# Bayesian Estimation of Entropy for Kumaraswamy Distribution and Its Application to Progressively First-Failure Censored Data

**Abstract:** Entropy can be mathematically defined as a measure of the uncertainty of random variables that represents the potential quantity of information. This article investigates the behavior of the entropy of random variables which follow a Kumaraswamy distribution using progressively first-failure censored (PFFC) data. In particular, we calculate the maximum likelihood estimation and the confidence interval of entropy by using the observed Fisher information matrix through the asymptotic distribution of the maximum likelihood estimator. Furthermore, we apply the Markov Chain Monte Carlo (MCMC) method which help us to estimates the entropy and to formulation the credible intervals in order to address this problem. Here, a numerical example of real data is presented to illustrate the performance of the proposed method. Finally, we perform Monte Carlo simulations to observe the behavior of the proposed procedure.

**Keywords:** Kumaraswamy distribution; Entropy; PFFC; Maximum likelihood method; Markov Chain Monte Carlo method; Credible intervals.

## 1 Introduction

Censoring is popular form in many areas, including engineering, social economy and pharmacology sciences. In particular, it is related to reliability and survival analysis. In real applications, it is complicated to study all the data because of time and cost constrains. Then, Type-I and Type-II are censorship schemes used in different test experiments of life. In Type-I, place $n$ units into life experiments and end the experiment after a given time while Type-II ends the experiment after the given number of units $m$ has failed. For more details about the theoretical strategies of progressive censoring, see Balakrishnan and Aggarwala [1]. The progressive censoring scheme was discussed in type-I, type-II and hybrid progressive censoring schemes. The experiment under a high realible products tested under the last censoring schemes may take a long period of time. One of the most solutions of this problem is to category the test units into multiple sets with the same number of units and the first failure in each category is recorded which called first failure censoring scheme see Jun et al. [2] and Fernàndez et al. [3]. Under the first failure censoring scheme, the expermint is terminated when the first failure in each set is recorded. The problem of removed sets from the experiment, before the final point was defined as a PFFC scheme, is discussed and developed by Wu and Kuş [4]. Both first failure and progressive censoring schemes are combined. Many researchers have discussed the study of the PFFC, e.g. Soliman et al. [5], Modhesh [6] and Yu et al. [7]. Now a days, the PFFC scheme has become a popular censoring scheme for life testing experiments as it saves considerable amount of cost and time of the experiment. In addition, other studies explain the PFFC, such as, Chaturvedi et al. [8], Kumar et al. [9] and Saini et al. [10]. Assuming $n$ independent groups with $k$ items within each group are subjected to a life test. $R_1$ groups and the group in which the first failure is observed are randomly removed from the test as soon as the first failure (say $x^{\mathbf{R}}_{1:m:n:k}$) has occurred, In general, $R_m$ $(m \leq n)$ groups and the group in which the $m^{th}$ failure is observed are randomly removed from the test when the $m^{th}$ failure (say $x^{\mathbf{R}}_{m:m:n:k}$) has occurred. The observations $x^{\mathbf{R}}_{1:m:n:k} < x^{\mathbf{R}}_{2:m:n:k} < ... < x^{\mathbf{R}}_{m:m:n:k}$ are called PFFC order statistics with progressive censoring scheme $\mathbf{R} = (R_1, R_2, ..., R_m)$. Where $m$ refers to the number of the first failure observed $(1 < m \leq n)$ and $n = m + R_1 + R_2 + ... + R_m$. The PFFC scheme containing four special cases; first-failure censored scheme, progressive type II censored, type II censored and complete sample.

If $X$ follows a two-parameters Kumaraswamy distribution, then the probability density function (pdf) of $X$ is defined as:

$$f(x) = \alpha\beta x^{\beta-1}(1 - x^{\beta})^{(\alpha-1)},\ 0 < x < 1,\ (\alpha > 0,\ \beta > 0), \tag{1}$$

and the cumulative distribution function (cdf) of X is notated by

$$F(x) = 1 - (1 - x^{\beta})^{\alpha},\ 0 < x < 1. \tag{2}$$

Then, the corresponding reliability rate functions of this distribution at some $t$, is given, by

$$S(t) = (1 - t^{\beta})^{\alpha},\ 0 < t < 1, \tag{3}$$

as well as, the failure rate functions is

$$H(t) = \alpha\beta t^{\beta-1}(1 - t^{\beta})^{-1},\ 0 < t < 1. \tag{4}$$

The two-parameters Kumaraswamy distribution is unimodal for $\alpha, \beta > 1$, uniantimodal for $\alpha, \beta < 1$, increasing for $\alpha \leq 1$ and $\beta > 1$, decreasing for $\alpha > 1$ and $\beta \leq 1$ and constant for $\alpha = \beta = 1$ . There are some similarities and differences between the beta and Kumaraswamy distributions investigated by Jones [11]. The Kumaraswamy distribution is applicable to a variety of hydrological problems and natural phenomena whose procedure values are limited on both sides. In hydrology and relevant filds, the Kumaraswamy distribution received appreciable interest, Cordeiro et al. ([12], [13]). In the general Kumaraswamy distribution is more similar of the same basic shape to beta distribution and hence, it's referred to as a "Beta-like" distribution. Also, the Kumaraswamy distribution is easier to use and more manageable in some situations. Kumaraswamy distribution has closed form of cdf, which make it preferred over the Beta. Unlike the beta, the cdf of Kumaraswamy distribution does not include the incomplete Beta function, which makes it much easier to work with (Michalowicz et al. [14]). The pdf and quantile functions are also of closed form, which make the Kumaraswamy distribution to be a more suitable option for multiple applications-particularly in simulation studies. Other no table advantages over the beta include an explicit formula for moments of order statistics and a simple formula for generating of random variables (Ishaq et al. [15]).

Entropy can be defined as a quantitive measure of the uncertainty of each probability distribution (Put a reference here please). In different fields, there are many applications that use entropy such as statistics, neurobiology, cryptography, bioinformatics, and computer science. For Instance, entropy is used to evaluate the probability distribution of electric charge among atoms observed under certain conditions. The Shannon entropy $H(X)$, is defined as

$$H(X) = -\int_{-\infty}^{\infty} f(x) \log(f(x)) dx, \tag{5}$$

where $X$ is a random variable of $f(x)$.

Canon entropy has been studied by many researchers, including in this research, for example, Wang and Gui [16] who studied some inferences of entropy based on Type II progressive for Burr Type XI distribution. Additionally, their inference Cho et. al [17] and Zhao et al. [18] studies the entropy based on doubly generalized Type-II hydrid censoring and under type-II censored data respectively.

The aim of this paper is to develop estimation process for entropy of Kumaraswamy distribution through a general censoring scheme (i.e. PFFC data). In the classical (frequentist) estimation, the maximum likelihood estimates are discussed with their asymptotic confidence intervals (ACIs). In contrast, the Bayesian estimation using (MCMC) method is built for credible intervals. In this paper, we use and analyze real data set for illustrated goals. Furthermore, we use the Monte Carlo simulation study to assess and compare the theoretical results.

The rest of this paper is elaborated as follows:

In Section 2, we discussed the maximum likelihood estimation of parameters and entropy with the PFFC data. At the same time, the corresponding ACIs of entropy are also provided. In Section 3, we work out the Bayesian estimation of parameters and entropy using MCMC methods. We analyzed real data to illustrate the novel methods in Section 4. In Section 5, we compare our results under different prior distributions using the Monte Carlo method. Finally, the paper is concluded in Section 6.

# 2 Estimation of the Parameters and Entropy

Here, we estimate both parameters $\alpha$ and $\beta$, and entropy by providing the maximum likelihood (ML) approach. Then, we compute the asymptotic confidence interval for entropy

## 2.1 Maximun Likelihood Estimation

Suppose $X_i = X^{\mathbf{R}}_{i:m:n:k}$, $i = 1, 2, ..., m$, is the PFFC order statistics from Kumaraswamy distribution, with censored scheme $\mathbf{R} = (R_1, R_2, ..., R_m)$. In other words, we choose $x_i$ to represent $x^{\mathbf{R}}_{i:m:n:k}$, $i = 1, 2, ..., m$. The joint pdf is

$$f_{1,2,...,m}(x^{\mathbf{R}}_{1:m:n:k}, x^{\mathbf{R}}_{2:m:n:k}, ..., x^{\mathbf{R}}_{m:m:n:k}) = Ck^m \prod_{i=1}^{m} f(x^{\mathbf{R}}_{i:m:n:k})(1 - F(x^{\mathbf{R}}_{i:m:n:k}))^{k(\mathbf{R}_i+1)-1} \tag{6}$$

$$0 < x^{\mathbf{R}}_{1:m:n:k} < x^{\mathbf{R}}_{2:m:n:k} < ... < x^{\mathbf{R}}_{m:m:n:k} < \infty,$$

where

$$C = n(n - R_1 - 1)(n - R_1 - R_2 - 1)...(n - R_1 - R_2 - ...R_{m-1} - m + 1).$$

Combining (1), (2), and (6), the likelihood function is

$$\ell(\underline{x}; \alpha, \beta) = Ck^m \alpha^m \beta^m \prod_{i=1}^{m} x_i^{\beta-1}(1 - x_i^{\beta})^{\alpha k(\mathbf{R}_i+1)-1}. \tag{7}$$

The log-likelihood function is defined as:

$$\begin{aligned} L(\underline{x}; \alpha, \beta) \quad &= \quad \log c + m \log k + m \log \alpha + m \log \beta + (\beta - 1) \sum_{i=1}^{m} \log(x_i) \\ &+ \sum_{i=1}^{m} (\alpha k(R_i + 1) - 1) \log(1 - x_i^{\beta}). \end{aligned} \tag{8}$$

By differentiating (8) with respect to $\alpha$ and $\beta$, and then equating each result to zero, the following two equations are satisfied to gain MLE

$$\frac{\partial L(\underline{x}; \alpha, \beta)}{\partial \alpha} = \frac{m}{\alpha} + \sum_{i=1}^{m} k(R_i + 1) \log(1 - x_i^{\beta}), \tag{9}$$

and

$$\frac{\partial L(\underline{x}; \alpha, \beta)}{\partial \beta} = \frac{m}{\beta} + \sum_{i=1}^{m} \log(x_i) - \sum_{i=1}^{m} (\alpha k(R_i + 1) - 1) \frac{x_i^{\beta} \log(x_i)}{(1 - x_i^{\beta})}. \tag{10}$$

From (9) the MLE of $\alpha$ denoted by $\hat{\alpha}$ is

$$\hat{\alpha} = -m/\hat{\eta}(\underline{x}, \hat{\beta}), \tag{11}$$

where $\hat{\eta}(\underline{x}, \hat{\beta}) = 1/\sum_{i=1}^{m} k(R_i + 1) \log(1 - x_i^{\hat{\beta}})$. By substituting (11) in (10), we get

$$\frac{m}{\hat{\beta}} + \sum_{i=1}^{m} \log(x_i) - (m/\hat{\eta}(\underline{x}, \hat{\beta})) \sum_{i=1}^{m} (R_i + 1) \frac{x_i^{\hat{\beta}} \log(x_i)}{(1 - x_i^{\hat{\beta}})} - \sum_{i=1}^{m} \frac{x_i^{\hat{\beta}} \log(x_i)}{(1 - x_i^{\hat{\beta}})} = 0. \tag{12}$$

To solve Eq. (12) by considering Eq. (11), we employed Newton-Raphson iteration to obtain the MLE of $\widehat{\beta}$, and then $\widehat{\alpha}$. Also, the starting value for the root-finding method can be obtained by using the graphical method [19]. In the line with the invariance property of the MLE, there is no difference to find the estimator for function parameters. For example, If we simplify the entropy in Eq. (5) under Kumaraswamy distribution into a function of $\alpha$ and $\beta$, then the MLE of entropy can be executed.

**Theorem 1**. Let $X$ be a random variable with pdf (1), then the entropy of $X$ is

$$H(X) = H(f) = -\ln(\alpha\beta) - (\frac{\beta-1}{\beta})[\psi(1) - \psi(\alpha+1)] - \frac{1}{\alpha} + 1, \tag{13}$$

where $\psi$ is defined by $\psi(z) = \frac{d}{dz}\ln\Gamma(z)$, which is also called the digamma function. $\Gamma$ is the gamma function.

**Proof:**

For the (pdf) given in Eq. (1), the log-density is

$$\ln(f(x)) = \ln(\alpha\beta) + (\beta-1)\ln x + (\alpha-1)\ln(1-x_i^{\beta}), \tag{14}$$

then the entropy is

$$\begin{aligned} H(X) &= -\int_0^1 f(x)\log(f(x))dx \\ &= -\ln(\alpha\beta) - (\beta-1)E[\ln x] - (\alpha-1)E[\ln(1-x_i^{\beta})]. \end{aligned} \tag{15}$$

Therefore, it is important to analyze and calculate $E[\ln x]$ and $E[\ln(1-x_i^{\beta})]$. The derivative of $\alpha$ on both sides of $\int_0^1 f(x)dx$ is given by

$$\int_0^1 [\beta x^{\beta-1}(1-x^{\beta})^{(\alpha-1)} + f(x)\ln(1-x_i^{\beta})]dx = 0, \tag{16}$$

from (16), we have

$$E[\ln(1-x_i^{\beta})] = \int_0^1 f(x)\ln(1-x_i^{\beta})dx = -\frac{1}{\alpha}, \tag{17}$$

remining calculate

$$E(X^r) = \int_0^1 X^r \alpha\beta x^{\beta-1}(1-x)^{\alpha-1}dx = 0. \tag{18}$$

Put $x^{\beta} = t$ in (18), which is yield

$$E(X^r) = \alpha\int_0^1 t^{\frac{r}{\beta}}(1-t)^{(\alpha-1)}dt = \alpha B\left(\frac{r}{\beta}+1, \alpha\right). \tag{19}$$

By the derivative of both sides of (19) with respect to $r$, we get

$$\frac{dE(X^r)}{dr} = E(X^r\ln x) = \frac{\alpha B\left(\frac{r}{\beta}+1,\alpha\right)}{\beta}[\psi(\frac{r}{\beta}+1) - \psi(\frac{r}{\beta}+\alpha+1)]. \tag{20}$$

When $r = 0$, which is yield

$$E(\ln x) = \frac{1}{\beta}[\psi(1) - \psi(\alpha+1)]. \tag{21}$$

Substituting (17) and (21) into (15), we get

$$H(X) = H(f) = -\ln(\alpha\beta) - (\frac{\beta-1}{\beta})[\psi(1) - \psi(\alpha+1)] - \frac{1}{\alpha} + 1.$$

### 2.2 Asymptotic Confidence Interval (ACI) for Entropy

To attain the ACI of information entropy, the observed Fisher information matrix of $\alpha$ and $\beta$ is set initially as

$$I(\hat{\alpha},\hat{\beta}) = \begin{pmatrix} -\frac{\partial^2 L}{\partial \alpha^2} & -\frac{\partial^2 L}{\partial \alpha \partial \beta} \\ -\frac{\partial^2 L}{\partial \beta \partial \alpha} & -\frac{\partial^2 L}{\partial \beta^2} \end{pmatrix}_{(\hat{\alpha},\hat{\beta})},$$

where

$$\frac{\partial^2 L}{\partial \alpha^2} = -\frac{m}{\alpha^2}, \tag{22}$$

$$\frac{\partial^2 L}{\partial \beta^2} = -\frac{m}{\beta^2} - \sum_{i=1}^{m} (\alpha k(R_i+1)-1) \frac{x_i^{\beta} \log^2(x_i)}{(1-x_i^{\beta})^2}, \tag{23}$$

$$\frac{\partial^2 L}{\partial \alpha \partial \beta} \equiv \frac{\partial^2 L}{\partial \beta \partial \alpha} = -k \sum_{i=1}^{m} (R_i+1) \frac{x_i^{\beta} \log(x_i)}{(1-x_i^{\beta})}. \tag{24}$$

In point of fact, we use $I^{-1}(\hat{\alpha},\hat{\beta})$ to estimate $I^{-1}(\alpha,\beta)$. Namely, the asymptotic normality of the MLE can be applied to compute the approximate confidence intervals for $\alpha$ and $\beta$. Consequently, $(1-\gamma)100\%$ confidence intervals for $\alpha$ and $\beta$ turn into

$$(\hat{\alpha} \pm Z_{\gamma/2}\sqrt{var(\hat{\alpha})},\ \hat{\beta} \pm Z_{\gamma/2}\sqrt{var(\hat{\beta})}), \tag{25}$$

where $Z_{\gamma/2}$ is the percentile of the standard normal distribution with right-tail probability $\gamma/2$.

Due to the fact that we want to acquire the asymptotic confidence interval of entropy in (13), which is a function of $\alpha$ and $\beta$, we are obligated to use the delta method to calculate the variance of entropy. In which, we calculate the variance of a linear function using the linear approximation of a function. Let

$$A = \left( \frac{\partial H(f)}{\partial \alpha}, \frac{\partial H(f)}{\partial \beta} \right), \tag{26}$$

where

$$\frac{\partial H(f)}{\partial \alpha} = -\frac{1}{\alpha} + (\frac{\beta-1}{\beta})\acute{\psi}(\alpha+1)] + \frac{1}{\alpha^2},$$

$$\frac{\partial H(f)}{\partial \beta} = -\frac{1}{\alpha} - \frac{1}{\beta^2}[\psi(1) - \psi(\alpha+1)].$$

Then, the variance of entropy is donated by

$$Var(\hat{H}(f)) = [AI^{-1}(\hat{\alpha},\hat{\beta})A^t]_{(\hat{\alpha},\hat{\beta})},$$

where $A^t$ is transposed A. Then, the $100(1-\gamma)\%$ ACI of entropy is

$$(\hat{H}(f) \pm Z_{\gamma/2}\sqrt{\hat{H}(f)}). \tag{27}$$

## 3 Bayes Estimation and Credible Intervals

In this section, when both parameters $\alpha$ and $\beta$ are unknown, Bayesian estimation would be used to estimate the entropy and corresponding credible intervals. To compute the Bayes estimation, only the squared error loss (SEL) function can be presumed mainly. Nevertheless, some other loss functions could be easily combined. Many statisticians use the Bayesian approach to evaluate parameters and associated functions for various distributions. For instance, Wang and Gui [20] used the Bayesian of entropy for Burr Type XII distribution under progressive Type-II censored data. In addition, Yu et al. [7] considered ML and Bayesian methods on the Shannon entropy of inverse Weibull distribution under the PFFC. Furthermore, Shi et al. [21] attained estimation for entropy and parameters of generalized Bilal distribution under an adaptive Type II progressive hybrid censoring scheme. In this paper, we presume the prior distribution of unknown parameters in order to examine the Bayes estimators. Here, both $\alpha$ and $\beta$ may have independent gamma priors with the pdfs

$$\pi_1(\alpha|a,b) \propto \alpha^{a-1}e^{-b\alpha} \quad \text{if } \alpha > 0, \tag{25}$$

and

$$\pi_2(\beta|c,d) \propto \beta^{c-1}e^{-d\beta} \quad \text{if } \beta > 0. \tag{26}$$

Where the gamma distribution is widely used in the field of Business, Engineering and reliability theory. Because that, gamma distribution is the continuous model variable that should have a positive and skewed pdf. The gamma distribution is a kind of statistical distributions which is related to the, Kumaraswamy distribution (beta distribution). Other, different shape of gamma distribution which make it useful to cover widely applications.

The product of (25), (26) and (7) give the joint posterior density for $\alpha$ and $\beta$ which written as:

$$\begin{aligned}\pi^*_{\alpha,\beta}(\alpha,\beta|data) \quad &\propto \quad \alpha^{m+a-1}\beta^{m+c-1}\prod_{i=1}^{m}\left(\frac{x_i^{\beta}}{1-x_i^{\beta}}\right) \\ &\times \exp\left[-\left\{\alpha\left(b-\sum_{i=1}^{m}k\,(R_i+1)\log(1-x_i^{\beta})\right)+d\beta\right\}\right]. \end{aligned}\tag{27}$$

Therefore, we use MCMC method (namely the Gibbs sampler) in order to simulate samples from the posterior distribution such that the inferences of samples can be drawn. Using Eq. (27), the posterior distributions for $\alpha$ is

$$\pi^*_{\alpha}(\alpha|\beta,data) \propto \alpha^{m+a-1}\exp\left[-\alpha\left(b-\sum_{i=1}^{m}k\,(R_i+1)\log(1-x_i^{\beta})\right)\right]. \tag{28}$$

Likewise, the marginal posterior density of $\beta$ is proportional to

$$\begin{aligned}\pi^*_{\beta}(\beta|\alpha,data) \quad &\propto \quad \beta^{m+c-1}\prod_{i=1}^{m}\left(\frac{x_i^{\beta}}{1-x_i^{\beta}}\right) \\ &\times \exp\left[-\left\{d\beta-\alpha\left(\sum_{i=1}^{m}k\,(R_i+1)\log(1-x_i^{\beta})\right)\right\}\right]. \end{aligned}\tag{29}$$

The gamma density is shown in Eq. (28) with shape $(m+a)$ and scale parameter $\left(b - \sum_{i=1}^{m} k\,(R_i + 1)\log\right.$ parameters. Accordingly, by using any gamma generating routine, we could generate samples of $\alpha$ easily. The conditional posterior density of $\beta$ Eq. (29) for the PFFC Kumarswamy model, show that standard sampling scheme is not feasible since the conditional distribution is not of a known form. We adopt Bayesian inference for $\beta$ using Metropolis-Hastings algorithm for more details one can see Chib and Greenberg [22], taking into account the conditional distribution as a target density. Thus, we proceed as follows:

It can be seen that Eq. (28) is a gamma density with shape parameter $(m + a)$ and scale parameter $\left(b - \sum_{i=1}^{m} k\,(R_i + 1)\log(1 - x_i^{\beta})\right)$ and, therefore, samples of $\alpha$ can be easily generated using any gamma generating routine. The conditional posterior density of $\beta$ Eq. (29) for the PFFC Kumarswamy model, show that standard sampling scheme is not feasible since the conditional distribution is not of a known form. Bayesian inference for the parameter $\beta$ can be performed by Metropolis-Hastings algorithm (see, for example, Chib and Greenberg [22]) considering the conditional distributions as the target densities. Thus, we proceed as follows:

**Step1:** Start with any point $\beta^{(0)}$ and stage indicator $t = 1$.

**Step2:** Generate $\alpha^{(t)}$ from $Gamma(m + a, b - \sum_{i=1}^{m} k\,(R_i + 1)\log(1 - x_i^{\beta^{(t-1)}}))$.

**Step3:** Using Metropolis-Hastings algorithm, with a target distribution $\pi_{\beta}^{*}(\beta|\alpha, data)$, generate $\beta^{(t)}$ with the proposal distribution $N(\beta^{(t-1)}, \sigma^2)$ .

**Step4:** Compute $H(f)^{(t)} = H(\alpha^{(t)}, \beta^{(t)})$.

**Step5:** Set $t = t + 1$.

**Step6:** Repeat steps $2 - 5$ $N$ times.

Now the approximate mean of $H(f)$ is

$$\hat{H}(f) = E(H(f)|data) = \frac{1}{N - M} \sum_{i=M+1}^{N} H(f)^{(i)},$$

Where $M$ is defined as the burn-in period (i.e. a number of iterations before the stationary distribution is obtained).
To calculate the credible interval of $H(f)$, order $H(f)_{M+1}, H(f)_{M+2}, \ldots, H(f)_N$ as $H(f)_{(1)}$, $H(f)_{(2)}, \ldots, H(f)_{(N-M)}$. Then the $100(1 - 2\gamma)\%$ symmetric credible interval is
$\left(H(f)_{(\gamma(N-M))},\ H(f)_{((1-\gamma)(N-M))}\right)$ .

# 4 Application of real data on PFFC

In this section, the real-life data set is used which was originally described by Hoel [23]. From a laboratory experiment, this data was gained where male mice got a radiation dose of 300 roentgen. For each meal mice, the reason for death was specified using autopsy. Limiting the analysis to two reasons for death, we consider reticulum cell sarcoma data for purposes of analysis. In this experiment, the data of $m = 38$ meal mice were transformed into the interval [0, 1]. The transformed data is presented below:

(0.420424, 0.421751, 0.529178, 0.656499, 0.696286, 0.710875, 0.787798, 0.790451, 0.802387, 0.811671, 0.823607, 0.832891, 0.83687, 0.843501, 0.852785, 0.85809, 0.859416, 0.860743, 0.876658, 0.87931, 0.883289, 0.888594, 0.921751, 0.924403, 0.928382, 0.935013, 0.944297, 0.945623, 0.97878, 0.992042, 0.998674).

The Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution was computed in order to check the validity of our model. The MLE of $\alpha$ and $\beta$ based on reticulum cell sarcoma data are (1.3086, 5.1244). The result of the K-S distance and the associated p-values are 0.119 and 0.6549. Based on the $p$-values, the estimated Kumarswamy model provides an excellent fit to the given data. The data are randomly ranged into 19 groups with $k = 2$ within each group. Assume that the predetermined PFFC plan is applied using the progressive censoring scheme $\mathbf{R} = (1^4, 0^{11})$. The following PFFC data of size ($m = 15$) out of 19 groups of reticulum cell sarcoma were observed: (0.420424, 0.421751, 0.529178, 0.656499, 0.710875, 0.732095, 0.738727, 0.740053, 0.787798, 0.802387, 0.832891, 0.843501, 0.860743, 0.921751, 0.924403).

We compute different estimates of the entropy namely, ML estimates and Bayes MCMC estimates by using 10000 MCMC samples and discarding the first 1000 values as 'burn-in'. We presume the non-informative gamma priors ($a = b = c = d = 0$), as well as 95% confidence intervals and credible intervals with the corresponding lengths since we have no prior information about the unknown parameters. The results are displayed in Table 1.

We notice from Table 1 that the length credible interval in the Bayes estimate and the length confidence interval in the maximum likelihood method approximately the same. Figure 1 shows the histograms and the posterior density functions for the MCMC output of entropy are plotted. Figure 2 plot the MCMC output of entropy, using 10000 MCMC samples. The posterior mean is represented by the middle line and the lower and upper bounds of the 95% probability intervals are represented by the marginal lines.

Table 1: Point estimates, 95% confidence and credible intervals for the entropy.

| Method of Estimation | Entropy | Estimates | Interval | Length |
|---|---|---|---|---|
| ML | $\widehat{H}(f)$ | -0.7310 | (-0.9818, -0.4802) | 0.5016 |
| Bayes MCMC | $\widetilde{H}(f)$ | -0.7583 | (-0.9893, -0.5237) | 0.4655 |

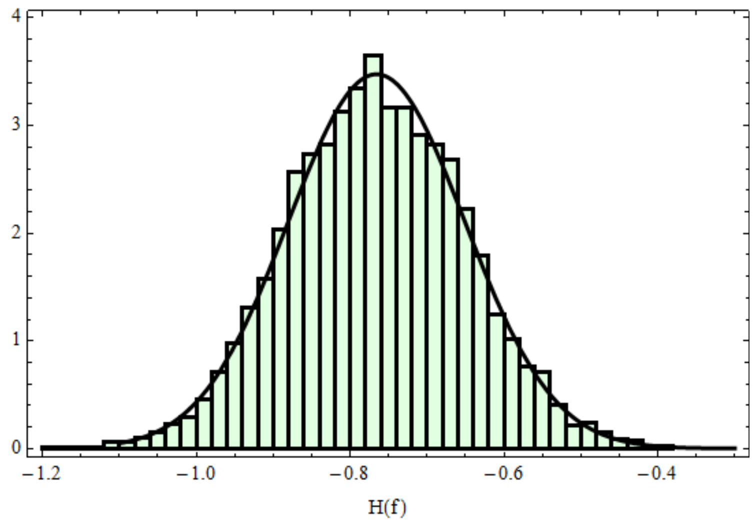


Figure 1:Simulation number of the entropy generated by MCMC method.

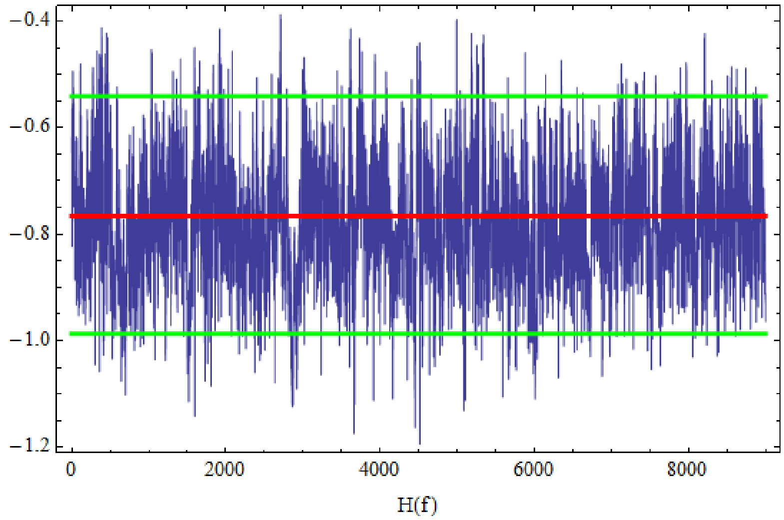


Figure 2: Histogram of entropy generated by MCMC method.

# 5 Monte Carlo Simulations

In this section, the entropy estimators obtained by different methods were computed by using the Monte Carlo simulation method. The PFFC samples are generated and simulated from several censoring schemes of $(k, n, m,\ R_i)$ and parameter values using the Kumaraswamy distribution when $\alpha = 2$ and $\beta = 0.5$. In other words, 1000 of these samples were simulated using Kumaraswamy distribution. In addition, these samples were generated by using the algorithm described in Balakrishnan and Sandhu [24]. In this study, we consider the following main patterns of different schemes as:

$S^{(1)}_{k:m:n}$ : All the removals are done at the first failure, i.e. $R_1 = n - m, \quad R_i = 0$ for $i \neq 1$.

$S^{(2)}_{k:m:n}$ : The removals are at middle observations, i.e. $R_{\frac{m}{2}} = n - m$, $R_i = 0$ for $i \neq \frac{m}{2}$.

$S^{(3)}_{k:m:n}$ : The removals are at the last observation, i.e. $R_m = n - m, \quad R_i = 0$ for $i \neq m$.

Table 2: The average MLEs, average Bayesian estimations of $H(f)$ and their mean squared errors for different censoring schemes are reported, when $\alpha = 0.5$, $\beta = 2.0$ and $H(f) = -0.6931$.

| Scheme | MLE | | (Bayes-MCMC prior 0) | | (Bayes-MCMC prior 1) | |
|---|---|---|---|---|---|---|
| | *Mean* | *MSE* | *Mean* | *MSE* | *Mean* | *MSE* |
| $S^{(1)}_{1:20:50}$ | -0.692 | 0.1015 | -0.7021 | 0.1205 | -0.7142 | 0.0853 |
| $S^{(2)}_{1:20:50}$ | -0.6701 | 0.0974 | -0.6678 | 0.0998 | -0.6781 | 0.0739 |
| $S^{(3)}_{1:20:50}$ | -0.6877 | 0.0808 | -0.6498 | 0.0798 | -0.7059 | 0.0725 |
| $S^{(1)}_{1:40:50}$ | -0.6898 | 0.0313 | -0.6981 | 0.0338 | -0.7042 | 0.0315 |
| $S^{(2)}_{1:40:50}$ | -0.7197 | 0.0317 | -0.7252 | 0.0320 | -0.7143 | 0.0324 |
| $S^{(3)}_{1:40:50}$ | -0.7287 | 0.0414 | -0.7409 | 0.0401 | -0.7409 | 0.0399 |
| $S^{(1)}_{1:35:70}$ | -0.7675 | 0.0414 | -0.7607 | 0.0321 | -0.7534 | 0.0388 |
| $S^{(2)}_{1:35:70}$ | -0.7171 | 0.0460 | -0.7254 | 0.0455 | -0.7326 | 0.0434 |
| $S^{(3)}_{1:35:70}$ | -0.7057 | 0.0765 | -0.7204 | 0.0607 | -0.7204 | 0.0607 |
| $S^{(1)}_{1:60:70}$ | -0.7262 | 0.0223 | -0.7276 | 0.0205 | -0.7378 | 0.0267 |
| $S^{(2)}_{1:60:70}$ | -0.7312 | 0.0272 | -0.7357 | 0.0268 | -0.7437 | 0.0299 |
| $S^{(3)}_{1:60:70}$ | -0.7208 | 0.0382 | -0.7306 | 0.0388 | -0.7342 | 0.0388 |
| $S^{(1)}_{4:20:50}$ | -0.7158 | 0.1018 | -0.6776 | 0.0816 | -0.6996 | 0.0695 |
| $S^{(2)}_{4:20:50}$ | -0.6724 | 0.1076 | -0.7106 | 0.0865 | -0.6596 | 0.0433 |
| $S^{(3)}_{4:20:50}$ | -0.7987 | 0.2093 | -0.6821 | 0.0931 | -0.6716 | 0.0448 |
| $S^{(1)}_{4:40:50}$ | -0.6972 | 0.0363 | -0.7115 | 0.0201 | -0.7094 | 0.0294 |
| $S^{(2)}_{4:40:50}$ | -0.7129 | 0.0598 | -0.7098 | 0.0419 | -0.7146 | 0.0303 |
| $S^{(3)}_{4:40:50}$ | -0.7062 | 0.0553 | -0.6982 | 0.0589 | -0.6982 | 0.0489 |
| $S^{(1)}_{4:35:70}$ | -0.7103 | 0.0568 | -0.7115 | 0.0412 | -0.7015 | 0.0347 |
| $S^{(2)}_{4:35:70}$ | -0.7296 | 0.0808 | -0.7280 | 0.0462 | -0.7202 | 0.0448 |
| $S^{(3)}_{4:35:70}$ | -0.7474 | 0.1406 | -0.7237 | 0.0407 | -0.7092 | 0.0461 |
| $S^{(1)}_{4:60:70}$ | -0.6687 | 0.0237 | -0.6900 | 0.0193 | -0.6821 | 0.0186 |
| $S^{(2)}_{4:60:70}$ | -0.7041 | 0.0267 | -0.7057 | 0.0241 | -0.7092 | 0.0226 |
| $S^{(3)}_{4:60:70}$ | -0.7148 | 0.0327 | -0.7122 | 0.0301 | -0.7222 | 0.0237 |

After that, we calculate the MLEs and the 95% confidence intervals based on the observed Fisher information matrix. For all cases, we utilized 1000 replicates to find different estimates of each scheme. For comparison purposes, we also computed the Bayes estimates and 95% credible intervals relying on 10000 MCMC samples and then ignore the first 1000 values as “burn-in”. We determine the average maximum likelihood estimates (AMLEs), average Bayesian estimates (ABEs), mean squared errors (MSEs), coverage percentages (CPs) and average confidence interval lengths (ACILs). when the hyperparameters are 0, we used the non-informative gamma priors for the two shape parameters. That means, the prior 0: $a = b = c = d = 0$. In contrast, the informative prior includes prior 1, $a = 1, b = 2, c = 2$ and $d = 1$, when $\alpha = 2$ and $\beta = 0.5$. We have selected hyperparameters in the informative prior distributions in such away that the mean of the priors equal to the parameters, nominal values. Our results are represented in Tables 2 and 3.

Table 3: The 95% average confidence interval lengths and the corresponding coverage percentages of $H(f)$ for different censoring schemes are reported, when $\alpha = 0.5$, $\beta = 2.0$ and $H(f) = -0.6931$.

| Scheme | MLE | | (Bayes-MCMC prior 0) | | (Bayes-MCMC prior 1) | |
|---|---|---|---|---|---|---|
| | ACILs | CPs | ACILs | CPs | ACILs | CPs |
| $S^{(1)}_{1:20:50}$ | 1.1024 | 0.868 | 1.2124 | 0.877 | 1.0327 | 0.887 |
| $S^{(2)}_{1:20:50}$ | 0.9943 | 0.849 | 0.9637 | 0.899 | 0.9221 | 0.906 |
| $S^{(3)}_{1:20:50}$ | 0.9747 | 0.830 | 0.943 | 0.841 | 0.9254 | 0.887 |
| $S^{(1)}_{1:40:50}$ | 0.7861 | 0.943 | 0.7812 | 0.941 | 0.7729 | 0.962 |
| $S^{(2)}_{1:40:50}$ | 0.7504 | 0.943 | 0.7415 | 0.952 | 0.7453 | 0.961 |
| $S^{(3)}_{1:40:50}$ | 0.7983 | 0.925 | 0.7961 | 0.956 | 0.7803 | 0.962 |
| $S^{(1)}_{1:35:70}$ | 0.9051 | 0.906 | 0.8865 | 0.911 | 0.8751 | 0.906 |
| $S^{(2)}_{1:35:70}$ | 0.8137 | 0.925 | 0.8133 | 0.942 | 0.7953 | 0.962 |
| $S^{(3)}_{1:35:70}$ | 0.7969 | 0.887 | 0.7843 | 0.915 | 0.7747 | 0.925 |
| $S^{(1)}_{1:60:70}$ | 0.6676 | 0.962 | 0.6659 | 0.943 | 0.6644 | 0.925 |
| $S^{(2)}_{1:60:70}$ | 0.6588 | 0.925 | 0.6667 | 0.958 | 0.6567 | 0.962 |
| $S^{(3)}_{1:60:70}$ | 0.6343 | 0.945 | 0.6745 | 0.964 | 0.6393 | 0.970 |
| $S^{(1)}_{4:20:50}$ | 1.2543 | 0.875 | 1.1724 | 0.893 | 0.9561 | 0.903 |
| $S^{(2)}_{4:20:50}$ | 1.2606 | 0.849 | 1.1130 | 0.867 | 0.9741 | 0.906 |
| $S^{(3)}_{4:20:50}$ | 1.8307 | 0.832 | 1.3210 | 0.899 | 1.0233 | 0.911 |
| $S^{(1)}_{4:40:50}$ | 0.8273 | 0.925 | 0.7643 | 0.953 | 0.7677 | 0.962 |
| $S^{(2)}_{4:40:50}$ | 0.8867 | 0.925 | 0.7837 | 0.956 | 0.7746 | 0.962 |
| $S^{(3)}_{4:40:50}$ | 0.9165 | 0.936 | 0.7775 | 0.973 | 0.7794 | 0.971 |
| $S^{(1)}_{4:35:70}$ | 0.9128 | 0.907 | 0.7989 | 0.937 | 0.7836 | 0.943 |
| $S^{(2)}_{4:35:70}$ | 1.0464 | 0.938 | 0.8765 | 0.952 | 0.8514 | 0.943 |
| $S^{(3)}_{4:35:70}$ | 0.9987 | 0.943 | 0.8922 | 0.955 | 0.8740 | 0.948 |
| $S^{(1)}_{4:60:70}$ | 0.6645 | 0.946 | 0.6432 | 0.969 | 0.6372 | 0.971 |
| $S^{(2)}_{4:60:70}$ | 0.7307 | 0.962 | 0.6836 | 0.959 | 0.6781 | 0.968 |
| $S^{(3)}_{4:60:70}$ | 0.7461 | 0.955 | 0.6777 | 0.971 | 0.6837 | 0.965 |

# 6 Conclusions

The statistical inferences for the information entropy of Kumaraswamy distribution using PFFC samples were investigated in this paper. The maximum likelikood estimation and Bayesian estimation are investigated. For Bayesian estimation, we apply MCMC method to approximate the Bayesian estimates under square error loss function. We construct the approximate intervals based on MLEs . In addition, we use the MCMC method to derive the credible intervals. Based on the results of the simulation study some of the points are clear from this experiment. We observe the following:

**(i)** From the results obtained in Tables 2-3, the MSEs and ACILs of entropy all significantly decrease as the sample size $m$ increases.

**(ii)** It was also found that different censoring schemes do have impacts on the estimated results, among which Scheme 1 performed best.

**(iii)** Bayesian estimations of entropy with the informative prior and censoring Scheme 1 were good estimates based on the MSEs and ACILs.

**(iv)** Form Table 3 all empirical results are close to confidence level 0.95, when the effective sample $m$ increases

**(v)** As expected the Bayes estimates under the non-informative prior, and the MLEs are quite close to each other. Moreover, Bayes estimators based on informative prior perform better than the MLEs, in terms of both MSEs and ACILs.

# References


[1] N. Balakrishnan and R. Aggarwala, Progressive Censoring –Theory, Methods, and Applications, Birkhäauser, Basel, Switzerland, 2000.

[2] C. H. Jun, S. Balamurali and S. H. Lee, "Variables sampling plans forWeibull distributed lifetimes under sudden death testing," IEEE Transactions on Reliability, vol. 55, no. 1, pp. 53-58, 2006.

[3] A, J, Fernández, C. J. Pérez-González, M. Aslam and C. H. Jun, "Design of progressively censored group sampling plans for Weibull distributions: An optimization problem," *European Journal of Operational Research,* vol. 311, no. 3, pp. 525–532, 2011.

[4] S.-J. Wu and C. Kuş, "On estimation based on progressive first-failure-censored sampling," *Computational Statistics and Data Analysis.* vol. 53 no. 10, pp, 3659-3670, 2009.

[5] A. A. Soliman, A. H Abd Ellah, N. A. Abou-Elheggag and A. A. Modhesh, "Bayesian and non-Bayesian inferences of the Burr-XII distribution for progressive first-failure censored data,". *Metron*, vol. 70, no. 1, pp. 1–25, 2012.

[6] A. A. Modhesh, "Analysis of Lifetime Performance Index with Compound Rayleigh Distribution in Presence of Progressive First-Failure-Censoring Scheme," *Journal of Statistics Applications & Probability*,vol. 6, no. 3, pp. 459-470, 2017.

[7] 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. 1-21, 2019.

[8] A. Chaturvedi, R. Garg and S. Saini, "Estimation and testing procedures for the reliability characteristics of Kumaraswamy-G distributions based on the progressively first failure censored samples," *OPSEARCH*, https://doi.org/10.1007/s12597-021-00523-7, 2021.

[9] K. Kumar, H. Krishna and R. Garg. "Estimation of P(Y < X) in Lindley distribution using progressively first failure censoring," *International Journal of System Assurance Engineering and Managemen*t, vol. 6, pp. 330–341, 2015.

[10] S. Saini, A. Chaturvedi and R. Garg, "Estimation of stress–strength reliability for generalized Maxwell failure distribution under progressive first failure censoring," *Journal of Statistical Computation and Simulation,* vol. 91, no. 7, 1366-1393, 2021.

[11] M. C. Jones, "Kumaraswamy's distribution: a beta-type distribution with some tractability advantages. Statistical," *Methodology*, vol. 6, no. 1, pp. 70-91, 2009.

[12] G. M. Cordeiro, E .M. M. Ortega and S. Nadarajah, "The Kumaraswamy Weibull distribution with application to failure data," *Journal of The Franklin Institute*, vol. 347, no. 8, pp. 1399-1429, 2010.

[13] Cordeiro, G. M., Nadarajah, S. and E. M. M. Ortega, "The Kumaraswamy Gumbel distribution," *Statistical Methods & Applications*, vol. 21, no. 2, pp. 139-168, 2012.

[14] J. V. Michalowicz, J. M. Nichols and F. Bucholtz, Handbook of Differential Entropy, Chapman and Hall_CRC Press), 2013.

[15] A. I. Ishaq, A. Usman, T. Musa and S. Agboola, "On some properties of Generalized Transmuted Kumaraswamy distribution," *Pakistan Journal of Statistics and Operation Research*; Lahore vol. 15, no 3, pp. 577-586, 2019.

[16] 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, https://doi.org/10.3390/math9040313, (2021).

[17] 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.

[18] G. Q. Zhao, W. Liang and S. Y. HEE, "Empirical entropy for right censored data," Acta *Mathematicae Applicatae Sinica, English Series*, vol. 31, no. 2, pp. 395-404, 2015.

[19] N. Balakrishnan and M. Kateri, "On the maximum likelihood estimation of parameters of Weibull distribution based on complete and censored data," *Statistics and Probability Letters,* vol. 78, no. 17, pp. 2971 2975, 2008.

[20] 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, 313, 2021.

[21] X. Shi, Y. Shi and K. Zhou, "Estimation for Entropy and Parameters of Generalized Bilal Distribution under Adaptive Type II Progressive Hybrid Censoring Scheme," *Entropy,* vol. 23, no. 2, 206, 2021.

[22] S. Chib and E.Greenberg, "Understanding the Metropolis-Hastings algorithm," *The American Statistician,* vol. 49, no. 4, pp. 327-335, 1995.

[23] D. G. Hoel, "A representation of mortality data by competing risks," *Biometrics*, vol. 28, no. 2, pp. 475-488, 1972.

[24] N. Balakrishnan and R. A. Sandhu,A simple "Simulation algorithm for generating progressively Type-II censored samples," T*he American Statistician*, vol. 49, no. 2, pp. 229-230, 1995.