sample size in every scenario. As anticipated, efficiency increases as the size of the set r grows. Consequently, using RSS, when feasible, can reduce the required sample size by at least 20% to achieve the same level of precision as SRS, leading to significant cost savings.

Furthermore, the precision and efficiency of all estimators are influenced by the choice of p. Specifically, when estimating Shannon and Tsallis entropies, efficiency tends to decrease as p increases. In contrast, the efficiency

of other entropy estimators improves as p increases from 0.5 to 3.0. As anticipated, both the bias and variance of all estimators are lower when using Ranked Set Sampling (RSS) compared to Simple Random Sampling (SRS). However, no consistent pattern emerges regarding the effect of p on bias. Additionally, RSS offers narrower confidence intervals and improved coverage compared to SRS.

Table 2. MLE, Bias, Variance, and Efficiency Estimators of RSS with Respect to SRS for $\alpha = 2, \beta = 2, p = 1.5$

		SRS			RSS			Efficiency
Entropy	Exact	MLE	Bias	Variance	MLE	Bias	Variance	
				$\mathbf{r = 3, m = 30, n = 90}$				
Shannon	0.6137	0.6022	-0.0115	0.0097	0.6047	-0.0090	0.0077	1.2597
Renyi	0.4598	0.4485	-0.0114	0.0075	0.4510	-0.0089	0.0057	1.3042
HCE	0.7013	0.6833	-0.0180	0.0140	0.6873	-0.0140	0.0107	1.3118
Arimoto	0.4263	0.4155	-0.0108	0.0056	0.4179	-0.0084	0.0043	1.3091
Tsallis	0.4108	0.3782	-0.0326	0.0304	0.3822	-0.0286	0.0283	1.0752
				$\mathbf{r = 4, m = 30, n = 120}$				
Shannon	0.6137	0.6059	-0.0078	0.0076	0.6083	-0.0054	0.0050	1.5172
Renyi	0.4598	0.4520	-0.0079	0.0058	0.4543	-0.0055	0.0037	1.5725
HCE	0.7013	0.6886	-0.0127	0.0108	0.6926	-0.0087	0.0069	1.5761
Arimoto	0.4263	0.4187	-0.0076	0.0043	0.4211	-0.0052	0.0027	1.5746
Tsallis	0.4108	0.3871	-0.0237	0.0234	0.3944	-0.0164	0.0185	1.2618
				$\mathbf{r = 5, m = 30, n = 150}$				
Shannon	0.6137	0.6057	-0.0080	0.0060	0.6103	-0.0034	0.0038	1.5881
Renyi	0.4598	0.4520	-0.0078	0.0046	0.4563	-0.0035	0.0028	1.6532
HCE	0.7013	0.6891	-0.0122	0.0086	0.6956	-0.0057	0.0052	1.6624
Arimoto	0.4263	0.4190	-0.0074	0.0034	0.4229	-0.0034	0.0021	1.6591
Tsallis	0.4108	0.3914	-0.0194	0.0178	0.3992	-0.0116	0.0134	1.3311

Table 3. MLE, Bias, Variance, and Efficiency Estimators of RSS with Respect to SRS for $\alpha = 2$, $\beta = 2$, $p = 3.0$.

		SRS			RSS			Efficiency
Entropy	Exact	MLE	Bias	Variance	MLE	Bias	Variance	
				$r = 2, m = 30, n = 60$				
Shannon	0.6137	0.6046	-0.0091	0.0100	0.6055	-0.0082	0.0077	1.3018
Renyi	0.2798	0.2706	-0.0092	0.0060	0.2717	-0.0081	0.0043	1.3927
HCE	0.5714	0.5479	-0.0236	0.0152	0.5521	-0.0193	0.0110	1.3862
Arimoto	0.2553	0.2459	-0.0093	0.0042	0.2473	-0.0080	0.0030	1.3876
Tsallis	0.2143	0.2001	-0.0142	0.0037	0.2022	-0.0121	0.0032	1.1580
				$\mathbf{r = 4, m = 30, n = 120}$				
Shannon	0.6137	0.6071	-0.0066	0.0073	0.6082	-0.0055	0.0049	1.4705
Renyi	0.2798	0.2729	-0.0069	0.0044	0.2745	-0.0053	0.0028	1.5760
HCE	0.5714	0.5541	-0.0174	0.0108	0.5590	-0.0124	0.0067	1.6060
Arimoto	0.2553	0.2483	-0.0070	0.0031	0.2501	-0.0052	0.0019	1.5824
Tsallis	0.2143	0.2043	-0.0100	0.0026	0.2063	-0.0080	0.0019	1.3511
				$\mathbf{r = 5, m = 30, n = 150}$				
Shannon	0.6137	0.6068	-0.0069	0.0061	0.6097	-0.0040	0.0036	1.6791
Renyi	0.2798	0.2729	-0.0069	0.0036	0.2759	-0.0039	0.0020	1.8250
HCE	0.5714	0.5552	-0.0090	0.0089	0.5624	-0.0090	0.0048	1.8610
Arimoto	0.2553	0.2485	-0.0068	0.0025	0.2515	-0.0038	0.0014	1.8330
Tsallis	0.2143	0.2053	-0.0090	0.0021	0.2085	-0.0058	0.0014	1.5055

5. Illustration using WBCD data

For the diagnosis of breast cancer, we utilized summary features from digitized images of a fine needle aspirate (FNA) of breast masses, which serve as biomarkers. This section applies the proposed entropy measures to assess

Table 4. MLE, 95% Confidence Interval Length and Coverage Probability of SRS and RSS for different values of p.

		Shannon		Renyi		HCE		Arimoto		Tsallis	
r, m, n	Scheme	Length	Cov	Length	Cov	Length	Cov	Length	Cov	Length	Cov
					p = 0.5						
$r = 2, m = 30, n = 60$	SRS	0.3878	0.9422	0.6065	0.9376	1.2492	0.9348	1.7767	0.9290	1.0349	0.9414
	RSS	0.3878	0.9682	0.6065	0.9614	1.2475	0.9556	1.7701	0.9496	1.0335	0.9874
$r = 3, m = 30, n = 90$	SRS	0.3354	0.9434	0.5237	0.9426	1.0735	0.9370	1.5173	0.9298	0.8893	0.9400
	RSS	0.3363	0.9790	0.5262	0.9716	1.0815	0.9682	1.5304	0.9648	0.8959	0.9956
					p = 1.5						
$r = 2, m = 30, n = 60$	SRS	0.3875	0.9418	0.3401	0.9428	0.4637	0.9448	0.2927	0.9440	0.2716	0.5726
	RSS	0.3876	0.9668	0.3400	0.9716	0.4630	0.9744	0.2924	0.9736	0.2712	0.5886
$r = 3, m = 30, n = 90$	SRS	0.3359	0.9436	0.2946	0.9416	0.4010	0.9432	0.2533	0.9428	0.2349	0.5632
	RSS	0.3361	0.9806	0.2947	0.9820	0.4006	0.9844	0.2532	0.9832	0.2347	0.6216
					p = 3.0						
$r = 2, m = 30, n = 60$	SRS	0.3877	0.9426	0.3019	0.9430	0.4726	0.9532	0.2521	0.9470	0.1772	0.8738
	RSS	0.3878	0.9656	0.3018	0.9714	0.4701	0.9808	0.2518	0.9732	0.1763	0.8976
$r = 3, m = 30, n = 90$	SRS	0.3361	0.9498	0.2614	0.9510	0.4063	0.9540	0.2179	0.9542	0.1524	0.8826
	RSS	0.3358	0.9810	0.2611	0.9850	0.4035	0.9884	0.2174	0.9870	0.1513	0.9212

the uncertainty between benign and malignant patients using data from the Wisconsin Breast Cancer Database (WBCD), created by the University of Wisconsin ([33]). The dataset consists of 569 observations and 30 features, with the variable "Diagnosis" serving as the gold standard, where $B =$ benign ($n = 357$) and $M =$ malignant ($n = 212$). Among the 30 features, we selected the perimeter-worst biomarker due to its superior diagnostic performance. This biomarker demonstrates high sensitivity (0.920) and specificity (0.919), with a Youden Index of 0.839, outperforming other biomarkers in differentiating between benign and malignant cases.

The legitimacy of the Burr XII model for both benign and malignant data is assessed based on $(\alpha_1, \beta_1) = (1.6305, 9.4945)$ for the benign group and $(\alpha_2, \beta_2) = (1.2270, 11.1030)$ for the malignant group, using Kolmogorov-Smirnov (K-S), Anderson-Darling (A-D), and chi-squared tests. The results, presented in Table 5 at a significance level of 0.05, provide strong evidence that the Burr XII model fits both datasets well.

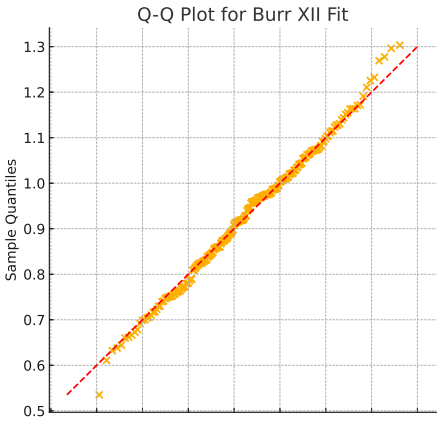
Table 5. Goodness-of-fit tests for the benign and malignant data.

Data	**K-S (p-value)**	**A-D (p-value)**	**Chi-squared (p-value)**
Benign	0.0379 (0.6711)	0.4960 (0.4995)	6.7356 (0.5654)
Malignant	0.0381 (0.9073)	0.2716 (0.4731)	6.6782 (0.4631)

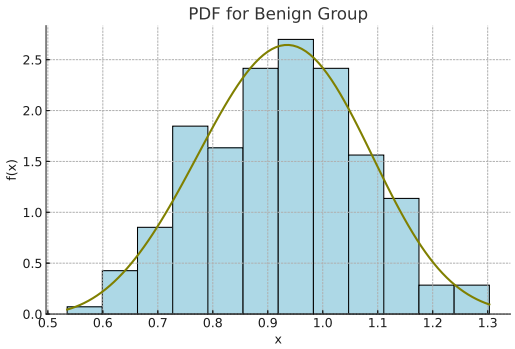
Additionally, the Q-Q and fitted pdf plots for the benign and malignant datasets, shown in Figures 1 and 2, respectively, further confirm the Burr XII distribution is a suitable model for both datasets.

Table 6 shows the results of the uncertainty entropy measures of the selected biomarker for the benign and malignant patients. We used the whole data ($N = 569$) as the population data to calculate the exact values of the Burr XII parameters and hence the exact values of those entropy measures. We draw a SRS of size $n = 90$ and RSS sample of size ($r = 3, m = 30$) from the benign data and SRS of size $n = 60$ and RSS of size ($r = 3, m = 20$) from the malignant data.

Figures 1 and the goodness-of-fit test indicated that the biomarker data for the benign group follow Burr XII distribution with $(\alpha = 1.6305, \beta = 9.4945)$. Similarly, Figure 2 and the goodness-of-fit test indicated that the biomarker data for the malignant group follow Burr XII distribution with $(\alpha = 1.2270, \beta = 11.1030)$.

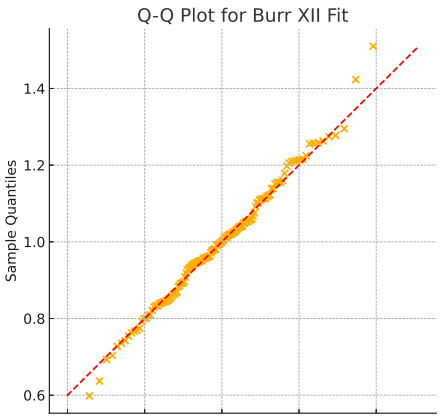


(a) Q-Q Plot for Benign Group

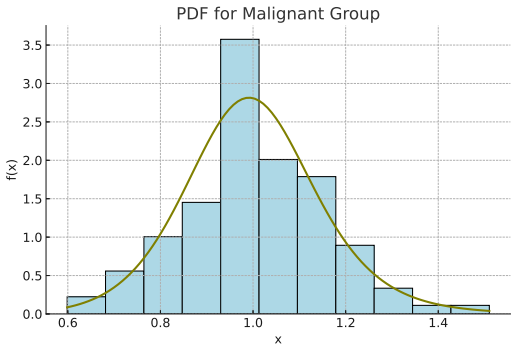


(b) PDF for Benign Group

Figure 1. Q-Q Plot and PDF for the Benign Group



(a) Q-Q Plot for Malignant Group



(b) PDF for Malignant Group

Figure 2. Q-Q Plot and PDF for the Malignant Group

Table 6. Maximum Likelihood Estimates (MLE), Asymptotic Bias, and Variance of Entropy Measures for Benign and Malignant WBCD Data (perimeter worst)

	Benign patients data (N=357, $\alpha = 1.6305$, $\beta = 9.4945$, $p = 0.20$)						
	SRS (n=90)			RSS (r=3, m=30)			
Entropy	**Exact**	**MLE**	**Bias**	**Variance**	**MLE**	**Bias**	**Variance**
SE	-0.4740	-0.5667	-0.0927	0.0072	-0.4337	0.0402	0.0071
RE	0.3126	0.2238	-0.0888	0.0124	0.3871	0.0746	0.0113
HCE	0.3833	0.2645	-0.1189	0.0206	0.4898	0.1065	0.0165
AE	0.6228	0.3618	-0.2610	0.0741	0.9261	0.3033	0.0351
TE	0.3551	0.4394	0.0843	0.0177	0.4381	0.0830	0.0142

	Malignant patients data (N=212, $\alpha = 1.2270$, $\beta = 11.1030$, $p = 0.20$)						
	SRS (n=60)			RSS (r=3, m=20)			
Entropy	**Exact**	**MLE**	**Bias**	**Variance**	**MLE**	**Bias**	**Variance**
SE	-0.5039	-0.4522	0.0518	0.0112	-0.4880	-0.0159	0.0110
RE	0.3772	0.3884	0.0112	0.0227	0.3735	0.0037	0.0196
HCE	0.4752	0.4917	0.0165	0.0493	0.4700	-0.0052	0.0329
AE	0.8801	0.9320	0.0519	0.5083	0.8639	0.0162	0.1226
TE	0.4402	0.2668	-0.1735	0.0423	0.3815	-0.0587	0.0283

Table 6 demonstrates that using Ranked Set Sampling (RSS) reduces both the bias and variance of the maximum likelihood estimators (MLE) of the entropies, thereby improving the efficiency of these estimators, as evidenced by our simulation results. However, a limitation of Shannon entropy is observed, as the MLE for this entropy produced negative values for both benign and malignant datasets. This issue, particularly in the continuous case, complicates the interpretation of Shannon entropy as a measure of uncertainty for certain probability distributions. To address this issue, we need to impose non-negativity constraints in entropy calculations, where appropriate, consider using alternative uncertainty measures like Gini impurity or variance if negative entropy values are persistent and problematic or recognize that in specific contexts (e.g., relative entropy), negative values may carry meaningful implications and should be interpreted accordingly.

In contrast, the other entropy measures yielded positive values. Furthermore, it should be noted that the *perimeter_worst* biomarker for benign data showed lower uncertainty values compared to its malignant counterpart. Therefore, the *perimeter_worst* biomarker demonstrates a strong ability to differentiate between diseased and non-diseased cases.

6. Final remarks and discussion

In conclusion, entropy, a foundational concept introduced by Shannon in 1948, remains central to quantifying uncertainty in random variables. It effectively captures the average information content, with a higher entropy signifying greater uncertainty and a broader probability distribution, while a lower entropy points to a more concentrated distribution. Entropy has seen wide applications across numerous scientific disciplines, from reliability studies, where it aids in understanding failure distributions, to the insurance industry, where it informs risk assessment and extreme event evaluation. Its relevance extends to neurobiology, cryptography, quantum computing, and bioinformatics, underscoring its broad utility in theoretical and applied research.

This paper explored the point and interval estimation of five key entropies; Shannon, Rényi, Havrda and Charvát, Arimoto, and Tsallis, for the Burr XII distribution under both Simple Random Sampling (SRS) and Ranked Set Sampling (RSS). We focused on evaluating the performance of maximum likelihood estimators (MLE) for these entropies and comparing their efficiency under the two sampling schemes. Through extensive simulations, we demonstrated that RSS consistently offers better efficiency, reducing the bias and variance of entropy estimators, and providing narrower confidence intervals compared to SRS. These findings confirm that RSS is a cost-effective sampling method, reducing the required sample size by at least 20% to achieve the same accuracy as SRS.

The application of these entropy measures to the Wisconsin Breast Cancer Database further highlighted the practical implications of our work. In particular, the perimeter-worst biomarker, which exhibited lower uncertainty in benign cases compared to malignant ones, proved effective in distinguishing between diseased and non-diseased cases. Although the MLE of Shannon entropy resulted in negative values for both datasets, an inherent limitation of Shannon entropy in continuous distributions, other entropy measures provided positive and interpretable values, reinforcing the utility of diverse entropy measures in real-life applications.

Overall, this study demonstrates the value of entropy as a tool for uncertainty quantification and the superior performance of RSS over SRS in estimating entropy measures. The continued refinement of sampling methods like RSS, coupled with advancements in entropy-based estimators, will further enhance the efficiency of data-driven research across various scientific fields.

REFERENCES

1. A. A. Al-Babtain, I. Elbatal, C. Chesneau, and M. Elgarhy, *Estimation of different types of entropies for the Kumaraswamy distribution*, PLoS One, vol. 16, no. 3, pp. e0249027, 2021.
2. E. K. Al-Hussaini and Z. F. Jaheen, *Bayesian estimation of the parameters, reliability, and failure rate functions of the Burr type XII failure model*, Journal of Statistical Computation and Simulation, vol. 41, no. 1-2, pp. 31–40, 1992.
3. M. F. Al-Saleh and M. A. Al-Kadiri, *Double ranked set sampling*, Statistics and Probability Letters, vol. 48, no. 2, pp. 205–212, 2000.
4. S. Arimoto, *Information-theoretical considerations on estimation problems*, Information and Control, vol. 19, no. 3, pp. 181–194, 1971.
5. R. A. Bantan, M. Elgarhy, C. Chesneau, and F. Jamal, *Estimation of entropy for inverse Lomax distribution under multiple censored data*, Entropy, vol. 22, no. 6, pp. 601, 2020.
6. G. Casella and R. Berger, *Statistical inference*, CRC Press, 2024.
7. Y. Cho, H. Sun, and K. Lee, *An estimation of the entropy for a Rayleigh distribution based on doubly-generalized Type-II hybrid censored samples*, Entropy, vol. 16, no. 7, pp. 3655–3669, 2014.
8. T. M. Cover, *Elements of information theory*, John Wiley & Sons, 1999.
9. E. Cramer and C. Bagh, *Minimum and maximum information censoring plans in progressive censoring*, Communications in Statistics-Theory and Methods, vol. 40, no. 14, pp. 2511–2527, 2011.
10. G. Dong, M. K. Shakhatreh, and D. He, *Bayesian analysis for the Shannon entropy of the Lomax distribution using noninformative priors*, Journal of Statistical Computation and Simulation, vol. 94, no. 6, pp. 1317–1338, 2024.
11. A. Eftekharian, H. Samawi, and M. Amiri, *On Kernel-based estimation of distribution function and its quantities based on ranked set sampling*, Journal of Statistical Computation and Simulation, vol. 93, no. 11, pp. 1772–1798, 2023.
12. A. Golan, *Information and entropy econometrics: A review and synthesis*, Foundations and Trends in Econometrics, vol. 2, no. 1–2, pp. 1–145, 2008.
13. A. Helu, H. Samawi, and M. Alslman, *Comparing some iterative methods of parameter estimation for progressively censored Lomax data*, Thailand Statistician, vol. 22, no. 3, pp. 533–546, 2024.
14. A. Helu, *Inference on overlapping coefficient in two exponential populations based on adaptive Type-II progressive hybrid censoring*, Afrika Matematika, vol. 36, no. 1, pp. 9, 2025.
15. J. Havrda and F. Charvát, *Quantification method of classification processes. Concept of structural a-entropy*, Kybernetika, vol. 3, no. 1, pp. 30–35, 1967.

16. M. J. Jozani and J. Ahmadi, *On uncertainty and information properties of ranked set samples*, Information Sciences, vol. 264, pp. 291–301, 2014.

17. A. Kaur, G. P. Patil, A. K. Sinha, and C. Taillie, *Ranked set sampling: An annotated bibliography*, Environmental and Ecological Statistics, vol. 2, pp. 25–54, 1995.

18. M. S. Kotb, H. A. Newer, and M. M. Mohie El-Din, *Bayesian inference for the entropy of the Rayleigh model based on ordered ranked set sampling*, Annals of Data Science, pp. 1–24, 2024.

19. K. Kumar, I. Kumar, and H. K. T. Ng, *On estimation of Shannon's entropy of Maxwell distribution based on progressively first-failure censored data*, Stats, vol. 7, no. 1, pp. 138–159, 2024.

20. M. S. Madukaife and H. D. Phuc, *Estimation of Shannon differential entropy: An extensive comparative review*, arXiv preprint arXiv:2406.19432, 2024.

21. G. A. McIntyre, *A method for unbiased selective sampling using ranked set*, Australian Journal of Agricultural Research, vol. 3, pp. 385–390, 1952.

22. H. A. Muttlak, *Median ranked set sampling*, Journal of Applied Statistical Science, vol. 6, no. 4, pp. 245–255, 1997.

23. A. Namdari and Z. Li, *A review of entropy measures for uncertainty quantification of stochastic processes*, Advances in Mechanical Engineering, vol. 11, no. 6, pp. 687814019857350, 2019.

24. G. P. Patil, A. K. Sinha, and C. Taillie, *Ranked set sampling: A bibliography*, Environmental and Ecological Statistics, vol. 6, pp. 91–98, 1999.

25. J. C. Principe, D. Xu, and D. Erdogmuns, *Renyi·s entropy, divergence and their nonparametric estimators*, Information Theoretic Learning: Renyi's Entropy and Kernel Perspectives, pp. 47–102, 2010.

26. H. Ren, Q. Gong, and X. Hu, *Estimation of entropy for generalized Rayleigh distribution under progressively Type-II censored samples*, Axioms, vol. 12, no. 8, pp. 776, 2023.

27. A. Rényi, *On measures of entropy and information*, Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, vol. 4, pp. 547–562, University of California Press, 1961.

28. D. W. Robinson, *Entropy and uncertainty*, Entropy, vol. 10, no. 4, pp. 493–506, 2008.

29. H. M. Samawi, M. S. Ahmed, and W. Abu Dayyeh, *Estimating the population mean using extreme ranked set sampling*, Biometrical Journal, vol. 38, no. 5, pp. 577–586, 1996.

30. H. M. Samawi and O. A. Al-Sageer, *On the estimation of the distribution function using extreme and median ranked set sampling*, Biometrical Journal, vol. 43, no. 3, pp. 357–373, 2001.

31. C. E. Shannon, *A mathematical theory of communication*, The Bell System Technical Journal, vol. 27, no. 3, pp. 379–423, 1948.

32. M. Shrahili, A. R. El-Saeed, A. S. Hassan, I. Elbatal, and M. Elgarhy, *Estimation of entropy for log-logistic distribution under progressive Type-II censoring*, Journal of Nanomaterials, vol. 2022, no. 1, pp. 2739606, 2022.

33. D. Sigmon and S. Fatima, *Fine needle aspiration*, Treasure Island (FL): StatPearls Publishing, 2022.

34. V. P. Singh, *Entropy theory and its application in environmental and water engineering*, John Wiley & Sons, 2013.

35. A. A. Soliman, N. A. Abou-Elheggag, A. H. Abd Ellah, and A. A. Modhesh, *Bayesian and non-Bayesian inferences of the Burr-XII distribution for progressive first-failure censored data*, Metron, vol. 70, pp. 1–25, 2012.

36. S. Tahmasebi, M. Longobardi, M. R. Kazemi, and M. Alizadeh, *Cumulative Tsallis entropy for maximum ranked set sampling with unequal samples*, Physica A: Statistical Mechanics and its Applications, vol. 556, pp. 124763, 2020.

37. C. Tsallis, *Possible generalization of Boltzmann-Gibbs statistics*, Journal of Statistical Physics, vol. 52, pp. 479–487, 1988.

38. X. Wang and W. Gui, *Bayesian estimation of entropy for Burr Type-XII distribution under progressive Type-II censored data*, Mathematics, vol. 9, no. 4, pp. 313, 2021.

39. J. Yu, W. Gui, and Y. Shan, *Statistical inference on the Shannon entropy of inverse Weibull distribution under the progressive first-failure censoring*, Entropy, vol. 21, no. 12, pp. 1209, 2019.

40. E. Zamanzade and M. Mahdizadeh, *Entropy estimation from ranked set samples with application to test of fit*, Revista Colombiana de Estadística, vol. 40, no. 2, pp. 223–241, 2017.

41. E. Zamanzade, M. Mahdizadeh, and H. Samawi, *Nonparametric estimation of mean residual lifetime in ranked set sampling with a concomitant variable*, Journal of Applied Statistics, 2024. DOI: 10.1080/02664763.2023.2301334.