

Entropy Bayesian Analysis for the Generalized Inverse Exponential Distribution Based on URRSS

Amer I. Al-Omari¹, Amal S. Hassan², Heba F. Nagy², Ayed R. A. Al-Anzi^{3,*} and Loai Alzoubi¹

 ¹Faculty of Science, Department of Mathematics, Al al-Bayt University, Mafraq, 25113, Jordan
 ²Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, 12613, Egypt
 ³Department of Mathematics, College of Science and Human Studies at Hotat Sudair, Majmaah University, Majmaah, 11952, Saudia Arabia
 *Corresponding Author: Ayed R. A. Al-Anzi. Email: a.alanzi@mu.edu.sa Received: 31 March 2021; Accepted: 03 May 2021

> Abstract: This paper deals with the Bayesian estimation of Shannon entropy for the generalized inverse exponential distribution. Assuming that the observed samples are taken from the upper record ranked set sampling (URRSS) and upper record values (URV) schemes. Formulas of Bayesian estimators are derived depending on a gamma prior distribution considering the squared error, linear exponential and precautionary loss functions, in addition, we obtain Bayesian credible intervals. The random-walk Metropolis-Hastings algorithm is handled to generate Markov chain Monte Carlo samples from the posterior distribution. Then, the behavior of the estimates is examined at various record values. The output of the study shows that the entropy Bayesian estimates under URRSS are more convenient than the other estimates under URV in the majority of the situations. Also, the entropy Bayesian estimates perform well as the number of records increases. The obtained results validate the usefulness and efficiency of the URV method. Real data is analyzed for more clarifying purposes which validate the theoretical results.

> **Keywords:** Shannon entropy; generalized inverse exponential distribution; Bayesian estimators; loss function; ranked set sampling; markov chain

1 Introduction

Record values are crucial in many areas of real life applications comprising data relating to weather, sports, economics and life testing studies. Reference [1] constructed the theory of record values as a model for successive extremes in a sequence of independently and identically distributed (iid) random variables. Reference [2] mentioned that an observation is called upper (lower) record value if its value more (less than) that all of the preceding observations. Let x_i , $i \ge 1$ be a sequence of iid random variables with a cumulative distribution function (CDF), say F(x), and probability density function (PDF), say f(x), an observation x_i is called upper record value if its value exceeds all the preceding values, i.e., x_i is an URV if $x_i > x_i$, where i > j.



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Let $T_1 = t_1, T_2 = t_2, ..., T_n = t_n$ be the first *n* URV arising from any distribution with a certain PDF and CDF, the joint PDF of the first n *URV* is given by:

$$f(t_1, t_2, \dots, t_n) = f(t_n) \prod_{i=1}^{n-1} \frac{f(t_i, \theta)}{1 - F(t_i, \theta)}, \quad -\infty < t_1 < t_2 < \dots < t_n < \infty, \ \theta \in \Theta,$$
(1)

where Θ is the parameter space and $\theta \in \Theta$ may be a vector.

Another record sampling scheme, known as upper record ranked set sampling, has been provided by [3]. This scheme is valuable in some situations when the used observations are the last record data such as athletic, weather and Olympic data. The URRSS can be described as follows:

Consider *n* independent sequences of continuous random variables, the *i*th sequence sampling is stopped when the *i*th record value is noticed. The only observations that are handled by analysis are the last record values in each sequence. Let the last record value of the *i*th sequence in this situation, say $T_{i,i}$, then the accessible observations are $T = (T_{1,1}, T_{2,2}, ..., T_{n,n})^T$, that is

1:
$$\underline{T}_{(1)1} \rightarrow T_{1,1} = T_{(1)1}$$

2: $\overline{T}_{(1)2} \underline{T}_{(2)2} \rightarrow T_{2,2} = T_{(2)2}$
:
 $n: T_{(1)n} T_{(2)n} \dots T_{(n)n} \rightarrow T_{n,n} = T_{(n)n},$

where $T_{(i)j}$ is the *i*th record in the *j*th cycle. Let $\underline{T_{i,i}} = (t_{1,1}, t_{2,2}, \dots, t_{n,n})^T$, $T_{i,i}$ be a set of observed URRSS, then according to [3], the joint PDF of $\overline{T_{i,i}}$, is given by

$$f(t_{1,1}, t_{2,2}, \dots, t_{n,n}) = \prod_{i=1}^{n} \frac{\left[-\ln(1 - F(t_{i,i};\theta))\right]^{i-1}}{(i-1)!} f(t_{i,i};\theta).$$
⁽²⁾

For information about ranked set sampling, see [4] for imputation of the missing observations using RSS and [5] for mean estimation based on modified robust extreme ranked set sampling. References [6–8] for partial, mixed and varied RSS methods, respectively. Reference [9] proposed a new RSS technique for mean and variance estimation, as well as [10] investigated the estimation of a symmetric distribution function in multistage ranked set sampling.

Some researchers have considered inference about different distributions based on records. For instance, Bayesian estimators and predictions for some life distributions from record values are discussed by [11]. Stress-strength reliability estimator of exponentiated inverted Weibull distribution values has been discussed by [12] based on lower record. Reference [13] considered Bayesian and non-Bayesian estimators from power Lomax distribution using URV. Estimation of the two-parameter bathtub-shaped distribution is discussed by [14] from record data. Bayesian estimators of the generalized inverse exponential (GIE) distribution are discussed by [15] via URV. Stress strength reliability estimator for independent GIE distributions using URRSS is handled by [16]. Reference [17] discussed estimation and prediction for Nadarajah-Haghighi distribution based on record. Statistical inference for the power Lindley model is studied by [18] from record values and inter-record times. Reference [19] handled reliability estimator for Weibull distribution for multicomponent system based on URV.

CMC, 2021, vol.69, no.3

Reference [20] introduced the concept of entropy as a measure of information, which provides a quantitative measure of the uncertainty. It is also considered as a measure of randomness of a probabilistic system. Let X be a non-negative random variable with cumulative distribution function F(x) and probability density function f(x). The Shannon entropy, denoted by SH(X), of the random variable is defined by

$$SH(X) = -\int_{-\infty}^{\infty} f(x)\log f(x)dx.$$
(3)

It is seen that a very sharply peaked distribution has very low entropy, whereas if the probability is spread out, the entropy is much higher. In this sense, SH(X) is a measure of uncertainty associated with f(x). Entropy estimation for some life distributions has been discussed by many authors. For example, [21] obtained an entropy estimator using URV from the generalized halflogistic distribution. References [22,23] suggested some entropy estimators based on RSS and double RSS methods, respectively. Reference [24] investigated entropy estimation and goodnessof-fit tests for the Laplace and inverse Gaussian distributions based on pair RSS. Reference [25] discussed the entropy Bayesian estimators of Weibull distribution based on generalized progressive hybrid censoring scheme. Reference [26] proposed new measures of entropy and [27] discussed the entropy maximum likelihood and Bayesian estimators of inverse Weibull distribution under generalized progressive hybrid censoring scheme. Reference [28] provided an exact expression for entropy information contained in both types of progressively hybrid censored data and applied it in exponential distribution. Reference [29] discussed the estimation of entropy for generalized exponential distribution via record values. Reference [30] discussed entropy estimators of a continuous random variable using RSS. Reference [31] obtained the maximum likelihood estimator of Shannon entropy for inverse Weibull distribution under multiple censored data and [32] proposed entropy Bayesian estimators of Lomax distribution using record data, and [33] considered extropy properties of RSS.

To our knowledge, in the literature, there are no studies that had been performed about entropy estimation in view of URRSS. So, our interest in this study is estimating the Shannon entropy of the GIE distribution using Bayesian approach from URRSS and URV. The Shannon entropy Bayesian estimator is considered using gamma priors. The Bayesian estimator of entropy is induced related to symmetric and asymmetric loss functions. The proposed loss functions are squared error loss function (SELF), linear exponential loss function (LINEX) and precautionary loss function (PRLF). Bayesian entropy estimators under symmetric and asymmetric loss functions have complicated expressions, so we implemented the Markov Chain Monte Carlo (MCMC) technique.

The following sections are organized as follows. Formula of Shannon entropy for GIE distribution is provided in Section 2. Entropy Bayesian estimator is derived using URRSS from symmetric and asymmetric loss functions in Section 3. Based on URV, entropy Bayesian estimator for GIE distribution is discussed using the proposed loss functions in Section 4. Simulation issue and application to real data are given in Sections 5 and 6, respectively. The paper ends with some concluding remarks in Section 7.

2 Expression of Shannon Entropy

The two-parameter GIE distribution is provided by [34] which has many applications in various areas such as, accelerated life testing, queues, horse racing, sea currents and wind speeds. The PDF of the GIE model with the shape parameter θ and scale parameter β is given by

$$f(x;\theta,\beta) = \frac{\beta\theta}{x^2} e^{-\beta/x} (1 - e^{-\beta/x})^{\theta-1}; \quad x,\theta,\beta > 0.$$

$$\tag{4}$$

The CDF of the GIE distribution is given by

$$F(x;\theta,\beta) = 1 - (1 - e^{-\beta/x})^{\theta}.$$
(5)

Let X be a random variable follows a GIE distribution with PDF given in (4), hence the Shannon entropy of X is obtained by substituting (4) in (3) as follows:

$$SH(X) = -[\ln \theta + \ln \beta - 2I_1 + I_2 - I_3], \tag{6}$$

where

$$I_{1} = \int_{-\infty}^{\infty} \frac{\beta\theta}{x^{2}} (1 - e^{-\beta/x})^{\theta - 1} e^{-\beta/x} \ln x dx, \quad I_{2} = (\theta - 1) \int_{0}^{\infty} \frac{\beta\theta}{x^{2}} (1 - e^{-\beta/x})^{\theta - 1} e^{-\beta/x} \ln(1 - e^{-\beta/x}) dx,$$

and $I_3 = \int_{0}^{\infty} \frac{\beta\theta}{x^2} (1 - e^{-\beta/x})^{\alpha - 1} \frac{\beta}{x} e^{-\beta/x} dx$. To obtain I_1 , we use the binomial expansion as follows

$$I_1 = \sum_{j=0}^{\infty} (-1)^j {\binom{\theta-1}{j}} \int_0^\infty \frac{\beta\theta}{x^2} e^{\frac{-\beta(j+1)}{x}} \ln x dx = \sum_{j=0}^\infty \frac{(-1)^j \theta}{j+1} {\binom{\theta-1}{j}} [\ln(\beta(j+1)+\gamma],$$

where $\gamma = 0.577$ is Euler constant. To obtain I_2 , let $y = 1 - e^{-\beta/x}$, then $I_2 = \theta(\theta - 1)$ $\left[\left\{ \frac{y^{\theta}}{\theta} \ln y \right\}_0^1 - \int_0^1 \frac{y^{\theta - 1}}{\theta} dy \right] = \frac{-(\theta - 1)}{\theta}.$

Also, I_3 is obtained as follows

$$I_{3} = \sum_{j=0}^{\infty} (-1)^{j} {\binom{\theta-1}{j}} \int_{0}^{\infty} \frac{\theta\beta^{2}}{x^{3}} e^{\frac{-\beta(j+1)}{x}} dx = \sum_{j=0}^{\infty} \frac{(-1)^{j}\theta}{(j+1)^{2}} {\binom{\theta-1}{j}}.$$

Substituting I_1 , I_2 , and I_3 in (6), we obtain the Shannon entropy for GIE distribution as follows:

$$SH(x) = -\left[\ln\beta + \ln\theta - 2\sum_{j=0}^{\infty} \frac{(-1)^{j}\theta}{j+1} \binom{\theta-1}{j} \left[\ln(\beta(j+1) + \gamma + \frac{1}{j+1}] - \frac{\theta-1}{\theta}\right],\tag{7}$$

which is a function of the parameters θ and β .

3 Entropy Bayesian Estimation Based on URRSS

In this section, Bayesian estimator of the Shannon entropy for the GIE model is discussed in view of URRSS. Firstly, the Bayesian estimators of parameters must be computed in order to get the entropy Bayesian estimator. Then, entropy Bayesian estimator is obtained using (7) according to the invariance property. The Bayesian estimator based on gamma priors is considered. Three Bayesian estimators are obtained according to SELF, LINEX and PRLF. Furthermore, the Bayesian credible intervals are constructed.

Let $\underline{t_{i,i}} = (t_{1,1}, t_{2,2}, \dots, t_{n,n})$ be a set of observed URRSS from GIE distribution, then the likelihood function denoted by L_1 , is obtained by inserting PDF in (4) and CDF in (5) in (2), as follows

$$L_{1} = \theta^{\sum_{i=1}^{n} i} \beta^{n} \prod_{i=1}^{n} \frac{e^{-\beta/t_{i,i}} t_{i,i}^{-2}}{(i-1)!} [-\log(1-e^{-\beta/t_{i,i}})]^{i-1} (1-e^{-\beta/t_{i,i}})^{\theta-1}.$$

Assuming that the prior of parameters θ and β has a gamma distribution with parameters (a, b) and (c, d), respectively. Hence, the joint prior distribution of parameters, denoted by $\pi(\theta, \beta)$, assuming independence of parameters is as follows

$$\pi(\theta,\beta) = \frac{1}{\Gamma(a)\,\Gamma(c)}\,\theta^{a-1}\,\beta^{c-1}\,e^{-b\theta-d\beta}\,;\quad a,b,c,d,\theta,\beta>0.$$
(8)

The joint posterior under the assumption that β and θ are independent gamma priors is

$$\Pi_{1}^{\bullet}(\theta,\beta \mid \underline{t_{i,i}}) \propto \theta^{\sum_{i=1}^{n} i+a-1} \beta^{n+c-1} e^{-b\theta-d\beta} \prod_{i=1}^{n} \frac{e^{-\beta/t_{i,i}} t_{i,i}^{-2}}{(i-1)!} [-\log(1-e^{-\beta/t_{i,i}})]^{i-1} (1-e^{-\beta/t_{i,i}})^{\theta-1}.$$

Hence, the marginal posterior distributions of β and θ are given by

$$\begin{aligned} \Pi_{1}^{\bullet}(\theta \mid \underline{t_{i,i}}) &= D_{1} \; \theta^{\sum_{i=1}^{n} i+a-1} e^{-b\theta} \prod_{i=1}^{n} \; \frac{t_{i,i}^{-2}}{(i-1)!} \\ & \times \int_{0}^{\infty} \beta^{n+c-1} e^{-\beta(d+t_{i,i}^{-1})} [-\log\left(1-e^{-\beta/t_{i,i}}\right)]^{i-1} (1-e^{-\beta/t_{i,i}})^{\theta-1} d\beta \\ \Pi_{1}^{\bullet}(\beta \mid \underline{t_{i,i}}) &= D_{1} \beta^{n+c-1} e^{-d\beta} \prod_{i=1}^{n} \; \frac{e^{-\beta/t_{i,i}} t_{i,i}^{-2}}{(i-1)!} [-\log\left(1-e^{-\beta/t_{i,i}}\right)]^{i-1} \\ & \times \int_{0}^{\infty} \theta^{(\sum_{i=1}^{n} i)+a-1} e^{-b_{1}\theta} \left(1-e^{-\beta/t_{i,i}}\right)^{\theta-1} \; d\theta, \end{aligned}$$

where

$$D_1^{-1} = \prod_{i=1}^n \frac{t_{i,i}^{-2}}{(i-1)!} \int_0^\infty \int_0^\infty \theta_{i=1}^{\sum_{i=1}^n i+a-1} \beta^{n+c-1} e^{-b\theta - d\beta} e^{-\beta/t_{i,i}} \left[-\log\left(1 - e^{-\beta/t_{i,i}}\right)\right]^{i-1} \left(1 - e^{-\beta/t_{i,i}}\right)^{\theta - 1} d\theta d\beta.$$

Therefore, the Bayesian estimators of β and θ under SELF, denoted by $\tilde{\beta}_1$ and $\tilde{\theta}_1$, depending on URRSS are obtained as the posterior mean as follows:

$$\tilde{\theta}_1 = \int_0^\infty \theta \,\Pi_1^{\bullet}(\theta \mid \underline{t_{i,i}}) \,d\theta, \quad \tilde{\beta}_1 = \int_0^\infty \beta \,\Pi_1^{\bullet}(\beta \mid \underline{t_{i,i}}) \,d\beta.$$
(9)

The Bayesian estimators of β and θ under LINEX, denoted by $\ddot{\beta}_1$ and $\ddot{\theta}_1$, are given by

$$\ddot{\theta}_{1} = \frac{-1}{\delta} \log E(e^{-\delta\theta}) = \frac{-1}{\delta} \log \left[\int_{0}^{\infty} e^{-\delta\theta} \ \Pi_{1}^{\bullet}(\theta \mid \underline{t_{i,i}}) \, d\theta \right],\tag{10}$$

and

$$\ddot{\beta}_{1} = \frac{-1}{\delta} \log E(e^{-\delta\beta}) = \frac{-1}{\delta} \log \left[\int_{0}^{\infty} e^{-\delta\beta} \ \Pi_{1}^{\bullet}(\beta \mid \underline{t_{i,i}}) \, d\beta \right],\tag{11}$$

where δ is a real number. Additionally, the Bayesian estimators of β and θ under PRLF, say $\tilde{\tilde{\theta}}_1$ and $\tilde{\tilde{\beta}}_1$ are given as follows

$$\tilde{\tilde{\theta}}_{1} = \sqrt{E(\theta^{2} \mid t)} = \sqrt{\int_{0}^{\infty} \theta^{2} \Pi_{1}^{\bullet}(\beta \mid \underline{t_{i,i}}) d\theta},$$
(12)

and

$$\tilde{\tilde{\beta}}_{1} = \sqrt{E(\beta^{2} \mid t)} = \sqrt{\int_{0}^{\infty} \beta^{2} \Pi_{1}^{\bullet}(\beta \mid \underline{t_{i,i}}) d\beta}.$$
(13)

The integrals (9)–(13) are very hard to be solved analytically according to their convoluted forms. Therefore, we employ the MCMC technique to approximate these integrations. The Bayesian estimates together with credible intervals width under SELF, LINEX and PRLF loss functions are implemented using Metropolis-Hastings (M-H) algorithm. Therefore, the Bayes estimate of SH(X), denoted by $S\tilde{H}_1(x)$ under SELF is obtained as follows

$$S\tilde{H}_{1}(X) = -\left[\ln\tilde{\beta}_{1} + \ln\tilde{\theta}_{1} - 2\sum_{j=0}^{\infty} \frac{(-1)^{j}\tilde{\theta}_{1}}{j+1} \binom{\tilde{\theta}_{1} - 1}{j} [\ln(\tilde{\beta}_{1}(j+1)) + \gamma + 1] - \frac{\tilde{\theta}_{1} - 1}{\tilde{\theta}_{1}}\right]$$

3800

CMC, 2021, vol.69, no.3

Consequently, the Bayesian estimator of SH(X) under LINEX and PRLF are obtained by similar way after setting their estimators in (7). Additionally, we get the Bayesian credible interval of entropy using the same algorithm proposed by [35].

4 Entropy Bayesian Estimation Based on URV

This section provides the Bayesian estimators of θ and β for the GIE distribution based on URV. The Bayesian estimators are obtained assuming that the gamma priors are independent using SELF, LINEX and PRLF. Let $\underline{t} = (t_1, t_2, ..., t_n)$ be *n* observed URV from GIE distribution with PDF in (4) and CDF in (5), then the likelihood function, say L_2 , of the GIE distribution is obtained by inserting (4) and (5) in (1), as follows:

$$L_2 = (1 - e^{-\beta/t_n})^{\theta} \prod_{i=1}^n \beta \theta t_i^{-2} e^{-\beta/t_i} (1 - e^{-\beta/t_i})^{-1}.$$

Assuming that the prior of θ and β has a gamma distribution with parameters (a, b) and (c, d), respectively. Hence, the joint prior distribution of parameters, assuming independence is considered as provided in (8). Therefore, the joint posterior can be expressed as follows:

$$\Pi_{2}^{\bullet}(\theta,\beta \mid \underline{t}) \propto \theta^{n+a-1} \beta^{n+c-1} e^{-b\theta - d\beta} (1 - e^{-\beta/t_n})^{\theta} \prod_{i=1}^{n} t_i^{-2} e^{-\beta/t_i} (1 - e^{-\beta/t_i})^{-1}.$$

Consequently, expressions for the marginal posterior distributions of θ and β are as follows:

$$\Pi_{2}^{\bullet}(\theta \mid \underline{t}) = D_{2}\theta^{n+a-1}e^{-b\theta} \int_{0}^{\infty} \beta^{n+c-1}e^{-d\beta} \left(1 - e^{-\beta/t_{n}}\right)^{\theta} \prod_{i=1}^{n} t_{i}^{-2}e^{-\beta/t_{i}} \left(1 - e^{-\beta/t_{i}}\right)^{-1}d\beta,$$

$$\Pi_{2}^{\bullet}(\beta \mid \underline{t}) = D_{2}\beta^{n+c-1}e^{-d\beta} \prod_{i=1}^{n} t_{i}^{-2}e^{-\beta/t_{i}} \left(1 - e^{-\beta/t_{i}}\right)^{-1} \int_{0}^{\infty} \theta^{n+a-1}e^{-b\theta} \left(1 - e^{-\beta/t_{n}}\right)^{\theta}d\theta,$$

where

$$D_2^{-1} = \int_0^\infty \int_0^\infty \theta^{n+a-1} \beta^{n+c-1} e^{-b\theta-c\beta} (1-e^{-\beta/t_n})^\theta \prod_{i=1}^n t_i^{-2} e^{-\beta/t_i} (1-e^{-\beta/t_i})^{-1} d\theta d\beta.$$

Hence, Bayesian estimators of θ and β , under SELF, say $\tilde{\theta}_2$ and $\tilde{\beta}_2$, can be obtained as posterior mean as follows:

$$\tilde{\theta}_2 = E(\theta \mid \underline{t}) = \int_0^\infty \theta \ \Pi_2^{\bullet}(\theta \mid \underline{t}) \ d\theta, \ \tilde{\beta}_2 = E(\beta \mid \underline{t}) = \int_0^\infty \beta \ \Pi_2^{\bullet}(\beta \mid \underline{t}) \ d\beta.$$
(14)

Also, under LINEX, the Bayesian estimators of θ and β , say $\ddot{\theta}_2$ and $\ddot{\beta}_2$, are obtained as follows:

$$\ddot{\theta}_{2} = \frac{-1}{\delta} \log \left[\int_{0}^{\infty} e^{-\delta\theta} \Pi_{2}^{\bullet}(\theta \mid \underline{t}) \, d\theta \right], \quad \text{and} \quad \ddot{\beta}_{2} = \frac{-1}{\delta} \log \left[\int_{0}^{\infty} e^{-\delta\beta} \Pi_{2}^{\bullet}(\beta \mid \underline{t}) \, d\beta \right]. \tag{15}$$

Furthermore, considering PRLF, the Bayesian estimators of θ and β , say $\tilde{\tilde{\theta}}_2$ and $\tilde{\tilde{\beta}}_2$ are given as follows:

$$\tilde{\tilde{\theta}}_2 = \sqrt{\int_0^\infty \theta^2 \Pi_2^{\bullet}(\theta \mid \underline{t}) d\theta}, \quad \text{and} \quad \tilde{\tilde{\beta}}_2 = \sqrt{\int_0^\infty \beta^2 \Pi_2^{\bullet}(\beta \mid \underline{t}) d\beta}.$$
(16)

Again, the MCMC procedure is provided to approximate the integrals (14)–(16) based on M-H algorithm to compute the estimates and credible interval width considering symmetric and asymmetric loss functions.

Regarding to Eq. (7), the Bayesian estimator of SH(x), denoted by $S\tilde{H}_2(x)$ under SELF is obtained as follows

$$S\tilde{H}_2(X) = -\left[\ln\tilde{\beta}_2 + \ln\tilde{\theta}_2 - 2\sum_{j=0}^{\infty} \frac{(-1)^j\tilde{\theta}_2}{j+1} \begin{pmatrix}\tilde{\theta}_2 - 1\\j \end{pmatrix} \left[\ln(\tilde{\beta}_2(j+1)) - \gamma + \frac{1}{j+1}\right] - \frac{\tilde{\theta}_2 - 1}{\tilde{\theta}_2}\right].$$

By similar way, the Bayesian estimator of SH(X) under LINEX and PRLF are obtained after setting their estimators in (7). Furthermore, the Bayesian credible interval is obtained as mentioned in the Section 3.

5 Simulation Study

In this section, a simulation investigation is carried out to compare the performance of the entropy estimate of the GIE distribution based on URV and URRSS. The relative absolute bias (RAB), estimated risk (ER) and width (WD) of credible intervals for the Shannon entropy based on URV and URRSS for GIE distribution are used to evaluate the behaviour of the Bayesian estimates. In the simulation setup, the number of records are selected as n = 4, 5, 6, 7. The values of parameters are selected as $(\theta, \beta) = (4, 2), (2, 2)$ and (0.5, 2), where the associated true values of entropy are SH(x) = 0.8452, 1.3584 and 3.2896, respectively. The hyper-parameters for gamma prior are selected as a = b = 2 and c = d = 2. Also, we take $\delta = -2$, 2 for LINEX loss function. M-H algorithm will be used via R 3.1.2 program.

The M-H algorithm procedure is described as follows:

Let g(.) be the density of subject distribution. Initialize a starting value x_0 and the number of samples Nfor i = 2 to Nset $x = x_{i-1}$ generate u from U(0, 1) generate y from g(.) if $u \leq \frac{\pi_{\alpha}(y) g(x)}{\pi_{\alpha}(x) g(y)}$, then set $x_i = y$ else set $x_i = x$ end if end for

Tabs. 1-3 summarize the Bayes estimates and their measures (RAB, ER and WD) based on URV and URRSS. From the numerical outcomes given in Tabs. 1-3 and Figs. 1-6, we can conclude the following:

Table 1: Entropy estimates, RAB, ER and WD based on URV and URRSS at $(\theta, \beta) = (4, 2)$ and SH(x) = 0.8452

n	Loss function	Scheme	Estimate	RAB	ER	WD
4	LINEX ($\delta = -2$)	URV	0.8382	8.20E-03	9.59E-09	0.0601
		URRSS	1.0634	2.58E-01	9.53E-06	0.5085
	LINEX ($\delta = 2$)	URV	0.7993	5.43E-02	4.21E-07	0.1159
		URRSS	0.8465	1.58E-03	3.55E-10	1.1973
	PRLF	URV	0.9010	6.60E-02	6.23E-07	0.1051
		URRSS	0.7206	1.47E-01	3.10E-06	0.3732
	SELF	URV	0.8463	1.40E-03	2.78E-10	0.0024
		URRSS	0.5260	3.78E-01	2.04E-05	0.7587
5	LINEX ($\delta = -2$)	URV	0.8443	1.00E-03	1.33E-10	0.0066
	· · · · ·	URRSS	0.9265	9.62E-02	1.32E-06	0.1177
	LINEX ($\delta = 2$)	URV	0.8401	6.00E-03	5.19E-09	0.0129
		URRSS	0.8257	1.00E-02	1.52E-10	0.0592
	PRLF	URV	0.8513	7.30E-03	7.59E-09	0.0116
		URRSS	0.8896	5.26E-02	3.96E-07	0.1412
	SELF	URV	0.8510	7.00E-03	6.94E-09	0.0119
		URRSS	0.8091	4.26E-02	2.60E-07	0.0970
6	LINEX ($\delta = -2$)	URV	0.8444	9.00E-04	1.08E-10	0.0059
		URRSS	0.8392	7.00E-03	7.00E-09	0.0146
	LINEX ($\delta = 2$)	URV	0.8406	5.40E-03	4.21E-09	0.0116
		URRSS	0.8444	9.41E-04	1.27E-10	0.0018
	PRLF	URV	0.8507	6.60E-03	6.15E-09	0.0105
		URRSS	0.8569	1.39E-02	2.74E-08	0.0212
	SELF	URV	0.8505	6.30E-03	5.62E-09	0.0107
		URRSS	0.8496	5.31E-03	4.03E-09	0.0160
7	LINEX ($\delta = -2$)	URV	0.8451	1.00E-04	1.33E-12	0.0007
		URRSS	0.8457	6.48E-04	6.00E-11	0.0014
	LINEX ($\delta = 2$)	URV	0.8446	6.00E-04	5.19E-11	0.0013
		URRSS	0.8448	3.92E-04	2.20E-11	0.0009
	PRLF	URV	0.8458	7.00E-04	7.59E-11	0.0012
		URRSS	0.8451	4.69E-05	3.14E-13	0.0005
	SELF	URV	0.8457	7.00E-04	6.94E-11	0.0012
		URRSS	0.8458	7.79E-04	8.67E-11	0.0013

Note: E–a: stands for 10^{-a} .

n	Loss function	Scheme	Estimate	RAB	ER	WD
4	LINEX ($\delta = -2$)	URV	1.3421	1.20E-02	5.31E-08	0.1316
		URRSS	1.6688	2.29E-01	1.93E-05	0.8824
	LINEX ($\delta = 2$)	URV	1.2583	7.40E-02	2.00E-06	0.2550
		URRSS	1.3556	2.03E-03	1.52E-09	0.0184
	PRLF	URV	1.4823	9.10E-02	3.07E-06	0.2334
		URRSS	2.0788	5.30E-01	1.04E-04	1.1534
	SELF	URV	1.4763	8.70E-02	2.78E-06	0.2357
		URRSS	2.2136	6.30E-01	1.46E-04	1.4540
5	LINEX ($\delta = -2$)	URV	1.3567	1.00E-03	5.31E-10	0.0132
		URRSS	1.3475	8.00E-03	2.37E-08	0.0488
	LINEX ($\delta = 2$)	URV	1.3482	8.00E-03	2.08E-08	0.0258
		URRSS	1.3616	2.40E-03	2.13E-09	0.01491
	PRLF	URV	1.3707	9.10E-02	3.04E-08	0.0232
		URRSS	1.3448	1.00E-02	3.65E-08	0.0919
	SELF	URV	1.3701	9.00E-03	2.78E-08	0.0239
		URRSS	1.3912	2.40E-02	2.16E-07	0.1017
6	LINEX ($\delta = -2$)	URV	1.3480	8.00E-03	5.13E-10	0.0132
		URRSS	1.3567	1.00E-03	5.34E-10	0.0065
	LINEX ($\delta = 2$)	URV	1.3394	1.40E-02	7.14E–08	0.0884
		URRSS	1.3639	4.04E-03	6.03E-09	0.0208
	PRLF	URV	1.4233	4.80E-02	3.04E-08	0.0232
		URRSS	1.3573	1.00E-03	2.22E-10	0.0044
	SELF	URV	1.2667	9.00E-03	2.78E-08	0.0239
		URRSS	1.3607	2.00E-03	1.06E-09	0.0067
7	LINEX ($\delta = -2$)	URV	1.3582	0.00E+00	5.31E-12	0.0013
		URRSS	1.3579	0.00E+00	4.34E-11	0.0009
	LINEX ($\delta = 2$)	URV	1.3573	1.00E-03	2.08E-10	0.0026
		URRSS	1.3578	3.93E-04	5.71E-11	0.0103
	PRLF	URV	1.3596	1.00E-03	3.04E-10	0.0023
		URRSS	1.3582	0.00E+00	6.34E-12	0.0010
	SELF	URV	1.3595	1.00E-03	2.78E-10	0.0024
		URRSS	1.3582	0.00E+00	7.55E-12	0.0009

Table 2: Entropy estimates, RAB, ER and WD based on URV and URRSS at $(\theta, \beta) = (2, 2)$ and SH(x) = 1.3584

Note: E–a: stands for 10^{-a} .

- The ER of entropy estimates under SELF and LINEX based on URRSS is smaller than that of the corresponding under URV at n=6 for all values of (θ, β) (see Figs. 1 and 2).
- The ER of entropy estimates under LINEX ($\delta = 2$) under URRSS is smaller than that of the corresponding under URV at n = 7 for all values of (θ, β) (see Fig. 3).
- The ER of entropy estimates based on URSS is smaller than the corresponding under URV at n=5, and $(\theta, \beta)=(0.5, 2)$ for different loss functions (see Fig. 4).

п	Loss function	Scheme	Estimate	RAB	ER	WD
4	LINEX ($\delta = -2$)	URV	3.3481	1.78E-02	6.84E-07	0.1290
		URRSS	3.8466	1.69E-01	6.20E-05	0.6574
	LINEX ($\delta = 2$)	URV	3.1869	3.12E-02	2.11E-06	0.2583
		URRSS	3.4590	5.15E-02	5.74E-06	1.0862
	PRLF	URV	3.4135	3.77E-02	3.07E-06	0.2336
		URRSS	3.2859	1.13E-03	2.78E-09	0.5921
	SELF	URV	3.4089	3.62E-02	2.84E-06	0.2415
		URRSS	3.2191	2.14E-02	9.94E-07	0.8571
5	LINEX ($\delta = -2$)	URV	3.2733	5.00E-03	5.31E-08	0.1316
		URRSS	3.2774	3.73E-03	3.00E-08	0.0917
	LINEX ($\delta = 2$)	URV	3.1896	3.04E-02	2.00E-06	0.2550
		URRSS	3.3409	1.56E-02	5.25E-07	0.0686
	PRLF	URV	3.4135	3.77E-02	3.07E-06	0.2334
		URRSS	3.2916	6.02E-04	7.83E-10	0.0150
	SELF	URV	3.4075	3.58E-02	2.78E-06	0.2357
		URRSS	3.3371	1.44E-02	4.50E-07	0.0793
6	LINEX ($\delta = -2$)	URV	3.2880	5.00E-04	5.31E-10	0.0132
		URRSS	3.2835	1.87E-03	7.54E-09	0.0146
	LINEX ($\delta = 2$)	URV	3.2794	3.10E-03	2.08E-08	0.0258
		URRSS	3.2950	1.65E-03	5.91E-09	0.0108
	PRLF	URV	3.3019	3.70E-03	3.04E-08	0.0232
		URRSS	3.2905	2.65E-04	1.52E-10	0.0080
	SELF	URV	3.3014	3.60E-03	2.78E-08	0.0239
		URRSS	3.2874	6.84E-04	1.01E-09	0.0131
7	LINEX ($\delta = -2$)	URV	3.2895	0.00E + 00	1.33E-12	0.0007
		URRSS	3.2897	3.12E-05	2.10E-12	0.0008
	LINEX ($\delta = 2$)	URV	3.2916	5.94E-04	7.64E-10	0.0120
		URRSS	3.2902	1.66E-04	5.96E-11	0.0150
	PRLF	URV	3.2902	2.00E-04	7.59E-11	0.0012
		URRSS	3.2903	2.21E-04	1.06E-10	0.0016
	SELF	URV	3.2902	2.00E-04	6.94E-11	0.0012
		URRSS	3.2896	7.49E–08	1.21E-17	0.0005

Table 3: Entropy estimates, RAB, ER and WD based on URV and URRSS at $(\theta, \beta) = (0.5, 2)$ and SH(x) = 3.2896

Note: E–a: stands for 10^{-a} .

- The ER of entropy estimates based on URRSS is smaller than the corresponding under URV at n=7, (θ, β) = (2, 2) for the proposed loss functions except LINEX (δ = −2) (see Fig. 5).
- The WD of entropy estimates based on URV is smaller than the corresponding under URRSS at n=4 under PRLF for all values of (θ, β) (see Fig. 6).
- In general, as *n* increases, the ER, RAB and WD of estimate decrease for both record schemes.



• As the true value SH(x) increases, the ER increases in most of the situations.

Figure 1: ER of entropy estimate based on URV and URRSS at SELF and n = 6 for all values of (θ, β)



Figure 2: ER of entropy estimate based on URV and URRSS at n = 6, and LINEX ($\delta = 2$) for all values of (θ, β)



Figure 3: ER of entropy estimate based on URV and URRSS at n = 7 and LINEX ($\delta = 2$) for all values of (θ, β)



Figure 4: ER of entropy estimate based on URV and URRSS at n = 5, SH(x) = 3.289 and different loss functions



Figure 5: ER of entropy estimate based on URV and URRSS at n = 7, SH(x) = 1.358 and different loss functions



Figure 6: WD of entropy estimate based on URV and URRSS at n = 4 and PRLF for all values of (θ, β)

6 Application to Real Data

In this section, a real data set is analysed for illustrative purposes. The suggested data represent the lifetimes of steel specimens tested at different stress levels (for more details see [36]. Some preliminary data analysis is performed. The Kolmogorov-Smirnov (K-S) test is used for the data set to the fitted model. It is observed that the K-S distance are 0.083 with the corresponding P-value 0.917. It indicates that the GIE model provides reasonable fit to this data set. Also, the estimated PDF, CDF and PP plots for data are represented in Fig. 7. From these figures it can be concluded that the GIE distribution is an adequate model to fit these data.



Figure 7: Estimated PDF, CDF and PP plots of the GIE distribution for lifetimes of steel specimens data

38.5	38	37	36
60	100	141	173
83	128	143	218
140	186	194	288
		318	394
		585	

The extracted records from a part of this data are presented as

Based on the above record data, it can be shown that URRSS of size n = 4 is $(t_{1,1}, \ldots, t_{4,4}) =$ (60, 128, 194, 394) and the URV of size n = 4 is $(t_1, \ldots, t_4) =$ (60, 83, 140, 186). Considering this record data, the entropy Bayes estimator at n = 4 under SELF, LINEX and PRLF are obtained and listed in Tab. 4.

From Tab. 4, we can conclude that ER of entropy estimates under URRSS gets the smallest values compared to the corresponding under URV in case of PRLF and LINEX ($\delta = -2$) at true

value SH(X) = 3.2896. While at true value SH(X) = 1.3584, it is noted that the ER of entropy estimates under URRSS is smaller than the corresponding under URV in case of LINEX ($\delta = 2$) and SELF. Furthermore, one can conclude that the ER of entropy estimates under URRSS are smaller than the corresponding counterparts URV at LINEX ($\delta = -2$) for true value SH(X) = 0.8452.

True entropy	$^{(heta,eta)}$	Loss function	Scheme	Estimate	RAB	ER	WD
0.8452	(4, 2)	LINEX ($\delta = -2$)	URV	0.667	0.211	6.33E-06	0.267
			URRSS	0.874	0.035	1.71E-07	0.219
		LINEX ($\delta = 2$)	URV	0.894	0.058	4.75E-07	0.115
			URRSS	0.787	0.069	6.76E-07	0.119
		PRLF	URV	0.891	0.054	4.16E-07	0.168
			URRSS	0.772	0.086	1.06E-06	0.208
		SELF	URV	0.799	0.054	4.17E-07	0.176
			URRSS	0.952	0.126	2.27E-06	0.220
1.3584	(2, 2)	LINEX ($\delta = -2$)	URV	1.342	0.012	5.31E-08	0.132
			URRSS	1.301	0.042	6.57E-07	0.156
		LINEX ($\delta = 2$)	URV	1.258	0.074	2.00E-06	0.255
			URRSS	1.373	0.011	4.11E-08	0.117
		PRLF	URV	1.343	0.011	4.74E-08	0.118
			URRSS	1.398	0.029	3.20E-07	0.143
		SELF	URV	1.409	0.037	5.17E-07	0.138
			URRSS	1.338	0.015	8.05E-08	0.119
3.2896	(0.5, 2)	LINEX ($\delta = -2$)	URV	3.347	0.017	6.61E-07	0.114
			URRSS	3.274	0.005	4.80E-08	0.180
		LINEX ($\delta = 2$)	URV	3.243	0.014	4.31E-07	0.124
			URRSS	3.234	0.017	6.16E-07	0.159
		PRLF	URV	3.362	0.022	1.06E-06	0.218
			URRSS	3.310	0.006	8.01E-08	0.250
		SELF	URV	3.278	0.003	2.49E-08	0.178
			URRSS	3.244	0.014	4.11E-07	0.106

Table 4: Entropy estimates and their RAB, ER and WD of steel specimens at different stress levels data based on URV and URRSS

7 Summary and Conclusion

This paper provides Bayesian estimation of the Shannon entropy for the generalized inverse exponential distribution using URRSS and URV shemes. The entropy Bayesian estimators are considered using gamma prior functions for symmetric (SELF) and asymmetric (LINEX and PRLF) loss functions. In order to obtain the Bayesian estimators, we employed Markov Chain Monte Carlo method based on Metropolis-Hastings algorithm. The performance of the entropy estimates for the GIE distribution is investigated in terms of their relative absolute bias, estimated risk and the width of credible intervals. From simulation results, it turns out that, the entropy Bayesian estimator approaches the true value as the number of record increases. Generally, the

entropy and ERs are directly proportional, that is; if the real value of entropy increases, the ERs increase. The WD of Bayes credible intervals for estimated values of entropy URRSS is smaller than the corresponding estimated values based on URV for all loss functions for most values of record values in the majority of the cases. A data real example has been considered to illustrate the applicability of the proposed methodology for the considered record schemes.

Acknowledgement: The authors are grateful to the Editor and anonymous reviewers for their valuable comments and suggestions.

Funding Statement: A. R. A. Alanzi would like to thank the Deanship of Scientific Research at Majmaah University for financial support and encouragement.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] K. Chandler, "The distribution and frequency of record values," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 14, no. 2, pp. 220–228, 1952.
- [2] H. Arnold, B. Balakrishnan and N. Nagaraja, Records, New York: John Wiley & Sons, 1998.
- [3] M. Salehi and J. Ahmadi, "Record ranked set sampling scheme," *Metron*, vol. 72, no. 3, pp. 351–365, 2014.
- [4] C. N. Bouza and A. I. Al-Omari, "Ranked set estimation with imputation of the missing observations: The median estimator," *Revista Investigacion Operacional*, vol. 32, no. 1, pp. 30–37, 2011.
- [5] A. I. Al-Omari, "Estimation of mean based on modified robust extreme ranked set sampling," *Journal* of Statistical Computation and Simulation, vol. 81, no. 8, pp. 1055–1066, 2011.
- [6] A. Haq, J. Brown, E. Moltchanova and A. I. Al-Omari, "Partial ranked set sampling design," *Environmetrics*, vol. 24, no. 3, pp. 201–207, 2013.
- [7] A. Haq, J. Brown, E. Moltchanova and A. I. Al-Omari, "Mixed ranked set sampling design," *Journal of Applied Statistics*, vol. 41, no. 10, pp. 2141–2156, 2014.
- [8] A. Haq, J. Brown, E. Moltchanova and A. I. Al-Omari, "Varied L ranked set sampling scheme," *Journal of Statistical Theory and Practice*, vol. 9, no. 4, pp. 741–767, 2015.
- [9] E. Zamanzade and A. I. Al-Omari, "New ranked set sampling for estimating the population mean and variance," *Hacettepe Journal of Mathematics and Statistics*, vol. 45, no. 6, pp. 1891–1905, 2016.
- [10] M. Mahdizadeh and E. Zamanzade, "Estimation of a symmetric distribution function in multistage ranked set sampling," *Statistical Papers*, vol. 61, no. 2, pp. 851–867, 2020.
- [11] J. Ahmadi and M. Doostparast, "Bayesian estimation and prediction for some life distributions based on record values," *Statistical Papers*, vol. 47, no. 3, pp. 373–392, 2006.
- [12] A. S. Hassan, H. Z. Muhammed and M. S. Saad, "Estimation of stress-strength reliability for exponentiated inverted weibull distribution based on lower record values," *Journal of Advances in Mathematics* and Computer Science, vol. 11, no. 2, pp. 1–14, 2015.
- [13] E. A. Amin, "On record values of the power lomax distribution," *Advances and Applications in Statistics*, vol. 50, no. 6, pp. 497–515, 2017.
- [14] M. Z. Raqab, O. M. Bdair and F. M. Al-Aboud, "Inference for the two-parameter bathtub-shaped distribution based on record data," *Metrika*, vol. 81, no. 3, pp. 229–253, 2018.
- [15] A. S. Hassan, M. Abd-Allah and H. F. Nagy, "Bayesian analysis of record statistics based on generalized inverted exponential model," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 8, no. 2, pp. 323–335, 2018a.
- [16] A. S. Hassan, A. Marwa and H. F. Nagy, "Estimation of P (Y > X) using record values from the generalized inverted exponential distribution," *Pakistan Journal of Statistics and Operation Research*, vol. 14, no. 3, pp. 645–660, 2018b.

- [17] M. A. Selim, "Estimation and prediction for nadarajah-haghighi distribution based on record values," *Pakistan Journal of Statististics*, vol. 34, no. 1, pp. 77–90, 2018.
- [18] A. Pak and S. Dey, "Statistical inference for the power lindley model based on record values and inter-record times," *Journal of Computational and Applied Mathematics*, vol. 347, pp. 156–172, 2019.
- [19] A. S. Hassan, H. F. Nagy, H. Z. Muhammed and M. S. Saad, "Estimation of multicomponent stressstrength reliability following weibull distribution based on upper record values," *Journal of Taibah University for Science*, vol. 14, no. 1, pp. 244–253, 2020.
- [20] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.
- [21] J.-I. Seo, H.-J. Lee and S.-B. Kang, "Estimation for generalized half logistic distribution based on records," *Journal of the Korean Data and Information Science Society*, vol. 23, no. 6, pp. 1249–1257, 2012.
- [22] A. I. Al-Omari, "Estimation of entropy using random sampling," Journal of Computational and Applied Mathematics, vol. 261, pp. 95–102, 2014.
- [23] A. I. Al-Omari, "New entropy estimators with smaller root mean squared error," Journal of Modern Applied Statistical Methods, vol. 14, no. 2, pp. 88–109, 2015.
- [24] A. I. Al-Omari and A. Haq, "Entropy estimation and goodness-of-fit tests for the inverse Gaussian and laplace distributions using paired ranked set sampling," *Journal of Statistical Computation and Simulation*, vol. 86, no. 11, pp. 2262–2272, 2016.
- [25] Y. Cho, H. Sun and K. Lee, "Estimating the entropy of a weibull distribution under generalized progressive hybrid censoring," *Entropy*, vol. 17, no. 1, pp. 102–122, 2015.
- [26] A. I. Al-Omari, "A new measure of entropy of continuous random variable," Journal of Statistical Theory and Practice, vol. 10, no. 4, pp. 721–735, 2016.
- [27] K. Lee, "Estimation of entropy of the inverse weibull distribution under generalized progressive hybrid censored data," *Journal of the Korean Data & Information Science Society*, vol. 28, no. 3, pp. 659–668, 2017.
- [28] A. Bader, "On the entropy of progressive hybrid censoring schemes," *Applied Mathematics & Information Sciences*, vol. 11, no. 6, pp. 1811–1814, 2017.
- [29] M. Chacko and P. Asha, "Estimation of entropy for generalized exponential distribution based on record values," *Journal of the Indian Society for Probability and Statistics*, vol. 19, no. 1, pp. 79–96, 2018.
- [30] A. I. Al-Omari and A. Haq, "Novel entropy estimators of a continuous random variable," *International Journal of Modeling, Simulation, and Scientific Computing*," vol. 10, no. 2, pp. 1950004, 2019.
- [31] A. S. Hassan and A. N. Zaky, "Estimation of entropy for inverse weibull distribution under multiple censored data," *Journal of Taibah University for Science*, vol. 13, no. 1, pp. 331–337, 2019.
- [32] A. S. Hassan and A. N. Zaky, "Entropy bayesian estimation for lomax distribution based on record," *Thailand Statistician*, vol. 19, no. 1, pp. 95–114, 2021.
- [33] M. Raqab and G. Qiu, "On extropy properties of ranked set sampling," *Statistics*, vol. 53, no. 1, pp. 210–226, 2019.
- [34] A. Abouammoh and A. M. Alshingiti, "Reliability estimation of generalized inverted exponential distribution," *Journal of Statistical Computation and Simulation*, vol. 79, no. 11, pp. 1301–1315, 2019.
- [35] M.-H., Chen and Q.-M. Shao, "Monte carlo estimation of Bayesian credible and HPD intervals," *Journal of Computational and Graphical Statistics*, vol. 8, no. 1, pp. 69–92, 1999.
- [36] J. F. Lawless, *Statistical Models and Methods for Lifetime Data*, 2nd ed., Hoboken, New Jersey: John Wiley & Sons, Inc., 2003.