Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/130109

Estimation of Kumaraswamy Distribution Parameters Using the Principle of Maximum Entropy

*M. R. Mahmoud and A. M. Saad**

Department of Mathematical Statistics, Faculty of Graduate Studies for Statistical Research, Cairo University, Cairo, Egypt

Received: 22 Mar. 2023, Revised: 2 Jun. 2023, Accepted: 7 Jul. 2023
Published online: 1 Jan. 2024

Abstract: This paper proposes using maximum entropy approach to estimate the parameters of the Kumaraswamy distribution subject to moment constraints. Kumaraswamy [7] introduced the double pounded probability density function which was originally used to model hydrological phenomena. It was mentioned that this probability density function is applicable to bounded natural phenomena which have values on two sides. The distribution share several properties with the beta distribution and it has the extra advantages that is possesses a closed form distribution function, but it remained unknown to most statisticians until it was developed by Jones [6] as a beta-type distribution with some tractability advantages in particular as it has fairly simple quantile function and it has explicit formula for L-Moment. Using the principle of maximum entropy to propose new estimators for the Kumaraswamy parameters and compared with maximum likelihood and Bayesian estimation methods. A simulation study is performed to investigate the performance of the estimators in terms of their mean square errors and their efficiency.

Keywords: Principle of maximum entropy, Maximum Likelihood, Bayes, Kumaraswamy distribution.

1 Introduction

In 1948, the concept of entropy as an uncertainty measure was developed by Shannon [12]. Jaynes [5] utilized Shannon formula for the entropy to suggest the estimation of the unknown density by maximizing the entropy using the exact amount of information provided by the data. The maximum entropy approach for estimating density function is considered by many to be a very powerful tool for approximating the density function (such as, Miller and Horn [9], Hall and Presnell [3], Wu [22], Phillips et al [11] and Dudik et al [2]).

Singh et al [17] and Singh and Rajagopal [18] showed that according to principle of maximum entropy (POME) if we have reasons to assume the mathematical formula for the density function except for some unknown parameters. Then we can use the maximum entropy principle to estimate the parameters of this distribution utilizing the information available in the data. They suggested a way of utilizing the data together with POME to estimate the parameters and offered an approach and used to derive a number of distributions in the analysis of hydrological data compared with maximum likelihood estimation method, and they showed that their approach is good the compared method.

Singh and Chowdhury [14] used the POME for estimating the Gamma distribution parameters. Singh and Guo [15] and [16] used the POME for estimating the three and two parameters generalized Pareto distribution.

A lot of natural phenomena either they have upper and lower bounds or that we are only interested in the double bounded observable part. In hydrological journal, Kumaraswamy [7] introduced a distribution to describe double bounded random phenomena with application in hydrology. This kept this distribution away from the attention until it was brought to the attention by Jones [6]. The distribution shares several properties with the beta distribution but it has extra advantage that is possess a closed form cumulative distribution function. Furthermore, it has fairly simple quantile function which lead to a simple formula for generating random samples from the Kumaraswamy distribution. Also it has explicit formulas for the L-Moments (see jones [6]). Mitnik [10] obtained the distribution form moments and proved the closeness under exponentiation and linear transformation.

Sundar [19] applied the Kumaraswamy probability density function to describe the ocean wave statistics, and estimate the most probable maximum wave height, a new generalized distributions family described by Cordeiro and Castro [1] named as the Kw-inverse Gaussian, Kw-Weibull , Kw-gamma, Kw-normal, and Kw-Gumbel distribution. Hussian [4] performed estimation with respect to ranked set sampling for the Kumaraswamy distribution using Bayesian technique and maximum

*Corresponding author e-mail: ahlam_stat@yahoo.com

likelihood function. Tahir et al [20] constructed many families of distributions through a new flexible generalized family, and investigated the properties of a new flexible Kumaraswamy (NFKw) as a new model distribution, and estimate the model parameters by maximum-likelihood, then applied NFKw model for three data sets.

This paper proposes the approach of principle of maximum entropy, also follows Singh approach to estimate the two parameter Kumaraswamy distribution based on simple random sample. In order to check the behavior efficiency of estimators based on POME, Monte Carlo simulation has been conducted compared with the maximum likelihood and Bayes methods.

The present paper is organized as follows. In section 2, we propose the principle of maximum entropy approach. In section 3, derivation of parameter estimation by POME. In section 4, we consider estimation of Kumaraswamy distribution parameters. In section 5, simulation study. Finally, we summarize our results in section 6.

2. The Principle of Maximum Entropy (POME)

The Shannon entropy of a continuous distribution with the density function $f(x;\theta)$ given by the function $H(x)$ expressed as;

$$H(x) = -\int_{-\infty}^{\infty} f(x;\theta)\, ln\, f(x;\theta)\, dx \tag{1}$$

where $H(x)$ is the $f(x;\ \theta)$ entropy function, and it considered as an average of $-ln f(x;\ \theta)$.

The maximum entropy principle, introduced by Jaynes [5], as a rational approach for making inference based on current state of knowledge, states, in short, that the probability distribution that describes a system is that one with the largest entropy subject to given information. Since must use the maximum Shannon's entropy probability distribution, consistent to just known information and whatever constraints, so we choose the Shannon's maximum entropy function. The distribution $H(x)$ is referred as MaxEnt distribution. In analyzing the maximum entropy problem, Verdogo Lazo and Rathie [21] relied on the fact that if,

$$f(x) = exp\{\lambda_0 + \lambda_1 g_1(x) + \cdots + \lambda_n g_n(x)\}$$

then

$$H(x) = -\{\lambda_0 + \lambda_1 E g_1(X) + \cdots + \lambda_n E g_n(X)\}$$

Where

$$E g_i(X) = \int_{-\infty}^{\infty} g_i(X) f(x)\, dx$$

and constructed a table of deferential entropies for several continuous probability distributions.

3. Derivation of Parameter Estimation by POME

According to Singh el al. [17] "The POME method involves population expectations, whereas the MLE method involves sample averages. If population is replaced by a sample then the two methods would yield the same parameter estimates", and they derived some univariate distributions as a direct consequence of POME given a set of m constraints C_i as,

$$c_i = \int_{-\infty}^{\infty} g_i(x)\ f(x)\, dx, i = 1, 2, \dots, m$$

where $g_i(x)$ are some functions whose averages over $f(x)$ are specified, then the maximum of $H(x)$ in (1) subject to the conditions of c_i is given by,

$$f(x) = \exp\left[-\lambda_0 - \sum_{i=1}^{m} \lambda_i \, g_i(x)\right]$$

where $\lambda_i \; i = 0, 1, \ldots, m$, are the Lagrange multipliers which are determined in terms of c_i . Inserting $f(x)$ in the definition of total probability,

$$\int_{-\infty}^{\infty} \exp\left[-\lambda_0 - \sum_{i=1}^{m} \lambda_i \, g_i(x)\right] dx = 1$$

which leads to,

$$\lambda_0 = \ln \int_{-\infty}^{\infty} \exp\left\{-\sum_{i=1}^{m} \lambda_i \, g_i(x)\right\}$$

Then multipliers of Lagrange according to c_i by,

$$-\frac{\partial \lambda_0}{\partial \lambda_i} = c_i \tag{2}$$

and also,

$$\frac{\partial^2 \lambda_0}{\partial {\lambda_i}^2} = var\,[g_i(x)]; \tag{3}$$

$$\frac{\partial^2 \lambda_0}{\partial \lambda_i \partial \, \lambda_j} = cov\,\left[g_i(x),\; g_j(x)\right], \qquad i \neq j$$

With estimated multipliers of Lagrange from (2) and (3), the pdf in $f(x)$ is uniquely defined. Clearly, this procedure can be applied to derive any probability density function for which appropriate constraints can be found. Following this procedure, Singh et al. ([14], [17], [18], [15], [16], [13]) have derived a number of distributions used in hydrology as; Gamma, two and three Pareto parameters. Kumphon [8] also estimated the unknown Kappa distribution parameters using maximum likelihood also maximum entropy.

4. Estimation of Kumaraswamy Distribution Parameters

Kumaraswamy [7] introduced a double bounded continuous probability family of distributions takes values on [0,1] interval, denoted by $Kw(\alpha, \beta)$ distribution and has cumulative distribution function of the form,

$$F_{Kw}(x; \alpha, \beta) = 1 - (1 - x^{\alpha})^{\beta}$$

and probability density function given by,

$$f_{Kw}(x; \alpha, \beta) = \alpha \, \beta \, x^{\alpha-1} \, (1 - x^{\alpha})^{\beta-1}, \;\; 0 \leq x \leq 1 \;\; and \;\; \alpha, \beta > 0 \tag{4}$$

where α and β are shape parameters.

We used Kumaraswamy distribution in equation (4) to introduce a new POME estimator compared with the maximum likelihood and Bayesian methods of estimation.

4.1 Maximum Likelihood Estimation

Let $x_1 \dots x_n$ be a Kw distribution sample, with α and β shape parameters. The shape parameters likelihood function for the observed samples is

$$L_s(\alpha,\beta|\boldsymbol{X}) = \alpha^n\beta^n \prod_{j=1}^{n} x_j^{\alpha-1} \prod_{j=1}^{n}(1-x_j^{\alpha})^{\beta-1}$$

and the log-likelihood function for α, β parameters will be

$$logL(\alpha,\beta|\boldsymbol{X}) = n\log\alpha + n\log\beta + (\alpha-1)\sum_{j=1}^{n}\log x_j + (\beta-1)\sum_{j=1}^{n}\log\left(1-x_j^{\alpha}\right)$$

The estimators of the parameters denoted by $\hat{\alpha}$ and $\hat{\beta}$ obtained by likelihood solution of the following equations

$$\frac{n}{\alpha} + \sum_{j=1}^{n}\log x_j - (\beta-1)\sum_{j=1}^{n}\frac{x_j^{\alpha}\log\alpha}{\left(1-x_j^{\alpha}\right)} = 0$$

$$\frac{n}{\beta} + \sum_{j=1}^{n}\log\left(1-x_j^{\alpha}\right) = 0$$

So that we have,

$$\hat{\beta}_l = \frac{-n}{\sum_{j=1}^{n}\log\left(1-x_j^{\hat{\alpha}}\right)}$$

$$\frac{n}{\hat{\alpha}} + \sum_{j=1}^{n}\log x_j - \left(\hat{\beta}-1\right)\log\hat{\alpha}\sum_{j=1}^{n}\frac{x_j^{\hat{\alpha}}}{\left(1-x_j^{\hat{\alpha}}\right)} = 0$$

Numerical methods are used to give solution for the equations and their properties.

4.2 Bayesian Estimation

The shape parameters estimators of α and β, using Bayesian method, as α and β have prior distributions $Gamma(a_1, b_1)$ and $Gamma(a_2, b_2)$ where the probability density function denoted by,

$$\pi_1(\alpha) = \frac{{b_1}^{a_1}}{\Gamma(a_1)}\alpha^{a_1-1}\, e^{-b_1\alpha}$$

$$\pi_2(\beta) = \frac{{b_2}^{a_2}}{\Gamma(a_2)}\beta^{a_2-1}\, e^{-b_2\beta}$$

Where (a_1, b_1) and (a_2, b_2) are assumed to be known. Based on the above likelihood function obtained by $L_s(\alpha, \beta|\boldsymbol{X})$ and the prior form, then α and β joint density could be denoted by,

$$L_s(\alpha, \beta|\boldsymbol{X}) = L(\alpha, \beta/X)\pi_1(\alpha)\pi_2(\beta) = K_1 Z$$

where K_1 is constant and

$$Z = \alpha^{n+a_1-1}\,\beta^{n+a_2-1} exp\left[-b_1\alpha - b_2\beta + (\alpha-1)\sum_{j=1}^{n} log\, x_j + (\beta-1)\sum_{j=1}^{n} log\left(1 - x_j^{\alpha}\right)\right].$$

So that, the α and β joint posterior density function derived by,

$$\pi_s(\alpha, \beta|\boldsymbol{X}) = \frac{Z}{\int_0^{\infty}\int_0^{\infty} Z\, d\alpha\, d\beta}$$

Therefore, the posterior density of α and β respectively are,

$$\pi_{\alpha}(\alpha|\boldsymbol{X}) = \frac{\int_0^{\infty} Z\, d\beta}{\int_0^{\infty}\int_0^{\infty} Z\, d\alpha\, d\beta}$$

$$\pi_{\beta}(\beta|\boldsymbol{X}) = \frac{\int_0^{\infty} Z\, d\alpha}{\int_0^{\infty}\int_0^{\infty} Z\, d\alpha\, d\beta}$$

So that, the estimators for α and β by using Bayes technique are defined as $\hat{\alpha}_B$, $\hat{\beta}_B$.

$$\hat{\alpha}_B = E(\alpha|\boldsymbol{X}) = \frac{\int_0^{\infty}\int_0^{\infty} \alpha\, Z\, d\beta\ d\alpha}{\int_0^{\infty}\int_0^{\infty} Z\, d\alpha\, d\beta} \quad (5)$$

$$\hat{\beta}_B = E(\beta|\boldsymbol{X}) = \frac{\int_0^{\infty}\int_0^{\infty} \beta\, Z\, d\alpha\ d\beta}{\int_0^{\infty}\int_0^{\infty} Z\, d\alpha\, d\beta} \quad (6)$$

Simulation studies will be used to study the properties of these estimators.

4.3 Principle of Maximum Entropy

The Kumaraswamy distribution entropy obtained through three steps; First, the specification of constraints. Second, constructions the function form of the entropy. Finally, finding how constraints are related with the distribution parameters.

4.3.1 Conducting constraints.

The Kumaraswamy distribution entropy derived by inserting the probability density in equation (4) into (1), such that Shannon entropy can be expressed as,

$$H[f(x)] = -\int_0^1 f(x;\theta)[ln\, f(x;\theta)]dx$$

$H(f)$ could be considered as the mean of $-\ln\ f(x)$.

$$H(f) = -\int_0^1 \{\ln \alpha\beta f(x) + (\alpha-1)\ln x\, f(x) + (\beta-1)\ln(1-x^\alpha)f(x)\}dx$$

so,

$$H(f) = -\,ln\,\alpha\beta\int_0^1 f(x)\,dx - (\alpha-1)\int_0^1 ln\,x\,f(x)\,dx\ - (\beta-1)\int_0^1 ln(1-x^\alpha)f(x)dx$$

and,

$$c_i = \int_{-\infty}^{\infty} g_i(x)\ f(x)\,dx = E[\,g_i(x)],\quad i = 1,2,\dots,m$$

So that Kumaraswamy appropriate constraints are;

1. $\int_0^1 f(x)dx = 1$
2. $\int_0^1 \ln x\, f(x)dx = E\ln x$
3. $\int_0^1 \ln(1-x^\alpha)f(x)dx = E\ln(1-x^\alpha)$

4.3.2 Entropy function construction

Kumaraswamy distribution PDF corresponding to equation (4) and consistent with POME could be written as:

$$f(x) = exp\left[-\lambda_0 - \sum_{i=1}^{m}\lambda_i\, g_i(x)\right]$$

$$f(x) = exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(1-x^\alpha)] \qquad (7)$$

where λ_0 and λ_1, λ_2 are Lagrange multipliers. Then using $\int_0^1 f(x)dx = 1$, so

$$\int_0^1 exp[-\lambda_0 - \lambda_1 \ln x - \lambda_2 \ln(1 - x^\alpha)]dx = 1$$

such that,

$$exp(\lambda_0) = \int_0^1 exp[-\lambda_1 \ln x - \lambda_2 \ln(1 - x^\alpha)]\, dx \tag{8}$$

$$exp(\lambda_0) = \int_0^1 x^{-\lambda_1}\, (1 - x^\alpha)^{-\lambda_2} \tag{9}$$

$$exp(\lambda_0) = \frac{1}{(1 - \lambda_1)(1 - \lambda_2)} \tag{10}$$

The zeroth multiplier takes the form,

$$\lambda_0 = ln\left[\frac{1}{(1 - \lambda_1)(1 - \lambda_2)}\right] \tag{11}$$

Using $exp(\lambda_0)$ as in equation (10) in (7) identify,

$$f(x) = exp(-\lambda_0)\, x^{-\lambda_1}(1 - x^\alpha)^{-\lambda_2}$$

$$f(x) = (1 - \lambda_1)(1 - \lambda_2)x^{-\lambda_1}(1 - x^\alpha)^{-\lambda_2} \tag{12}$$

Comparing with Kumaraswamy pdf in (4) we get; $(1 - \lambda_1) = \alpha$ and $(1 - \lambda_2) = \beta$

Using logarithm of $f(x)$ in (12) leads to,

$$\ln f(x) = \ln(1 - \lambda_1)(1 - \lambda_2) - \lambda_1 \ln x \; - \lambda_2 \ln(1 - x^\alpha)$$

Therefore, the entropy $H(f)$ of the Kumaraswamy distribution using Lagrange multipliers follows:

$$H(f) = -\int_0^1 f(x)\, ln\, f(x)dx$$

$$\begin{aligned} H(f) &= -\int_0^1 [ln(1 - \lambda_1)(1 - \lambda_2) - \lambda_1\, ln\, x \; - \lambda_2\, ln(1 - x^\alpha)]\, f(x)dx \\ &= -ln(1 - \lambda_1)(1 - \lambda_2) + \lambda_1 E(\, ln\, x) \; + \lambda_2\, E\, ln(1 - x^\alpha) \end{aligned} \tag{13}$$

4.3.3 Finding the distribution parameters relation with the constraints equations

Following Singh et al [17] derived the relation between constraints and shape parameters by differentiating the equation of entropy denoted by $H(f)$, then equating these derivatives to zero, also using the constraints depending on distribution parameters and Lagrange multipliers,. Taking partial differentiation of (13) depending on λ_1 and λ_2 and zero equating the differentiation yields,

$$\frac{\partial H}{\partial \lambda_1} = \frac{1}{1-\lambda_1} + E(ln\,x) = 0 \tag{14}$$

$$\frac{\partial H}{\partial \lambda_2} = \frac{1}{1-\lambda_2} + E\;ln(1-x^{\alpha}) = 0 \tag{15}$$

equations (14) to (15) yields, respectively,

$$-\mathrm{E}(\ln x) = \frac{1}{1-\lambda_1} \tag{16}$$

$$-E\,\ln(1-x^{\alpha}) = \frac{1}{1-\lambda_2} \tag{17}$$

Equations (16) and (17) are the POME estimation equations. Alternatively, the equation form (16) and (17) could be conducted by differentiating the zeroth Lagrange coefficient with regard to λ_1 and λ_2, then equating the derivative to zero. Equation (8) could be identified by,

$$\lambda_0 = \ln \int_0^1 exp[-\lambda_1 \ln x - \lambda_2 \ln(1-x^{\alpha})]\,dx,$$

$$\lambda_0 = \ln \int_0^1 x^{-\lambda_1}\,(1-x^{\alpha})^{-\lambda_2}\,dx \tag{18}$$

Taking the derivative of (18) respect to λ_1 and λ_2:

$$\frac{\partial \lambda_0}{\partial \lambda_1} = \frac{\int_0^1 x^{-\lambda_1}\,(1-x^{\alpha})^{-\lambda_2} \ln x^{-1}\,dx}{\int_0^1 x^{-\lambda_1}\,(1-x^{\alpha})^{-\lambda_2} dx}$$

$$\frac{\partial \lambda_0}{\partial \lambda_1} = E\;(\ln x^{-1}) = -E \ln x \tag{19}$$

and,

$$\frac{\partial\lambda_0}{\partial{\lambda_1}^2} = var \ln x \tag{20}$$

Also,

$$\frac{\partial\lambda_0}{\partial\lambda_2} = \frac{\int_0^1 x^{-\lambda_1} (1-x^\alpha)^{-\lambda_2} \, ln(1-x^\alpha)^{-1} \, dx}{\int_0^1 x^{-\lambda_1} (1-x^\alpha)^{-\lambda_2} dx}$$

$$\frac{\partial\lambda_0}{\partial\lambda_2} = -E \, ln(1-x^\alpha) \tag{21}$$

$$\frac{\partial\lambda_0}{\partial{\lambda_2}^2} = var \, ln(1-x^\alpha) \tag{22}$$

From equation (11),

$$\lambda_0 = -\ln(1-\lambda_1) - \ln(1-\lambda_2)$$

$$\frac{\partial\lambda_0}{\partial\lambda_1} = \frac{1}{1-\lambda_1} \tag{23}$$

and,

$$\frac{\partial\lambda_0}{\partial{\lambda_1}^2} = \frac{1}{(1-\lambda_1)^2} \tag{24}$$

$$\frac{\partial\lambda_0}{\partial\lambda_2} = \frac{1}{1-\lambda_2} \tag{25}$$

and

$$\frac{\partial\lambda_0}{\partial{\lambda_2}^2} = \frac{1}{(1-\lambda_2)^2} \tag{26}$$

Equating equation (19, 23), (20, 24), (21, 25) and (22, 26) leads to:

$$E\{\ln x\} = -\frac{1}{1-\lambda_1} \tag{27}$$

$$E\{\ln(1-x^{\alpha})\} = -\frac{1}{1-\lambda_2} \tag{28}$$

and,

$$var\ln x = \frac{1}{(1-\lambda_1)^2} \tag{29}$$

$$var\ln(1-x^{\alpha}) = \frac{1}{(1-\lambda_2)^2} \tag{30}$$

which is the same as equation (16, 17).

Using $(1-\lambda_1) = \alpha$ and $(1-\lambda_2) = \beta$ and the equations from (27) to (30) one gets,

$$\frac{E[\ln x]}{var\ln x} = \hat{\alpha} \tag{31}$$

$$\frac{E\{\ln(1-x^{\hat{\alpha}})\}}{var\ln(1-x^{\hat{\alpha}})} = \hat{\beta} \tag{32}$$

Equations (31) and (32) are the POME-based estimation equations.

5. Simulation Study

Since it is not possible to assess the estimation process through the analysis of real life data, it was decided to run a simulation study for this purpose. Simulation studies are performed on a large number generated data sets and the results of them are summarized. Safely and efficiently results of simulation studies are compared to actual parameters values used in the generation of data and are therefore verified easily for each group of parameters values and each sample size considered (low, moderate and large). 10000 samples were generated and analyzed estimates of the parameters were obtained via numerical solutions of the maximizing equations of the likelihood function.

Using simulated data with 10000 repetitions and different values of α and β to obtain principle of maximum entropy, maximum likelihood and Bayes estimates of the distribution parameters of the Kumaraswamy also the estimators performance are compared based on several values of the parameters $\alpha = 0.5$ and 2 and $\beta = 0.5, 1.5$ and 3.5. The comparison is conducted through the MSEs and the efficiency of $\hat{\alpha}$ and $\hat{\beta}$. The results are reported in Tables 1, 2, 3 and 4 in the following section.

6. Results and Conclusions

The problem of unknown parameters estimation based on POME is considered in this paper for Kumaraswamy distribution. For the comparison purpose, Bayesian and Maximum Likelihood estimation methods are used. It is observed from the simulation study, when choosing $\alpha = 2$ we found that in many cases the POME estimates perform better than the Bayes and ML estimates relative to their MSE, and the MSEs decreased when sample size increase. Based on $\alpha = 0.5$ it is observed that in all cases the POME estimates relative to MSE and efficiency can be better than Bayes estimates, and the MSE decreased when sample size increase, so the unknown parameters estimates efficiency of the Kumaraswamy distribution under POME approach increased with the sample size.

The results are reported in the following tables.

Table 1: MSEs for Kw distribution estimators for population with $\alpha = 2$, and the prior hyper-parameter $(a_1, a_2, b_1, b_2) = (2, 2, 3, 3)$.

n	β		POME	MSE	MLE	MSE	Bayes	MSE
20	0.5	$\hat{\alpha}$	2.4212	1.8255	1.0563	1.0122	0.4167	2.5121
		$\hat{\beta}$	0.6385	0.1060	0.4176	0.0257	0.4803	0.0094
	1.5	$\hat{\alpha}$	2.766	2.093	1.418	0.3926	0.5316	2.1635
		$\hat{\beta}$	1.921	0.9854	1.149	0.1347	0.6328	0.0185
	3.5	$\hat{\alpha}$	3.434	4.2635	1.0189	0.9694	1.4223	0.3617
		$\hat{\beta}$	4.4992	5.362	1.6194	3.5572	1.9964	2.3339
30	0.5	$\hat{\alpha}$	2.1859	0.9358	1.0127	1.0189	0.4069	2.5413
		$\hat{\beta}$	0.5951	0.0606	0.4008	0.0177	0.4691	0.0068
	1.5	$\hat{\alpha}$	2.5774	1.2237	1.4039	0.3911	0.8873	1.2493
		$\hat{\beta}$	1.7811	0.5461	1.1358	0.1406	1.1577	0.1447
	3.5	$\hat{\alpha}$	3.2278	2.7939	1.0152	0.9743	1.4694	0.3043
		$\hat{\beta}$	4.1365	2.8456	1.6108	3.5831	2.1193	1.9735
50	0.5	$\hat{\alpha}$	1.9880	0.4623	0.9941	1.0150	0.4003	2.5609
		$\hat{\beta}$	0.5550	0.0299	0.3933	0.0136	0.4616	0.0049
	1.5	$\hat{\alpha}$	2.4348	0.6735	1.3854	0.3972	0.8910	1.2371
		$\hat{\beta}$	1.6656	0.2735	1.1254	0.1451	1.1725	0.1257
	3.5	$\hat{\alpha}$	3.0561	1.8231	1.0135	0.9758	1.5084	0.2575
		$\hat{\beta}$	3.8916	1.514	1.6027	3.6079	2.2289	1.6672
80	0.5	$\hat{\alpha}$	1.8914	0.2769	0.9954	1.0094	0.3957	2.5735
		$\hat{\beta}$	0.5344	0.0174	0.3911	0.0131	0.4564	0.0040
	1.5	$\hat{\alpha}$	2.3623	0.4217	1.3772	0.3992	0.8916	1.2331
		$\hat{\beta}$	1.6128	0.1586	1.1188	0.1482	1.1787	0.1150
	3.5	$\hat{\alpha}$	2.9756	1.3712	1.0123	0.9772	1.5347	0.2275
		$\hat{\beta}$	3.7580	0.8544	1.6007	3.6123	2.3033	1.4704

Table 2: MSEs of the estimators of Kw distribution for population parameter $\alpha = 0.5$, and the prior hyper-parameter $(a_1, a_2, b_1, b_2) = (2, 2, 3, 3)$

n	β		POME	MSE	MLE	MSE	Bayes	MSE
20	0.5	$\hat{\alpha}$	0.6040	0.1140	0.9909	0.2524	0.6777	0.0648
		$\hat{\beta}$	0.6388	0.1101	0.6812	0.0579	0.8089	0.1561
	1.5	$\hat{\alpha}$	0.6893	0.1299	0.4462	0.0401	2.3178	3.5776
		$\hat{\beta}$	1.9206	0.9749	1.3697	0.2472	3.0323	2.7366
	3.5	$\hat{\alpha}$	0.8539	0.2609	0.3134	0.0427	4.0016	12.336
		$\hat{\beta}$	4.4815	5.3080	2.045	2.3576	4.3385	0.7196
30	0.5	$\hat{\alpha}$	0.5421	0.0578	1.0083	0.2649	0.6951	0.0621
		$\hat{\beta}$	0.5944	0.0592	0.6768	0.0483	0.8304	0.1534
	1.5	$\hat{\alpha}$	0.6433	0.0730	0.4438	0.0380	2.4872	4.2056
		$\hat{\beta}$	1.7778	0.5403	1.3622	0.2246	3.3626	3.8514
	3.5	$\hat{\alpha}$	0.8069	0.1746	0.3143	0.0411	4.4126	15.342
		$\hat{\beta}$	4.1365	2.8456	2.0411	2.3352	4.6430	1.3105
50	0.5	$\hat{\alpha}$	0.5003	0.0283	1.0179	0.2710	0.6843	0.0482
		$\hat{\beta}$	0.5562	0.3001	0.6787	0.0441	0.8179	0.1274
	1.5	$\hat{\alpha}$	0.6093	0.0405	0.4600	0.0308	2.6324	4.7679
		$\hat{\beta}$	1.6663	0.2766	1.3981	0.1774	3.6843	5.1150
	3.5	$\hat{\alpha}$	0.7629	0.1117	0.3170	0.0386	4.7096	17.727
		$\hat{\beta}$	3.8875	1.4860	2.0472	2.2712	4.8260	1.7589
80	0.5	$\hat{\alpha}$	0.4740	0.0176	1.0173	0.2692	0.6751	0.0394
		$\hat{\beta}$	0.5366	0.0174	0.6761	0.0397	0.8069	0.1104
	1.5	$\hat{\alpha}$	0.5882	0.0256	0.4782	0.0221	2.7180	5.0874
		$\hat{\beta}$	1.6099	0.1570	1.4432	0.1270	3.9125	6.0988
	3.5	$\hat{\alpha}$	0.7413	0.0840	0.3235	0.0342	7.7359	53.092
		$\hat{\beta}$	3.7412	0.8521	2.0743	2.1262	9.5440	36.563

Table 3: Efficiency of the POME estimators of the Kw distribution compared with ML and Bayes estimators for population parameter $\alpha = 2$, and the prior hyper-parameter $(a_1, a_2, b_1, b_2) = (2, 2, 3, 3)$.

n	β	eff $(\hat{\alpha}_{POME,ML})$	eff $(\hat{\alpha}_{POME,Bayes})$	eff $(\hat{\beta}_{POME,ML})$	eff $(\hat{\beta}_{POME,Bayes})$
20	0.5	0.5545	1.3761	0.2429	0.0890
	1.5	0.1875	1.274	0.1366	0.1313
	3.5	0.2273	0.6634	0.0848	0.4352
30	0.5	1.0888	2.7157	0.2918	0.1122
	1.5	0.3195	1.0209	0.2576	0.2650
	3.5	0.3487	0.10893	1.2591	0.6935
50	0.5	2.1953	5.5388	0.4561	0.1660
	1.5	0.5898	1.8367	0.5307	0.4597
	3.5	0.5352	0.1412	2.3818	1.1006
80	0.5	3.6444	9.2963	0.7516	0.2333
	1.5	0.9466	2.9241	0.9342	0.7251
	3.5	0.7126	0.1659	4.2278	1.7209

Table 4: Efficiency of the POME estimators of the Kw distribution compared with ML and Bayes estimators for population parameter $\alpha = 0.5$, and the prior hyper-parameter $(a_1, a_2, b_1, b_2) = (2, 2, 3, 3)$

n	β	eff $(\hat{\alpha}_{POME,ML})$	eff $(\hat{\alpha}_{POME,Bayes})$	eff $(\hat{\beta}_{POME,ML})$	eff $(\hat{\beta}_{POME,Bayes})$
20	0.5	2.2135	0.5686	0.5260	1.4178
	1.5	0.3089	27.531	0.2535	2.8069
	3.5	0.1636	47.2734	0.4441	0.1355
30	0.5	4.5832	1.0750	0.8157	2.5904
	1.5	0.5207	57.5703	0.4158	7.1279
	3.5	0.2359	87.865	0.8206	0.4605
50	0.5	9.5466	1.7002	1.4713	4.2449
	1.5	0.7597	117.60	0.6414	18.492
	3.5	0.3459	158.574	1.5283	1.1836
80	0.5	15.271	2.2385	2.2793	6.3276
	1.5	0.8629	198.00	0.8089	38.828
	3.5	0.4075	631.79	2.4951	42.908

References

[1] G. M. Cordeiro and M. Castro. A new family of generalized distributions. *Journal of Statistical Computation and Simulation*., 81, 883-898, (2011).
[2] M. Dudik, S. J. Phillips and R. E. Schapire. Maximum entropy density estimation with generalized regularization and an application to species distribution modeling. *Journal of Machine Learning Research*., 8, 1217–1260, (2007).
[3] P. Hall and B. Presnell. Density estimation under constraints. *Journal of Computational and Graphical Statistics*., 8, 259-277, (1999).
[4] M. A. Hussian. Bayesian and maximum likelihood estimation for Kumaraswamy distribution based on ranked set sampling. *American Journal of Mathematics and Statistics*., 4, 30-37, (2014).
[5] E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*., 106, 620-630, (1957).
[6] M. C. Jones. Kumaraswamy's distribution: a beta-type distribution with some tractability Advantages. *Statistical Methodology*., 6, 70-81, (2009).
[7] P. Kumaraswamy. A generalized probability density function for double bounded random processes. *Journal of Hydrology*., 46, 79–88, (1980).
[8] B. Kumphon. Maximum entropy and maximum likelihood estimation for the three-parameter Kappa distribution. *Open Journal of Statistics*., 2, 415-419, (2012).
[9] G. Miller and D. Horn. *Maximum entropy approach to probability density estimation*, Second International Conference on Knowledge-Based Intelligent Electronic Systems. Australia, 225-230, (1998).
[10] P. Mitnik. New Properties of the Kumaraswamy distribution. *Communication in Statistics-Theory and Methods*., 42, 741–755, (2013).
[11] S. J. Phillips, R. P. Anderson and R. E. Schapire. Maximum entropy modeling of species geographic distributions. *Ecological Modeling*., 190, 231–259, (2006).
[12] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*., 27, 379–423, 623–656, (1948).
[13] V. P. Singh. The entropy theory as a tool for modeling and decision making in environmental and water resources. *Water SA*., 26, 1-11, (2000).
[14] V. P. Singh and P. K. Chowdhury. On fitting gamma distribution to synthetic runoff hydrographs. *Nordic Hydrology*.. 16, 177-192, (1985).
[15] V. P. Singh and H. Guo. Parameter estimation for 3-parameter generalized pareto distribution by the principle of maximum entropy (POME). *Hydrological Sciences -Journal- des Sciences Hydrologiques.*, 40, 165-181, (1995).
[16] V. P. Singh and H. Guo. Parameter estimation for 2-parameter generalized pareto distribution by POME. *Stochastic Hydrology and Hydraulics*., 11, 211-227, (1997).
[17] V. P. Singh, A. K. Rajagopal, and K. Singh. Derivation of some frequency distributions using the principle of maximum entropy (POME). *Water Resources*., 9, 91-106, (1986).
[18] V. P. Singh and A. K. Rajagopal. Some recent advances in the application of the principle of maximum entropy (POME) in hydrology. *Water for the Future: Hydrology in Perspective.*, 164, 353-364, (1987).
[19] V. Sundar and K. Subbiah. Application of double bounded probability density function for analysis of ocean waves. *Ocean Engenering*., 16, 193–200, (1989).
[20] M. H. Tahir, M. A. Hussain and G. M. Cordeiro. A new flexible generalized family for constructing many families of distributions. *Journal of Applied Statistics*., 49, 1615-1635, (2022).
[21] A. C. G. Verdogo Lazo and P. N. Rathie. On the Entropy of Continuous Probability Distributions. *IEEE Transactions on Information Theory IT-24*., 1, 120-122, (1978).
[22] X. Wu. Calculation of maximum entropy densities with application to income distribution. *Journal of Econometrics*., 115, 347- 354, (2003).