**Article**

# Bayesian Estimations of Shannon Entropy and Rényi Entropy of Inverse Weibull Distribution

Haiping Ren [*] and Xue Hu

*Article*

# Bayesian Estimations of Shannon Entropy and Rényi Entropy of Inverse Weibull Distribution

**Haiping Ren [1,*] and Xue Hu [2]**

[1]  Department of Basic Subjects, Jiangxi University of Science and Technology, Nanchang 330013, China; 9520060004@ jxust.edu.cn

[2]  College of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China; 6120210110@mail.jxust.edu.cn

*  Correspondence: 9520060004@ jxust.edu.cn; Tel.:+86-15979014223

**Abstract:** In this paper, under the symmetric entropy and the scale squared error loss functions, we consider the maximum likelihood (ML) estimation and Bayesian estimation of the Shannon entropy and Rényi entropy of the two-parameter inverse Weibull distribution. In the ML estimation, the dichotomy is used to solve likelihood equation. In addition, the approximation confidence interval is given by the Delta method. Because the form of estimation results is more complex in the Bayesian estimation, the Lindley approximation method is used to achieve the numerical calculation. Finally, Monte Carlo simulations and a real data set are used to illustrate the results derived. By comparing the mean square error between the estimated value and the real value, it can be found that the performance of ML estimation of Shannon entropy is better than that of Bayesian estimation, and there is no significant difference between the performance of ML estimation of Rényi entropy and that of Bayesian estimation

## 1. Introduction

Information is an abstract concept. In the face of a large amount of data, it is easy to know how much data there is, but it is not clear how much information this data contain. Entropy is one of the important terms in physics. Shannon [1] introduced the concept of entropy into statistics, which represents the uncertainty of events. This entropy is generally called "Shannon entropy". Generally speaking, when we hear a message within expectation, we think it contains less information. When we hear an unexpected message, we think that the amount of information it conveys to us is huge. In statistics, the probability is usually used to describe the uncertainty of an event. Therefore, Shannon believes that probability can be used to describe the amount of information contained in an event. After that, Alred Rényi [2] generalized Shannon entropy and put forward the concept of Rényi entropy. Since then, the study of entropy has attracted a lot of attention [3,4]. For example, Chacko and Asha [5] considered the maximum likelihood (ML) estimation and Bayesian estimation of Shannon entropy for generalized exponential distribution by the importance sampling method based on record values. Liu and Gui [6] considered the ML estimation and Bayesian estimation of Shannon entropy for two-parameter Lomax distribution by the Lindley method and the Tierney-Kadane method under a generalized progressively hybrid censoring test. Shrahili et al. [7] considered the estimation of entropy of Log-Logistic distribution. The estimations of different entropy functions are obtained by the ML method, and the approximate confidence intervals are obtained by using various censoring methods and sample sizes. Mahmoud et al. [8] considered the estimation of entropy and residual entropy of two-parameter Lomax distribution based on the generalized type-II hybrid censoring scheme. The ML estimators and Bayesian estimators of entropy and residual entropy are obtained. The simulation study of estimating performance under

different sample sizes is described. Finally, the conclusion is discussed. Hassan and Mazen [9] estimated three entropy measures for the inverse Weibull distribution using progressively Type-II censored data, which are Shannon entropy, Rényi entropy, and q-entropy. The method of maximum likelihood and maximum product of spacing are used to estimate them. Mavis et al. [10] proposed and studied gamma-inverse Weibull distribution, and some mathematical properties were given including moments, mean deviations, Bonferroni and Lorenz curves, and entropies. Basheer [11] introduced a new generalized alpha power inverse Weibull distribution, and the Shannon entropy and Rényi entropy were obtained. Valeriia and Broderick [12] proposed the weighted inverse Weibull class of distributions and derived the expressions of Shannon entropy and Rényi entropy.

In 1982, Keller and Kamath [13] introduced the Inverse Weibull Distribution (IWD) to model the degradation of mechanical components of diesel engines. It is a useful lifetime probability distribution, and it can be used to represent various failure characteristics. Depending on the value of the shape parameter of the IWD, the risk function can be changed flexibly. The use of IWD for data fitting is therefore more appropriate in many cases. For example, Abhijit and Anindya [14] found that the use of IWD was superior to previous normal models when measuring concrete structures using ultrasonic pulse velocities. Chiodo et al. [15] proposed a new model generated from an appropriate mixture of IWD for modeling extreme wind speed scenarios. Langlands et al. [16] observed that breast cancer mortality data could be analyzed using IWD for modeling analysis. That is why two-parameter IWD has attracted more and more researchers to pay attention to and discuss in recent years [17,18]. For example, Asuman and Mahmut [19] considered the classical and Bayesian estimation of parameters and the reliability function of the IWD. In classical estimation, they derived the ML estimators and modified ML estimators. In Bayesian estimation, they utilized the Lindley method to calculate the Bayesian estimators of parameters under symmetric and asymmetric loss functions. Sultan et al. [20] discussed the estimation of parameters of IWD based on the progressive type-II censored sample. They put forward an approximate maximum likelihood method to obtain the ML estimator and used Lindley's approximation to obtain the Bayesian estimators. Amirzadi et al. [21] considered the Bayesian estimation of scale parameter and reliability in the inverse generalized Weibull distribution. In addition to general entropy, squared log error, and weight squared error function. They introduced a new loss function to carry out Bayesian estimation. Peng and Yan [22] studied the Bayesian estimation and prediction for shape and scale parameters of the IWD under a general progressive censoring test. Sindhu et al. [23] assumed different priors and loss functions, and discussed the Bayesian estimation of inverse Weibull mixture distributions based on doubly censored data. Mohammad and Sana [24] obtained the Bayes estimators and ML estimators for the unknown parameters of IWD under lower record values. Faud [25] developed a linear exponential loss function, and estimated parameter and reliability of IWD based on lower record values under this loss function. Li and Hao [26] considered the estimation of a stress-strength model when stress and strength are two independent IWDs with different parameters. Ismail and Tamimi [27] proposed a constant stress partially accelerated life test model and analyzed it using type-I censored data from IWD. Kang and Han [28] derived the approximate maximum likelihood estimators of parameters of IWD under multiply type-II censoring and also proposed a simple graphical method for a goodness-on-fit test. Saboori et al. [29] introduced generalized modified inverse Weibull distribution, and some statistical and probabilistic properties were derived.

This paper will consider the Bayesian estimation of Shannon entropy and Rényi entropy of two-parameter IWD based on complete samples. In Section 2, some related knowledge is introduced first, and then the specific expressions of Shannon entropy and Rényi entropy of two-parameter IWD are derived. In Section 3, the maximum likelihood estimators of the scale parameter and shape parameter of IWD are derived by the dichotomy method, and then the ML estimators of Shannon entropy and Rényi entropy are obtained. In Section 4, the gamma distribution is adopted as the prior distribution (PD) of the scale parameter. A non-informative PD is adopted as the PD of the shape parameter. And then the Bayesian estimators of Shannon entropy and Rényi entropy are obtained

based on the symmetric entropy loss function and scale squared error loss function. Lindley approximation is used to achieve the numerical calculation of the Bayesian estimators of entropy, on account of the complexity of these Bayesian estimators. In Section 5, Monte Carlo simulations are utilized to simulate and compare the estimators that are mentioned above. In Section 6, a real data set has been analyzed for illustrative purposes. Finally, the conclusions of the article are given in Section 7.

## 2. Preliminary Knowledge

The probability density function (pdf) of two-parameter IWD is defined as Eq. (1).

$$f(t;\omega,\upsilon) = \omega \upsilon t^{-\upsilon-1} \exp(-\omega t^{-\upsilon}), \omega > 0, \upsilon > 0, t > 0 , \tag{1}$$

and the cumulative distribution function (cdf) of two-parameter IWD is defined as Eq. (2)

$$F(t;\omega,\upsilon) = \exp(-\omega t^{-\upsilon}), \omega > 0, \upsilon > 0, t > 0 , \tag{2}$$

where the scale parameter is $\omega$, and the shape parameter is $\upsilon$.

Figure 1 shows the pdf of IWD under different values of shape and scale parameters, respectively.
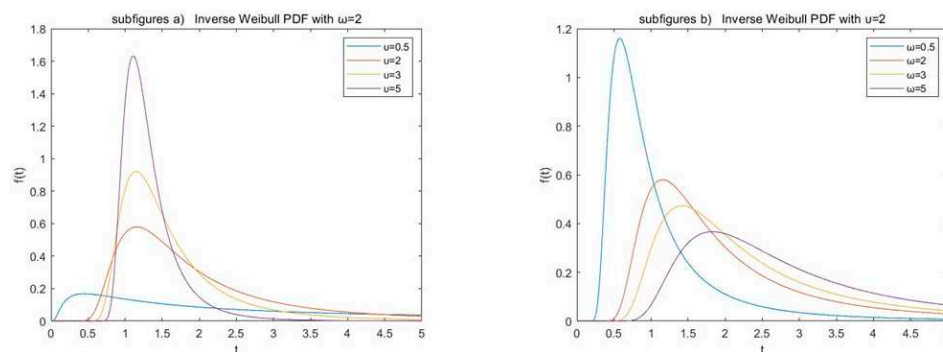


**Figure 1.** The curves of the pdf of IWD with respect to different values of parameters.

The Shannon entropy is defined in Equation (Eq. 3) [1]

$$H_s(t) = -\int_{-\infty}^{+\infty} f(t) \ln[f(t)] dt , \tag{3}$$

and the Rényi entropy is defined in Equation (Eq. 4) [2]

$$H_r(t) = \frac{1}{1-r} \ln \int_{-\infty}^{+\infty} f^r(t) dt, r > 0, r \neq 1, \tag{4}$$

where $f(t)$ is the pdf of a continuous random variable $T$.

**Theorem 1.** *Let $T_1, T_2, ..., T_n$ is a random sample that follows IWD with the pdf (1), $t_1, t_2, ..., t_n$ are the sample observations of $T_1, T_2, ..., T_n$.*

(i) Shannon entropy of IWD is showed in Eq. (5)

$$H_s = \frac{\upsilon+1}{\upsilon} (\ln \omega + \gamma) - \ln \omega \upsilon + 1 . \tag{5}$$

(ii) Rényi entropy of IWD is showed in Eq. (6)

$$H_r = \frac{1}{1-r} [-\frac{r\upsilon+r}{\upsilon} \ln r + \frac{r}{\upsilon} \ln \omega + r \ln \upsilon + \ln \Gamma(r + \frac{r}{\upsilon} + 1)], \tag{6}$$

where $\Gamma(\cdot)$ the gamma function and $\gamma$ is Euler constant.

**Proof.** The log density of pdf (1) of IWD is showed in Eq. (7)

$$\ln[f(t)] = \ln \omega v - (v+1)\ln t - \omega t^{-v}. \tag{7}$$

According to the log-density function (7), and Eq. (3), the Shannon entropy of IWD can be derived as follows:

$$H_s = -\int_0^{+\infty} f(t)[\ln \omega v - (v+1)\ln t - \omega t^{-v}]dt$$

$$= -(\ln \omega v)\int_0^{+\infty} f(t)dt + (v+1)\int_0^{+\infty} f(t)\ln t dt + \omega \int_0^{+\infty} t^{-v} f(t)dt.$$

$$= -\ln \omega v + (v+1)E(\ln T) + \omega E(T^{-v})$$

Obviously,

$$E(T^c) = \int_0^{+\infty} \omega v t^c t^{-v-1} e^{-\omega t^{-v}} dt$$

$$= \omega^{\frac{t}{v}} \int_0^{+\infty} (\omega t^{-v})^{-\frac{c}{v}} e^{-\omega t^{-v}} d(\omega t^{-v}).$$

$$= \omega^{\frac{c}{v}} \Gamma(1-\frac{c}{v})$$

Let $c = -v$,

$$E(T^{-v}) = \omega^{-1}\Gamma(2) = \frac{1}{\omega}.$$

Because

$$E(T^c \ln T) = \frac{dE(T^c)}{dc}$$

$$= \frac{1}{v} \omega^{\frac{c}{v}} \Gamma(1-\frac{c}{v})\ln \omega - \frac{1}{v} \omega^{\frac{c}{v}} \Gamma'(1-\frac{c}{v})$$

$$= \frac{1}{v} \omega^{\frac{c}{v}} [\Gamma(1-\frac{c}{v})\ln \omega - \Gamma'(1-\frac{c}{v})].$$

Let $c = 0$,

$$E(\ln T) = \frac{1}{v}(\ln \omega + \gamma).$$

Therefore, Shannon entropy of two-parameter IWD can be expressed as

$$H_s = -\ln \omega v + (v+1)E(\ln T) + \omega E(T^{-v})$$

$$= \frac{v+1}{v}(\ln \omega + \gamma) - \ln \omega v + 1 \qquad .$$

Obviously,

$$\int_{-\infty}^{+\infty} f^r(t)dt = \int_0^{+\infty} (\omega v t^{-v-1} e^{-\omega t^{-v}})^r dt$$

$$= (\omega v)^r \int_0^{+\infty} t^{-rv-r} e^{-r\omega t^{-v}} dt$$

$$= r^{\frac{-rv-r}{v}} \omega^{\frac{r}{v}} v^r \Gamma(1+\frac{r+rv}{v})$$

Then, according to Eq. (4), the Rényi entropy of two-parameter IWD can be expressed as

$$H_r = \frac{1}{1-r}[\frac{-rv-r}{v}\ln r + \frac{r}{v}\ln \omega + r\ln v + \ln \Gamma(1+\frac{rv+r}{v})].$$

## 3. Maximum Likelihood Estimation

Suppose that $T_1, T_2, ..., T_n$ is a random sample that follows IWD with the pdf (1), $t_1, t_2, ..., t_n$ are the sample observations of $T_1, T_2, ..., T_n$. Thus, the likelihood function (LF) can be derived as Eq. (8)

$$s(t; \omega, v) = \omega^n v^n (\prod_{i=1}^{n} t_i^{-v-1}) \exp(-\omega \sum_{i=1}^{n} t_i^{-v}) . \tag{8}$$

Then, the corresponding log LF of Eq. (8) is showed in Eq. (9)

$$S(t; \omega, v) = \ln s(t; \omega, v) = n \ln \omega v - (v+1) \sum_{i=1}^{n} \ln t_i - \omega \sum_{i=1}^{n} t_i^{-v} . \tag{9}$$

For convenience, we denote $S(t; \omega, v)$ as $S$. Thus, the likelihood equations can be expressed respectively as Eq. (10) and Eq. (11)

$$\frac{\partial S}{\partial \omega} = \frac{n}{\omega} - \sum_{i=1}^{n} t_i^{-v} = 0 , \tag{10}$$

$$\frac{\partial S}{\partial v} = \frac{n}{v} - \sum_{i=1}^{n} \ln t_i + \omega \sum_{i=1}^{n} t_i^{-v} \ln t_i = 0 . \tag{11}$$

The ML estimators $\hat{\omega}$ and $\hat{v}$ can be obtained by solving Eq. (10) and Eq. (11) with dichotomy, whose calculation steps are listed as follows:

(i) According to Eq. (10) and Eq. (11), there are

$$\hat{\omega} = n(\sum_{i=1}^{n} t_i^{-\hat{v}})^{-1} , $$

$$\tag{12}$$

$$\frac{n}{v} - \sum_{i=1}^{n} \ln t_i + n(\sum_{i=1}^{n} t_i^{-v})^{-1} \sum_{i=1}^{n} t_i^{-v} \ln t_i = 0 . \tag{13}$$

(ii) Denote $y(v) = \frac{n}{v} - \sum_{i=1}^{n} \ln t_i + n(\sum_{i=1}^{n} t_i^{-v})^{-1} \sum_{i=1}^{n} t_i^{-v} \ln t_i$ , given the accuracy $\varepsilon$ , determine the interval $[v_u, v_l]$ and verify $y(v_u) \cdot y(v_l) < 0$ .

(iii) Find the midpoint $v_m$ of the interval $[v_u, v_l]$ and calculate $y(v_m)$.

(iv) If $y(v_m) = 0$, $\hat{v} = v_m$.

(v) If $y(v_u) \cdot y(v_m) < 0$, $v_l = v_m$; If $y(v_l) \cdot y(v_m) < 0$, $v_u = v_m$.

(vi) If $|v_u - v_l| < \varepsilon$, $\hat{v}$ is equal to $v_u$ or $v_l$. If not, return to step (iii) to step (vi).

Due to the invariance of ML estimation, the ML estimators of Shannon entropy and Rényi entropy can be obtained by putting $\hat{\omega}$ and $\hat{v}$ into Eq. (3) and Eq. (4), and their mathematical expressions are showed in Eq. (14) and Eq. (15)

$$\hat{H}_{s1} = \frac{\hat{v}+1}{\hat{v}} (\ln \hat{\omega} + \gamma) - \ln \hat{\omega} \hat{v} + 1 , \tag{14}$$

$$\hat{H}_{r1} = \frac{1}{1-r} [\frac{-r\hat{v}-r}{\hat{v}} \ln r + \frac{r}{\hat{v}} \ln \hat{\omega} + r \ln \hat{v} + \ln \Gamma(1 + \frac{r\hat{v}+r}{\hat{v}})] . \tag{15}$$

*Next, the Delta method is used to derive the approximate confidence intervals (briefly, ACIs) of Shannon entropy and Rényi entropy.*

*Denote vector $D_s$ and $D_r$ as*

$$D_s = (\frac{\partial H_s}{\partial \omega}, \frac{\partial H_s}{\partial v})\Big|_{\omega = \hat{\omega}, v = \hat{v}} , \tag{16}$$

$$D_r = (\frac{\partial H_r}{\partial \omega}, \frac{\partial H_r}{\partial v})\Big|_{\omega = \hat{\omega}, v = \hat{v}} , \tag{17}$$

*in which $D_s$ and $D_r$ are calculated through Eq. (18) and Eq. (19).*

$$\frac{\partial H_s}{\partial \omega} = \frac{1}{\omega v}, \quad \frac{\partial H_s}{\partial v} = \frac{-\gamma - \ln \omega}{v^2} - \frac{1}{v} \tag{18}$$

$$\frac{\partial H_r}{\partial \omega} = \frac{r}{(1-r)\omega v}, \quad \frac{\partial H_r}{\partial v} = \frac{r}{(1-r)v}[1 + \frac{\ln r - \ln \omega}{v} - \frac{\Gamma'(1 + r + rv^{-1})}{v\Gamma(1 + r + rv^{-1})}] \tag{19}$$

*According to the Delta method, calculate the estimated variance of $\hat{H}_{s1}$ and $\hat{H}_{r1}$ as Eq. (20) and Eq. (21). $I$ is the Fisher information matrix of $\omega$ and $v$, and Eq. (22) gives the elements of $I$. $I^{-1}$ is the inverse matrix of $I$.*

$$V_s = D_s I^{-1} D_s^{\mathrm{T}}\big|_{\omega = \hat{\omega}, v = \hat{v}} \tag{20}$$

$$V_r = D_r I^{-1} D_r^{\mathrm{T}}\big|_{\omega = \hat{\omega}, v = \hat{v}} \tag{21}$$

$$\frac{\partial^2 S}{\partial \omega^2} = -\frac{n}{\omega^2}, \quad \frac{\partial^2 S}{\partial \omega \partial v} = \frac{\partial^2 S}{\partial v \partial \omega} = \sum_{i=1}^{n} t_i^{-v} \ln t_i$$

$$\frac{\partial^2 S}{\partial v^2} = -\frac{n}{v^2} - \omega \sum_{i=1}^{n} t_i^{-v} (\ln t_i)^2 \tag{22}$$

*Then, the $100(1-\alpha)\%$ ACI of Shannon entropy is Eq. (23), and the $100(1-\alpha)\%$ ACI of Rényi entropy is Eq. (24), where $z_{\alpha/2}$ is the upper ($\alpha/2$)th quantile of the standardized normal distribution.*

$$(\hat{H}_{s1} - z_{\alpha/2}\sqrt{V_s}, \hat{H}_{s1} + z_{\alpha/2}\sqrt{V_s}) \tag{23}$$

$$(\hat{H}_{r1} - z_{\alpha/2}\sqrt{V_r}, \hat{H}_{r1} + z_{\alpha/2}\sqrt{V_r}) \tag{24}$$

## 4. Bayesian Estimation

Bayesian estimation is a method of introducing prior information to deal with decision problems. The advantage is that it can include the prior information in statistical inference and improve the accuracy of the taken decision. From the time when Bayesian estimation was proposed to now, many researchers have adopted this method in estimating parameters and related functions. For example, Kundu and Howlader [30] considered the Bayesian inference and prediction of inverse Weibull distribution, based on type-II censored data. A Gibbs sampling procedure was used for MCMC samples, and Bayes estimation was computed by this sample. Sultan et al. [31] considered the Bayesian estimation of inverse Weibull parameters based on progressive type-II censored data. Because the Bayes estimators can't be obtained explicitly, the Lindley approximation was used to calculate them. Mohammad and Mina [32] presented the Bayesian inferences of parameters of inverse Weibull distribution based on type-I hybrid censored data and computed the Bayes estimates using Lindley approximation. Algarni et al. [33] considered the Bayes estimation of parameters for the inverse Weibull distribution employing a progressive type-I censored sample. Metropolis-Hasting (MH) algorithm was used to compute the Bayesian estimates.

In addition to the areas mentioned above, there are some recent applications of the Bayesian method. Zhou and Luo [34] developed a supplier's recursive multiperiod discounted profit model based on Bayesian information updating. Yulin et al. [35] put forward a Bayesian approach to tackle the misalignments for over-the-air computation. Taborsky et al. [36] presented a novel generic Bayesian probabilistic model to solve the problem of parameters marginalization under the constraint of forced community structure. Oliver [37] introduced the Bayesian toolkit and showed how geomorphic models might benefit from probabilistic concepts. Ran et al. [38] proposed a Bayesian approach to measure the loss of privacy in a mechanism. Luo et al. [39] used the Bayesian information criterion for model selection when revisiting the lifetime data of brake pads. Peng et al. [40] extended a general Bayesian framework to deal with the degradation analysis of sparse degradation observations and evolving observations. František et al. [41] illustrated Bayesian

estimation of how to benefit parametric survival analysis. Liu et al. [42] proposed fuzzy Bayesian knowledge tracing models to address continuous score scenarios.

In this paper, the Bayesian estimations of Shannon entropy and Rényi entropy of IWD are investigated under symmetric entropy (SE) and scale squared error (SSE) loss functions, which are widely used in Bayesian statistical inference [43–45].

(i) SE loss function is defined in Equation (Eq. 25) [43]

$$L_1(H,\hat{H}) = \frac{H}{\hat{H}} + \frac{\hat{H}}{H} - 2,\tag{25}$$

where $\hat{H}$ is the estimator of $H$.

**Lemma 1.** Suppose that $T$ is the historical data information about the entropy function $H$. Then, under the SE loss function (25), the Bayesian estimator $\hat{H}_1$ for any prior distribution is showed in Eq. (26)

$$\hat{H}_1 = [\frac{E(H|T)}{E(H^{-1}|T)}]^{\frac{1}{2}},\tag{26}$$

where $E(H|T)$ is the posterior expectation of $H$ and $E(H^{-1}|T)$ is the posterior expectation of $H^{-1}$.

**Proof.** Under the SE loss function (25), the Bayesian risk of $\hat{H}$ is

$$R(\hat{H}) = E_H(E(L_1(H,\hat{H})|T)).$$

To minimize $R(\hat{H})$, only need to minimize $E(L_1(H,\hat{H})|T)$. For convenience, let $g(\hat{H}) = E(L_1(H,\hat{H})|T)$.

Because

$$g(\hat{H}) = \hat{H}^{-1}E(H|T) + \hat{H}E(H^{-1}|T) - 2,$$

and the derivative is

$$g'(\hat{H}) = -\hat{H}^{-2}E(H|T) + E(H^{-1}|T).$$

The Bayesian estimator $\hat{H}_1$ can be obtained by $g'(\hat{H}) = 0$.

(ii) SSE loss function is defined in Equation (Eq. 27) [45]

$$L_2(H,\hat{H}) = \frac{(H-\hat{H})^2}{H^k},\tag{27}$$

where $k$ is a nonnegative integer.

**Lemma 2.** *Suppose that $T$ is the historical data information about the entropy function $H$. Then, under the SSE loss function (27), the Bayesian estimator $\hat{H}_2$ for any prior distribution is*

$$\hat{H}_2 = \frac{E(H^{1-k}|T)}{E(H^{-k}|T)},\tag{28}$$

where $E(H^{1-k}|T)$ is the posterior expectation of $H^{1-k}$ and $E(H^{-k}|T)$ is the posterior expectation of $H^{-k}$.

**Proof.** Under the SSE loss function (27), the Bayesian risk of $\hat{H}$ is

$$R(\hat{H}) = E_H(E(L_2(H,\hat{H})|T)).$$

To minimize $R(\hat{H})$, only need to minimize $E(L_2(H,\hat{H})|T)$. Similarly, let $h(\hat{H}) = E(L_2(H,\hat{H})|T)$.

Because

$$h(\hat{H}) = E(\frac{H^2 - 2H\hat{H} + \hat{H}^2}{H^k}|T)$$

$$= E(H^{2-k}|T) - 2\hat{H}E(H^{1-k}|T) + \hat{H}^2 E(H^{-k}|T)$$

and the derivative of $h(\hat{H})$ is

$$h'(\hat{H}) = -2E(H^{1-k}|T) + 2\hat{H}E(H^{-k}|T).$$

The Bayes estimator $\hat{H}_2$ can be obtained by $h'(\hat{H}) = 0$.

Assume that the scale parameter $\omega$ and shape parameter $\upsilon$ of two-parameter IWD are independent random variables, which $\omega$ obey $\Gamma(a,b)$ and $\upsilon$ obey non-informative PD as follows:

$$P_1(\omega) = \frac{a^b}{\Gamma(b)}\omega^{b-1}e^{-a\omega}, a > 0, b > 0, \tag{29}$$

$$P_2(\upsilon) \propto \frac{1}{\upsilon}. \tag{30}$$

Thus, the joint PD of $\omega$ and $\upsilon$ is

$$P(\omega,\upsilon) \propto \frac{a^b}{\upsilon\Gamma(b)}\omega^{b-1}e^{-a\omega}. \tag{31}$$

Referring to Bayesian formulation, the posterior distribution of $\omega$ and $\upsilon$ is

$$P(\omega,\upsilon|T) = \frac{P(\omega,\upsilon)s(t;\omega,\upsilon)}{\int_0^{+\infty}\int_0^{+\infty}P(\omega,\upsilon)s(t;\omega,\upsilon)d\omega d\upsilon}. \tag{32}$$

Thus, the Bayesian estimators of Shannon entropy and Rényi entropy under SE can be expressed as

$$\hat{H}_{s2} = [\frac{E(H_s|T)}{E(H_s^{-1}|T)}]^{\frac{1}{2}} = [\frac{\int_0^{+\infty}\int_0^{+\infty}H_s P(\omega,\upsilon|T)d\omega d\upsilon}{\int_0^{+\infty}\int_0^{+\infty}H_s^{-1}P(\omega,\upsilon|T)d\omega d\upsilon}]^{\frac{1}{2}}, \tag{33}$$

$$\hat{H}_{r2} = [\frac{E(H_r|T)}{E(H_r^{-1}|T)}]^{\frac{1}{2}} = [\frac{\int_0^{+\infty}\int_0^{+\infty}H_r P(\omega,\upsilon|T)d\omega d\upsilon}{\int_0^{+\infty}\int_0^{+\infty}H_r^{-1}P(\omega,\upsilon|T)d\omega d\upsilon}]^{\frac{1}{2}}. \tag{34}$$

The Bayesian estimators of Shannon entropy and Rényi entropy under SSE can be expressed as

$$\hat{H}_{s3} = \frac{E(H_s^{1-k}|T)}{E(H_s^{-k}|T)} = \frac{\int_0^{+\infty}\int_0^{+\infty}H_s^{1-k}P(\omega,\upsilon|T)d\omega d\upsilon}{\int_0^{+\infty}\int_0^{+\infty}H_s^{-k}P(\omega,\upsilon|T)d\omega d\upsilon}, \tag{35}$$

$$\hat{H}_{r3} = \frac{E(H_r^{1-k}|T)}{E(H_r^{-k}|T)} = \frac{\int_0^{+\infty}\int_0^{+\infty}H_r^{1-k}P(\omega,\upsilon|T)d\omega d\upsilon}{\int_0^{+\infty}\int_0^{+\infty}H_r^{-k}P(\omega,\upsilon|T)d\omega d\upsilon}. \tag{36}$$

From Eq. (33) to Eq. (36), it can be seen that the calculation of Bayesian estimators of Shannon and Rényi entropy are complex and difficult to calculate. Thus, Lindley approximation will be employed to achieve the approximate calculation results of $\hat{H}_{s3}$ and $\hat{H}_{r3}$.

*4.1. Bayesian Estimation by using Lindley approximation under SE loss function*

Referring to Lindley approximation, $I(t)$ can be defined as

$$I(t) = E[U(\omega,\upsilon)|T] = \frac{\int U(\omega,\upsilon)e^{S(t;\omega,\upsilon)+G(\omega,\upsilon)}d(\omega,\upsilon)}{\int e^{S(t;\omega,\upsilon)+G(\omega,\upsilon)}d(\omega,\upsilon)}, \tag{37}$$

where $U(\omega,\upsilon)$ is a function of independent variables $\omega$ and $\upsilon$, $S(t;\omega,\upsilon)$ is log LF defined in Eq. (9), $G(\omega,\upsilon)$ is the log of joint PD defined in Eq. (31).

If the sample size is large, Eq. (37) can be expressed as

$$I(t)=U(\hat{\omega},\hat{\upsilon})+\frac{1}{2}(A+B+C+D),\qquad(38)$$

where $\hat{\omega}$ and $\hat{\upsilon}$ are the ML estimators of $\omega$ and $\upsilon$, and

$$
\begin{aligned}
A &= (\hat{U}_{\omega\omega}+2\hat{U}_{\omega}\hat{G}_{\omega})\hat{\sigma}_{\omega\omega}+(\hat{U}_{\upsilon\omega}+2\hat{U}_{\upsilon}\hat{G}_{\omega})\hat{\sigma}_{\upsilon\omega}\\
B &= (\hat{U}_{\omega\upsilon}+2\hat{U}_{\omega}\hat{G}_{\upsilon})\hat{\sigma}_{\omega\upsilon}+(\hat{U}_{\upsilon\upsilon}+2\hat{U}_{\upsilon}\hat{G}_{\upsilon})\hat{\sigma}_{\upsilon\upsilon}\\
C &= (\hat{U}_{\omega}\hat{\sigma}_{\omega\omega}+\hat{U}_{\upsilon}\hat{\sigma}_{\omega\upsilon})(\hat{S}_{\omega\omega\omega}\hat{\sigma}_{\omega\omega}+\hat{S}_{\omega\upsilon\omega}\hat{\sigma}_{\omega\upsilon}+\hat{S}_{\upsilon\omega\omega}\hat{\sigma}_{\upsilon\omega}+\hat{S}_{\upsilon\upsilon\omega}\hat{\sigma}_{\upsilon\upsilon})\\
D &= (\hat{U}_{\omega}\hat{\sigma}_{\upsilon\omega}+\hat{U}_{\upsilon}\hat{\sigma}_{\upsilon\upsilon})(\hat{S}_{\omega\omega\upsilon}\hat{\sigma}_{\omega\omega}+\hat{S}_{\omega\upsilon\upsilon}\hat{\sigma}_{\omega\upsilon}+\hat{S}_{\upsilon\omega\upsilon}\hat{\sigma}_{\upsilon\omega}+\hat{S}_{\upsilon\upsilon\upsilon}\hat{\sigma}_{\upsilon\upsilon})
\end{aligned}\qquad(39)
$$

$\sigma_{ij}$ ( $i,j=\omega,\upsilon$ ) is the element of inverse matrix of $-S_{ij}$ .

The $\hat{U}_{\omega\omega}$ is denoted that taking the second derivative of $U(\omega,\upsilon)$ with respect to $\omega$ and putting $\hat{\omega}$ into it. Similarly, others can be expressed as

$$S_{\omega\omega\upsilon}=S_{\omega\upsilon\omega}=S_{\upsilon\omega\omega}=0,\ \ S_{\omega\omega\omega}=\frac{3n}{\omega^3}$$

$$S_{\omega\upsilon\upsilon}=S_{\upsilon\omega\upsilon}=S_{\upsilon\upsilon\omega}=-\sum_{i=1}^{n}t_i^{-\upsilon}\ln t_i,\ \ S_{\upsilon\upsilon\upsilon}=\frac{2n}{\upsilon^3}+\omega\sum_{i=1}^{n}t_i^{-\upsilon}(\ln t_i)^3.\qquad(40)$$

$$G_{\omega}=\frac{b-1}{\omega}-a,\ \ G_{\upsilon}=\frac{1}{\upsilon}$$

Under the SE loss function, the step of numerical calculation of Shannon entropy $\hat{H}_{s2}$ by Lindley approximation is shown as follows:

When $U(\omega,\upsilon)=H_s$ ,

$$U_{\omega}=\frac{1}{\omega\upsilon},\ \ U_{\upsilon}=\frac{-\gamma-\ln\omega}{\upsilon^2}-\frac{1}{\upsilon}$$

$$U_{\omega\omega}=-\frac{1}{\omega^2\upsilon},\ \ U_{\upsilon\upsilon}=\frac{2\ln\omega+2\gamma}{\upsilon^3}+\frac{1}{\upsilon^2},\ \ U_{\omega\upsilon}=U_{\upsilon\omega}=-\frac{1}{\omega\upsilon^2}\qquad(41)$$

Put Eq. (40) and Eq. (41) into Eq. (38), $E(H_s\,|\,T)$ is obtained.

Similarly, when $U(\omega,\upsilon)=H_s^{-1}$ ,

$$U_{\omega}=-H_s^{-2}\frac{1}{\omega\upsilon},\ \ U_{\upsilon}=H_s^{-2}(\frac{\gamma+\ln\omega}{\upsilon^2}+\frac{1}{\upsilon})$$

$$U_{\omega\omega}=2H_s^{-3}\frac{1}{\omega^2\upsilon^2}+H_s^{-2}\frac{1}{\omega^2\upsilon}$$

$$U_{\upsilon\upsilon}=2H_s^{-3}(\frac{-\gamma-\ln\omega}{\upsilon^2}-\frac{1}{\upsilon})^2-H_s^{-2}(\frac{2\ln\omega+2\gamma}{\upsilon^3}+\frac{1}{\upsilon^2})\qquad(42)$$

$$U_{\omega\upsilon}=U_{\upsilon\omega}=2H_s^{-3}(\frac{-\gamma-\ln\omega}{\omega\upsilon^3}-\frac{1}{\omega\upsilon^2})+H_s^{-2}\frac{1}{\omega\upsilon^2}$$

Then, put Eq. (40) and Eq. (42) into Eq. (38), $E(H_s^{-1}|T)$ is obtained. Thus, the numerical calculation of Shannon entropy $\hat{H}_{s2}$ is calculated by Eq. (33).

Under the SE loss function, the numerical calculation of Rényi entropy $\hat{H}_{r2}$ by Lindley approximation is shown as follows:

When $U(\omega,\upsilon)=H_r$ ,

$$U_\omega = \frac{r}{(1-r)\omega\upsilon}, \quad U_{\omega\omega} = \frac{-r}{(1-r)\omega^2\upsilon}, \quad U_{\omega\upsilon} = U_{\upsilon\omega} = \frac{-r}{(1-r)\omega\upsilon^2}$$

$$U_\upsilon = \frac{r}{(1-r)\upsilon}[1 + \frac{\ln r - \ln\omega}{\upsilon} - \frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}] \tag{43}$$

$$U_{\upsilon\upsilon} = \frac{2r(\ln\omega - \ln r)}{(1-r)\upsilon^3} - \frac{[\Gamma'(1+r+r\upsilon^{-1})]^2 - \Gamma''(1+r+r\upsilon^{-1})\Gamma(1+r+r\upsilon^{-1})}{(1-r)r^{-2}\upsilon^4[\Gamma(1+r+r\upsilon^{-1})]^2} - \frac{r}{(1-r)\upsilon^2}$$

*Put Eq. (40) and Eq. (43) into Eq. (38), $E(H_r \,|\, T)$ is obtained.*

When $U(\omega,\upsilon) = H_r^{-1}$,

$$U_\omega = \frac{-rH_r^{-2}}{(1-r)\omega\upsilon}, \quad U_\upsilon = \frac{-rH_r^{-2}}{(1-r)\upsilon}[1 + \frac{\ln r - \ln\omega}{\upsilon} - \frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]$$

$$U_{\omega\omega} = \frac{2r^2H_r^{-3}}{(1-r)^2\omega^2\upsilon^2} + \frac{rH_r^{-2}}{(1-r)\omega^2\upsilon},$$

$$U_{\omega\upsilon} = U_{\upsilon\omega} = \frac{r^2 2H_r^{-3}}{(1-r)^2\omega\upsilon^2}[1 + \frac{\ln r - \ln\omega}{\upsilon} - \frac{1}{\upsilon}\frac{\Gamma'(1+r+r\upsilon^{-1})}{\Gamma(1+r+r\upsilon^{-1})}] \tag{44}$$

$$U_{\upsilon\upsilon} = \frac{[\Gamma'(1+r+r\upsilon^{-1})]^2 - \Gamma''(1+r+r\upsilon^{-1})\Gamma(1+r+r\upsilon^{-1})}{(1-r)r^{-2}\upsilon^4[\Gamma(1+r+r\upsilon^{-1})]^2 H_r^2} + \frac{rH_r^{-2}}{(1-r)\upsilon^2} - \frac{2r(\ln\omega - \ln r)}{(1-r)\upsilon^3 H_r^2}$$

$$+ \frac{2r^2 H_r^{-3}}{(1-r)^2\upsilon^2}[1 + \frac{\ln r - \ln\omega}{\upsilon} - \frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]$$

*Put Eq. (40) and Eq. (44) into Eq. (38), $E(H_r^{-1}\,|\,T)$ is obtained. Thus, the numerical calculation of Shannon entropy $\hat{H}_{r2}$ is calculated by Eq. (34).*

### 4.2. Bayesian Estimation by using Lindley approximation under SSE loss function

Under the SSE loss function, the step of numerical calculation of Shannon entropy $\hat{H}_{s3}$ by Lindley approximation is shown as follows:

When $U(\omega,\upsilon) = H_s^{1-k}$,

$$U_\omega = (1-k)H_s^{-k}\frac{1}{\omega\upsilon}, \quad U_\upsilon = (1-k)H_s^{-k}(\frac{-\gamma - \ln\omega}{\upsilon^2} - \frac{1}{\upsilon})$$

$$U_{\omega\omega} = -k(1-k)H_s^{-k-1}(\frac{1}{\omega\upsilon})^2 - (1-k)H_s^{-k}\frac{1}{\omega^2\upsilon}$$

$$U_{\upsilon\upsilon} = k(1-k)H_s^{-k-1}(\frac{\gamma + \ln\omega}{\upsilon^2} + \frac{1}{\upsilon})^2 + (1-k)H_s^{-k}(\frac{2\ln\omega + 2\gamma}{\upsilon^3} + \frac{1}{\upsilon^2}) \tag{45}$$

$$U_{\omega\upsilon} = U_{\upsilon\omega} = k(1-k)H_s^{-k-1}\frac{1}{\omega\upsilon}(\frac{\gamma + \ln\omega}{\upsilon^2} + \frac{1}{\upsilon}) - (1-k)H_s^{-k}\frac{1}{\omega\upsilon^2}$$

Then, putting Eq. (40) and Eq. (45) into Eq. (38), $E(H_s^{1-k}|T)$ is obtained.

When $U(\omega,\upsilon) = H_s^{-k}$,

$$U_\omega = -kH_s^{-k-1}\frac{1}{\omega\upsilon}, \quad U_\upsilon = kH_s^{-k-1}(\frac{\gamma + \ln\omega}{\upsilon^2} + \frac{1}{\upsilon})$$

$$U_{\omega\omega} = k(k+1)H_s^{-k-2}(\frac{1}{\omega\upsilon})^2 + kH_s^{-k-1}\frac{1}{\omega^2\upsilon}$$

$$U_{\upsilon\upsilon} = k(k+1)H_s^{-k-2}(\frac{-\gamma - \ln\omega}{\upsilon^2} - \frac{1}{\upsilon})^2 - kH_s^{-k-1}(\frac{2\ln\omega + 2\gamma}{\upsilon^3} + \frac{1}{\upsilon^2}) \tag{46}$$

$$U_{\omega\upsilon} = U_{\upsilon\omega} = k(k+1)H_s^{-k-2}\frac{1}{\omega\upsilon}(\frac{-\gamma - \ln\omega}{\upsilon^2} - \frac{1}{\upsilon}) + kH_s^{-k-1}\frac{1}{\omega\upsilon^2}$$

Then, putting Eq. (40) and Eq. (46) into Eq. (38), $E(H_s^{-k}|T)$ is obtained. Thus, the numerical calculation of Shannon entropy $\hat{H}_{s3}$ is calculated by Eq. (35).

Under the SSE loss function, the step of numerical calculation of Rényi entropy $\hat{H}_{r3}$ by Lindley approximation is shown as follows:

When $U(\omega,\upsilon)=H_r^{1-k}$,

$$U_\omega=\frac{(1-k)rH_r^{-k}}{(1-r)\omega\upsilon}, \quad U_\upsilon=\frac{(1-k)rH_r^{-k}}{(1-r)\upsilon}[1+\frac{\ln r-\ln\omega}{\upsilon}-\frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]$$

$$U_{\omega\omega}=(k-1)[\frac{kr^2H_r^{-k-1}}{(1-r)^2\omega^2\upsilon^2}+\frac{rH_r^{-k}}{(1-r)\omega^2\upsilon}]$$

$$U_{\omega\upsilon}=U_{\upsilon\omega}=\frac{kr^2H_r^{-k-1}(k-1)}{(1-r)^2\omega\upsilon^2}[1+\frac{\ln r-\ln\omega}{\upsilon}-\frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]+\frac{rH_r^{-k}(k-1)}{(1-r)\omega\upsilon^2} \qquad (47)$$

$$U_{\upsilon\upsilon}=\frac{r^2kH_r^{-k-1}(k-1)}{(1-r)^2\upsilon^2}[1+\frac{\ln r-\ln\omega}{\upsilon}-\frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]^2+\frac{2r(\ln\omega-\ln r)(1-k)}{(1-r)\upsilon^3H_r^k}$$

$$-\frac{[\Gamma'(1+r+r\upsilon^{-1})]^2-\Gamma''(1+r+r\upsilon^{-1})\Gamma(1+r+r\upsilon^{-1})}{(1-r)(1-k)^{-1}r^{-2}\upsilon^4[\Gamma(1+r+r\upsilon^{-1})]^2H_r^k}-\frac{(1-k)H_r^{-k}r}{(1-r)\upsilon^2}$$

*Put Eq. (40) and Eq. (47) into Eq. (38), $E(H_r^{1-k}|T)$ is obtained.*

When $U(\omega,\upsilon)=H_r^{-k}$,

$$U_\omega=-\frac{rkH_r^{-k-1}}{(1-r)\omega\upsilon}, \quad U_\upsilon=\frac{rkH_r^{-k-1}}{(1-r)\upsilon}[\frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}-1-\frac{\ln r-\ln\omega}{\upsilon}]$$

$$U_{\omega\omega}=\frac{kr}{(1-r)\omega^2\upsilon}[\frac{rH_r^{-k-2}(k-1)}{\upsilon}+H_r^{-k-1}]$$

$$U_{\omega\upsilon}=U_{\upsilon\omega}=\frac{r^2k(k-1)H_r^{-k-2}}{(1-r)^2\omega\upsilon^2}[1+\frac{\ln r-\ln\omega}{\upsilon}-\frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]+\frac{rkH_r^{-k-1}}{(1-r)\omega\upsilon^2} \qquad (48)$$

$$U_{\upsilon\upsilon}=\frac{r^2k(k-1)H_r^{-k-2}}{(1-r)^2\upsilon^2}[1+\frac{\ln r-\ln\omega}{\upsilon}-\frac{\Gamma'(1+r+r\upsilon^{-1})}{\upsilon\Gamma(1+r+r\upsilon^{-1})}]^2-\frac{2rk(\ln\omega-\ln r)}{(1-r)\upsilon^3H_r^{k+1}}$$

$$+\frac{[\Gamma'(1+r+r\upsilon^{-1})]^2-\Gamma''(1+r+r\upsilon^{-1})\Gamma(1+r+r\upsilon^{-1})}{(1-r)r^{-2}\upsilon^4k^{-1}H_r^{k+1}[\Gamma(1+r+r\upsilon^{-1})]^2}+\frac{rkH_r^{-k-1}}{(1-r)\upsilon^2}$$

Put Eq. (40) and Eq. (48) into Eq. (38), $E(H_r^{-k}|T)$ is obtained. Thus, the numerical calculation of Rényi entropy $\hat{H}_{r3}$ is calculated by Eq. (36).

## 5. Monte Carlo simulation

In this chapter, Monte Carlo simulation is used to generate random samples that obey two-parameter IWD, and repeat 1000 experiments respectively with different sample sizes ($n=10,20,30,40,50,60,70,80,90,100$). The true values of the parameters in the two-parameter IWD are taken as $\omega=1$ and $\upsilon=2$, the parameters of the gamma distribution are taken as $a=5$ and $b=1$, the parameters in SSE are taken as $k=10$, and the parameters of the Rényi entropy are taken as $r=0.5$. Then, the mean squared error (briefly, MSE) is used to compare the performance of each estimator. The results of Shannon entropy are shown in Table 1, and the results of Rényi entropy are shown in Table 2. For showing the performance of $100(1-\alpha)\%$ ACIs, the coverage probability is calculated and the results are shown in Table 3.

For convenience, $H_{s0}$ and $H_{r0}$ represent the true values of Shannon entropy and Rényi entropy, $\hat{M}_{s1}$ and $\hat{M}_{r1}$ represent mean values of 1000 ML estimates of entropy respectively, $\hat{M}_{s2}$ and $\hat{M}_{r2}$ represent mean values of 1000 Bayesian estimates of entropy respectively under SE loss

function, $\hat{M}_{s3}$ and $\hat{M}_{r3}$ represent mean values of 1000 Bayesian estimates of entropy respectively under SSE loss function, $MSE_{s1}$ and $MSE_{r1}$ represent MSEs of ML estimates of entropy respectively, $MSE_{s2}$ and $MSE_{r2}$ represent MSEs of Bayesian estimates of entropy respectively under SE loss function, $MSE_{s3}$ and $MSE_{r3}$ represent MSEs of Bayesian estimates of entropy respectively under SSE loss function. The $\hat{M}_{sj}$ and $MSE_{sj}$ ( $j = 1, 2, 3$ ) are calculated by Eq. (49) and Eq. (50), where $m = 1000$ and $\hat{H}_{sj,i}$ represents the *i*-th ML estimate or Bayesian estimate of Shannon entropy. The $\hat{M}_{rj}$ and $MSE_{rj}$ ( $j = 1, 2, 3$ ) are calculated by Eq. (51) and Eq. (52), where $m = 1000$ and $\hat{H}_{rj,i}$ represents the *i*-th ML estimate or Bayesian estimate of Rényi entropy.

$$\hat{M}_{sj} = \frac{1}{m}\sum_{i=1}^{m}\hat{H}_{sj,i} \ , \tag{49}$$

$$MSE_{sj} = \frac{1}{m}\sum_{i=1}^{m}(\hat{H}_{sj,i} - H_{s0})^2 \ , \tag{50}$$

$$\hat{M}_{rj} = \frac{1}{m}\sum_{i=1}^{m}\hat{H}_{rj,i} \ , \tag{51}$$

$$MSE_{rj} = \frac{1}{m}\sum_{i=1}^{m}(\hat{H}_{rj,i} - H_{r0})^2 \ . \tag{52}$$

**Table 1.** Estimates and MSEs of Shannon entropy ( $H_{s0}$ =1.1727 ).

| Sample size ( $n$ ) | Estimate | | | MSE | | |
|---|---|---|---|---|---|---|
| | $\hat{M}_{s1}$ | $\hat{M}_{s2}$ | $\hat{M}_{s3}$ | $MSE_{s1}$ | $MSE_{s2}$ | $MSE_{s3}$ |
| 10 | 1.0604 | 0.8259 | 0.8914 | 0.1903 | 0.3666 | 0.2065 |
| 20 | 1.1183 | 0.9631 | 0.9683 | 0.0863 | 0.1282 | 0.1186 |
| 30 | 1.1388 | 1.0301 | 1.0076 | 0.0558 | 0.0751 | 0.0766 |
| 40 | 1.1355 | 1.0526 | 1.0292 | 0.0445 | 0.0574 | 0.0592 |
| 50 | 1.1461 | 1.0788 | 1.0379 | 0.0323 | 0.0404 | 0.0472 |
| 60 | 1.1503 | 1.0938 | 1.0506 | 0.0287 | 0.0343 | 0.0399 |
| 70 | 1.1579 | 1.1093 | 1.0646 | 0.0244 | 0.0282 | 0.0334 |
| 80 | 1.1623 | 1.1196 | 1.0694 | 0.0197 | 0.0224 | 0.0284 |
| 90 | 1.1653 | 1.1272 | 1.0803 | 0.0171 | 0.0191 | 0.0256 |
| 100 | 1.1628 | 1.1284 | 1.0777 | 0.0161 | 0.0183 | 0.0244 |

**Table 2.** Estimates and MSEs of Rényi entropy ( $H_{r0}$ =1.5641 ).

| Sample size ( $n$ ) | Estimate | | | MSE | | |
|---|---|---|---|---|---|---|
| | $\hat{M}_{r1}$ | $\hat{M}_{r2}$ | $\hat{M}_{r3}$ | $MSE_{r1}$ | $MSE_{r2}$ | $MSE_{r3}$ |
| 10 | 1.6681 | 1.7793 | 1.7682 | 0.0525 | 0.1075 | 0.0954 |
| 20 | 1.6056 | 1.6512 | 1.6587 | 0.0178 | 0.0218 | 0.0186 |
| 30 | 1.5999 | 1.6278 | 1.6229 | 0.0129 | 0.0136 | 0.0112 |
| 40 | 1.5903 | 1.6113 | 1.6082 | 0.0103 | 0.0103 | 0.0075 |
| 50 | 1.5829 | 1.5992 | 1.5972 | 0.0072 | 0.0071 | 0.0064 |
| 60 | 1.5809 | 1.5954 | 1.5896 | 0.0055 | 0.0057 | 0.0049 |
| 70 | 1.5765 | 1.5885 | 1.5878 | 0.0046 | 0.0046 | 0.0045 |
| 80 | 1.5781 | 1.5886 | 1.5857 | 0.0044 | 0.0041 | 0.0034 |
| 90 | 1.5752 | 1.5845 | 1.5779 | 0.0038 | 0.0038 | 0.0032 |
| 100 | 1.5731 | 1.5814 | 1.5775 | 0.0032 | 0.0032 | 0.0031 |

**Table 3.** The coverage probability of $100(1-\alpha)\%$ ACIs with different $\alpha$ .

| Sample size ($n$) | Shannon entropy | | Rényi entropy | |
|---|---|---|---|---|
| | $\alpha = 0.1$ | $\alpha = 0.05$ | $\alpha = 0.1$ | $\alpha = 0.05$ |
| 10 | 0.9637 | 0.9752 | 0.9662 | 0.9791 |
| 20 | 0.9798 | 0.9894 | 0.9789 | 0.9884 |
| 30 | 0.9829 | 0.9916 | 0.9847 | 0.9930 |
| 40 | 0.9839 | 0.9941 | 0.9860 | 0.9953 |
| 50 | 0.9857 | 0.9946 | 0.9894 | 0.9957 |
| 60 | 0.9876 | 0.9947 | 0.9936 | 0.9954 |
| 70 | 0.9875 | 0.9947 | 0.9925 | 0.9965 |
| 80 | 0.9875 | 0.9940 | 0.9929 | 0.9972 |
| 90 | 0.9894 | 0.9955 | 0.9934 | 0.9966 |
| 100 | 0.9865 | 0.9950 | 0.9929 | 0.9971 |

Based on the above tables, the following conclusions can be drawn:

(1) For Shannon entropy, the ML estimation performs better than the Bayesian estimation. While for Rényi entropy, the performance of ML estimation is similar to Bayesian estimation.

(2) In Bayesian estimation, it is better to select the SE to estimate Shannon entropy. On the contrary, it is better to select the SSE to estimate Rényi entropy.

(3) The sample size has a great influence on Shannon entropy than Rényi entropy. When the sample size increases gradually, the Bayesian estimation of Shannon entropy under SE is close to the ML estimation, but it has no obvious effect on Rényi entropy.

(4) In Table 3, it can be noted that the coverage probability of ACIs is quite close to confidence levels.

## 6. Real data analysis

There is a real data set given by Bjerkdal [46], which represents the survival time (in days) of guinea pigs after the injection of different doses of tubercle bacilli. Kundu and Howlader [47] proved that this set of data sets using the IWD fitting effect is very good, therefore, this data set can be seen as a sample of IWD. In Reference [46], the regimen number refers to the common logarithm of bacillary units contained in 0.5 ml of challenge solution. In other words, regimen 6.6 represents $4.0*10^6$ bacillary units per 0.5 ml. Corresponding to regimen 6.6, the 72 observed observations are listed as follows:

12, 15, 22, 24, 24, 32, 32, 33, 34, 38, 38, 43, 44, 48, 52, 53, 54, 54, 55, 56, 57, 58, 58, 59, 60, 60, 60, 60, 61, 62, 63, 65, 65, 67,68, 70, 70, 72, 73, 75, 76, 76, 81, 83, 84, 85, 87, 91, 95, 96, 98, 99, 109, 110, 121, 127, 129, 131, 143, 146, 146, 175, 175, 211,233, 258, 258, 263, 297, 341, 341, 376.

Using the proposed estimates described in the above sections, the ML estimates and Bayesian estimates of Shannon entropy and Rényi entropy are displayed in Table 4. It's obvious that the ML estimates of entropies are all smaller than Bayesian estimates under SE respectively, and the Bayesian estimates under SSE of entropies are all smaller than ML estimates respectively.

**Table 4.** The estimates and ACIs of entropies based on the real data set.

| | ML estimates | Bayesian estimates | | $100(1-\alpha)\%$ ACIs | |
|---|---|---|---|---|---|
| | | Under SE | Under SSE | $\alpha = 0.1$ | $\alpha = 0.05$ |
| Shannon entropy | 5.6307 | 5.6998 | 4.8706 | (5.1858, 6.0757) | (5.1328, 6.1287) |
| Rényi entropy | 5.4129 | 4.7280 | 4.8706 | (5.1877, 5.6381) | (5.1609, 5.6649) |

## 7. Conclusions

This paper considers the Bayesian estimations of Shannon entropy and Rényi entropy based on two-parameter IWD. First, the expressions of these entropies of two-parameter IWD are derived in Theorem 1. For ML estimation, due to the invariance of ML estimation, the ML estimators of parameters are obtained by the dichotomy method at first. And then, the ML estimators of entropies can be obtained. Additionally, the approximate confidence intervals are given by the Delta method. For Bayesian estimation, the symmetric entropy loss function and scale squared error loss function are chosen. However, the forms of Bayesian estimators are complex and difficult to calculate. Lindley approximation is used to solve this problem. Finally, the mean square errors of the above estimators are used to compare their performances. For Shannon entropy, it's better to use ML estimator. And for Rényi entropy, the performances of the ML estimator and Bayesian estimator are analogous.

## References

1. Shannon, C. E. A mathematical theory of communication. *Bell. Labs. Tech. J.* **1948**, *27*, 379-423
2. Rényi, A. On measures of entropy and information. *Berkeley Symp on Math. Statist. and Prob.* **1970**, *4*, 547-561
3. Alexander Bulinsk; Denis Dimitrov. Statistical estimation of the Shannon entropy. *Acta Math Sin* **2019**, *35*, 17-46
4. Wolf Ramona. Information and entropies. *Lect Notes Phys* **2021**, *988*, 53-89
5. Chacko, M.; Asha, P. S. Estimation of entropy for generalized exponential distribution based on record values. *J. Indian Soc. Prob. St.* **2018**, *19*, 79-96
6. Liu, S.; Gui, W. Estimating the entropy for Lomax distribution based on generalized progressively hybrid censoring. *Symmetry* **2019**, *11*, 1-17
7. Shrahili, M.; El-Saeed A.R.; Hassan, A.S.; Elbatal, I. Estimation of entropy for Log-Logistic distribution under progressive type II censoring. *J. Nanomater* **2022**, *3*, 1-10
8. Mahmoud, M.R.; Ahmad, M.A.M.; Mohamed, B.S.Kh. Estimating the entropy and residual entropy of a Lomax distribution under generalized type-II hybrid censoring. *Math. Stat.* **2021**, *9*, 780-791
9. Hassan O.; Mazen N. Product of spacing estimation of entropy for inverse Weibull distribution under progressive type-II censored data with applications. *J. Taibah Univ. Sci.* **2022**, *16*, 259-269
10. Mavis, P.; Gayan, W. L.; Broderick, O. O. A New Class of Generalized Inverse Weibull Distribution with Applications. *J Appl Math Bioinformatics* **2014**, *4*, 17-35
11. Basheer, A. M. Alpha power inverse Weibull distribution with reliability application. *J Taibah Univ Sci* **2019**, *13*, 423-432
12. Valeriia, S.; Broderick, O. O. Weighted Inverse Weibull Distribution: Statistical Properties and Applications. *Theor Math Appl* **2014**, *4*, 1-30
13. Keller A Z; Kamath A R R. Alternative reliability models for mechanical systems. 3rd International Conference on Reliability and Maintainability, Toulose, France, 1982.
14. Abhijit, C; Anindya, C. Use of the Fréchet distribution for UPV measurements in concrete. *NDT E Int.* **2012**, *52*, 122-128
15. Chiodo, E; Falco, P D; Noia, L P D; Mottola, F. Inverse loglogistic distribution for Extreme Wind Speed modeling: Genesis, identification and Bayes estimation. *AIMS Energy* **2018**, *6*, 926-948
16. A O Langlands; S J Pocock; G R Kerr; S M Gore. Long-term survival of patients with breast cancer: a study of the curability of the disease. *Brit Med J* **1979**, *2*, 1247-1251
17. Ellah A. Bayesian and non-Bayesian estimation of the inverse Weibull model based on generalized order statistics. *Intell Inf Manag* **2012**, *4*, 23-31
18. Singh, S K; Singh, U; Kumar, D. Bayesian estimation of parameters of inverse Weibull distribution. *J Appl Stat* **2013**, *40*, 1597-1607
19. Asuman, Y.; Mahmut, K. Reliability estimation and parameter estimation for inverse Weibull distribution under different loss functions. *Kuwait J. Sci.* **2022**, *49*, 1-24
20. Sultan, K.S.; Alsadat, N.H.; Kundu, D. Bayesian and maximum likelihood estimations of the inverse Weibull parameters under progressive type-II censoring. *J. Stat. Comput. Sim.* **2014**, *84*, 2248-2265

21.  Amirzadi, A.; Jamkhaneh, E.B.; Deiri, E. A comparison of estimation methods for reliability function of inverse generalized Weibull distribution under new loss function. *J. Stat. Comput. Sim.* **2021**, *91*, 2595-2622

22.  Peng, X.; Yan, Z. Z. Bayesian estimation and prediction for the inverse Weibull distribution under general progressive censoring, *Commu. Stat.- Theor. M.* **2016**, *45*, 621-635,

23.  Sindhu, T.N; Feroze, N; Aslam, M. Doubly censored data from two-component mixture of inverse Weibull distributions: Theory and Applications. *J. Mod. Appl. Stat. Meth.* **2016**, *15*, 322-349

24.  Mohammad, F.; Sana, S. Bayesian estimation and prediction for the inverse Weibull distribution based on lower record values. *J. Stat. Appl. Probab.* **2021**, *10*, 369-376

25.  Faud, S. Al-Duais. Bayesian analysis of Record statistic from the inverse Weibull distribution under balanced loss function. *Math. Probl. Eng.* **2021**, *2021*, 1-9, Article ID 6648462. https://doi.org/10.1155/2021/6648462

26.  Li, C.P.; Hao, H.B.. Reliability of a stress-strength model with inverse Weibull distribution. *Int. J. Appl. Math.* **2017**, *47*, 302-306

27.  Ismail, A.; Al Tamimi, A. Optimum constant-stress partially accelerated life test plans using type-I censored data from the inverse Weibull distribution. *Strength Mater* **2017**, *49*, 847–855.

28.  Kang, S.B.; Han, J.T. The graphical method for goodness of fit test in the inverse Weibull distribution based on multiply type-II censored samples. *SpringerPlus* **2015**, *4*, 768. https://doi.org/10.1186/s40064-015-1554-x

29.  Saboori, H.; Barmalzan, G.; Ayat, S.M. Generalized modified inverse Weibull distribution: Its properties and applications. *Sankhya B* **2020**, *82*, 247–269.

30.  Debasis Kundu; Hatem Howlader. Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data. *Comput Stat Data Anal* **2010**, *54*, 1547-1558

31.  Sultan, K.S.; Alsadat, N.H.; Kundu, D.. Bayesian and maximum likelihood estimations of the inverse Weibull parameters under progressive type-II censoring. *J Stat Comput Simul* **2014**, *84*, 2248-2265

32.  Mohammad K; Mina A. Estimation of the Inverse Weibull Distribution Parameters under Type-I Hybrid Censoring. *Austrian J. Stat.* **2021**, *50*, 38-51

33.  Ali Algarni; Mohammed Elgarhy; Abdullah M Almarashi; Aisha Fayomi; Ahmed R El-Saeed. Classical and Bayesian Estimation of the Inverse Weibull Distribution: Using Progressive Type-I Censoring Scheme. *Adv. Civ. Eng.* **2021**, *2021*, 1-15

34.  Zhou, J.H.; Luo, Y. Bayes information updating and multiperiod supply chain screening. *Int J Prod Econ* **2023**, *256*, 108750-108767

35.  Yulin, S; Deniz, G; Soung, C L. Bayesian Over-the-Air Computation. *IEEE J. Sel. Areas Commun*. **2023**, *41*, 589-606

36.  Taborsky, P; Vermue, L; Korzepa, M; Morup, M. The Bayesian Cut. *IEEE Trans. Pattern Anal. Mach. Intell*. **2021**, *43*, 4111-4124

37.  Oliver Korup. Bayesian geomorphology. *Earth Surf Process Landf.* **2021**, *46*, 151-172

38.  Ran, E; Kfir, E; Mu, X. S. Bayesian privacy. *Theor. Econ.* **2021**, *16*, 1557-1603

39.  Luo, C.L.; Shen, L.J.; Xu, A.C. Modelling and estimation of system reliability under dynamic operating environments and lifetime ordering constraints. *Reliab. Eng. Syst. Saf.* **2022**, *218*, 108136-108145

40.  Peng, W. W.; Li, Y. F.; Yang, Y. J.; Mi, J. H.; Huang, H.Z. Bayesian Degradation Analysis With Inverse Gaussian Process Models Under Time-Varying Degradation Rates. *IEEE Trans Reliab* **2017**, *66*, 84-96

41.  František, B; Frederik, A; Julia, M. H. Informed Bayesian survival analysis. *BMC Medical Res. Methodol.* **2022**, *22*, 238-260

42.  Liu F.; Hu X.G.; Bu C.Y.; Yu K. Fuzzy Bayesian Knowledge Tracing. *IEEE Trans Fuzzy Syst* **2022**, *30*, 2412-2425

43.  Xu, B.; Wang, D.H.; Wang, R.T. Estimator of scale parameter in a subclass of the exponential family under symmetric entropy loss. *Northeast. Math. J.* **2008**, *24*, 447-457

44.  Li, Q.; Wu, D. Bayesian analysis of Rayleigh distribution under progressive type-II censoring. *J. Shanghai Polytech. Univ.* **2019**, *36*,114-117

45.  Song, L.X.; Chen, Y.S.; Xu, J.M. Bayesian estimation of Pission distribution parameter under scale squared error loss function. *J. Lanzhou Univ. Tech.* **2008**, *34*, 152-154.

46.  Bjerkdal, Tor. Acquisition of resistance in guinea pigs infected with different doses of virulent tubercle bacilli. *Am J Epidemiol*, **1960**, *72*, 130-148

47.  Debasis Kundu; Hatem Howlader. Bayesian inference and prediction of the inverse Weibull distribution for Type-II censored data. *Comput Stat Data An* **2010**, *54*, 1547-1558