Research Article

Uma distribution with properties and applications

Abstract

The stochastic natures of lifetime data are really a challenge for statistician to search a suitable distribution for modeling and analysis of lifetime data. Keeping in mind the stochastic natures of lifetime data, a new lifetime distribution named Uma distribution has been suggested. Its several statistical properties, estimation of parameter and applications have been discussed. Applications of the distribution have been presented with three datasets and the goodness of fit of Uma distribution has been compared with exponential, Lindley, Shanker, Akash and Sujatha distributions.

Keywords: lifetime distributions, statistical properties, estimation of parameter, applications

Volume 11 Issue 5 - 2022

Rama Shanker

Department of Statistics, Assam University, Silchar, Assam, India

Correspondence: Rama Shanker, Department of Statistics, Assam University, Silchar, Assam, India, Email shankerrama2009@gmail.com

Received: December 09, 2022 | **Published:** December 22, 2022

Introduction

Due to stochastic nature of lifetime data, the search for lifetime distribution in the field of lifetime data analysis is expanding exponentially and getting popularity among policy makers to model data. In recent decades several lifetime distributions have been suggested in statistics literature. For example, Lindley distribution by Lindley,[1] Shanker distribution by Shanker,[2] Akash distribution by Shanker,[3] and Sujatha distribution by Shanker,[4] is some among others. Shanker et al.[5] discussed the modeling of lifetime data using exponential and Lindley distributions and observed that there are some datasets in which these two distributions do not give good fit. Further, Shanker et al.[6] put an effort to have comparative study on modeling of lifetime data using exponential, Lindley and Akash distribution and found that Akash distributions gives much better fit than both exponential and Lindley distribution but still there are some data sets in which these three distributions do not give good fit. Then, Shanker and Hagos,[7] tried to model the real lifetime datasets using exponential, Lindley, Shanker and Akash and observed that still there are some datasets in which these distributions do not give good fit. Flexibility and tractability are the two important characteristics of a lifetime distributions and if the existing distributions are not flexible or tractable for the given dataset, then the search for a new distribution starts. Sometimes, data are being transformed to satisfy some assumptions of the distribution so that distribution fits well. But this is not useful practice because the original nature of the dataset is lost. Therefore, the most preferable is to search a distribution which fits the given data well than to modify the existing distributions.

While testing the goodness of fit of some well-known one parameter lifetime distributions available in literature, it has been observed that the existing distributions do not fit the data well. In this paper, in the search for a new distribution, we propose a new distribution named Uma distribution which fits the data well over the existing distributions. The statistical properties, estimation of parameter and applications of the distribution has been presented systematically. It is hoped and expected that the distribution will draw attention of researchers to model lifetime data and preferred over the existing one parameter lifetime distributions.

Uma distribution

Taking the convex combination of exponential (θ), gamma $(2,\theta)$ and gamma $(4,\theta)$ with respective mixing proportions $\frac{\theta^3}{\theta^3+\theta^2+6}, \frac{\theta^2}{\theta^3+\theta^2+6}$ and $\frac{6}{\theta^3+\theta^2+6}$, a probability density function (pdf) can be expressed as

$$f(x;\theta)=\frac{\theta^4}{\theta^3+\theta^2+6}\left(1+x+x^3\right)e^{-\theta x};x>0,\theta>0$$

We would call this distribution as 'Uma distribution'. Since it is a convex combination of exponential and gamma distributions, it is expected to give better fit over exponential and gamma distribution and other distributions developed using convex combinations of exponential and gamma distribution. The cumulative distribution function (cdf) of Uma distribution can be obtained as

$$F(x;\theta)=1-\left[1+\frac{\theta x\left(\theta^2x^2+3\theta x+\theta^2+6\right)}{\theta^3+\theta^2+6}\right]e^{-\theta x};x>0,\theta>0$$

The behaviour of the pdf and the cdf of Uma distribution for varying values of parameter θ have been presented in Figures 1,2 respectively.

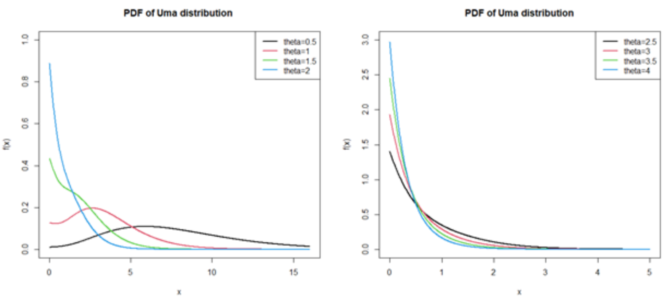

Figure 1 Graphs of the pdf of Uma distribution for selected values of the parameter.

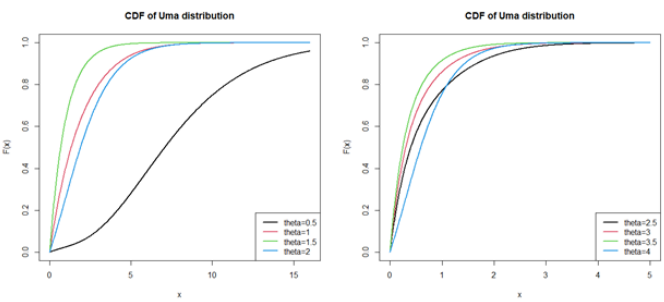

Figure 2 Graphs of the cdf of Uma distribution for selected values of the parameter.

Reliability properties

Hazard rate function

The hazard rate function of a random variable X having pdf $f(x;\theta)$ and cdf $F(x;\theta)$ is defined as

$$h(x)=\lim_{\Delta x\to 0}\frac{P(X<x+\Delta x|X>x)}{\Delta x}=\frac{f(x;\theta)}{1-F(x;\theta)}$$

Thus, the hazard rate function of Uma distribution can be obtained as

$$h(x)=\frac{\theta^4(1+x+x^3)}{\theta^3(x^3+x+1)+\theta^2(3x^2+1)+6\theta x+6}.$$

This gives $h(0)=\frac{\theta^4}{\theta^3+\theta^2+6}=f(0)$. The behaviour of the hazard rate function of Uma distribution for various values of parameter θ is shown in the following Figure 3.

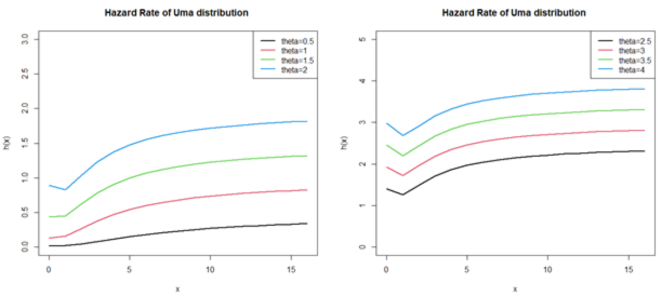

Figure 3 Graphs of the hazarad rate function of Uma distribution for selected values of the parameter.

Mean residual life function

Let X be a random variable over the support $(0,\infty)$ representing the lifetime of a system. Mean Residual life (MRL) function measures the expected value of the remaining lifetime of the system, provided it has survived up to time x. Let us consider the conditional random variable $X_x=(X-x\mid X>x);x>0$. Then, the MRL function, denoted by $m(x)$, is defined as

$$m(x)=E(X_x)=\frac{1}{S(x)}\int_x^\infty[1-F(t)]dt$$

The MRL function of Uma distribution can thus be obtained as

$$m(x)=\frac{1}{\{\theta^3(x^3+x+1)+\theta^2(3x^2+1)6\theta x+6\}e^{-\theta x}}\int_x^\infty\left[\theta^3(t^3+t+1)+\theta^2(3t^2+1)6\theta t+6\right]e^{-\theta t}dt$$

$$=\frac{\theta^3(x^3+x+1)+\theta^2(6x^2+2)+18\theta x+24}{\theta\{\theta^3(x^3+x+1)+\theta^2(3x^2+1)+6\theta x+6\}}.$$

This gives $m(0)=\frac{\theta^3+2\theta^2+24}{\theta(\theta^3+\theta^2+6)}=\mu_1'$. The behaviour of the mean residual life function of Uma distribution for various values of parameter θ is shown in the following Figure 4.

Reverse hazard rate and Mill's ratio

The reverse hazard rate of a random variable X having pdf $f(x;\theta)$ and cdf $F(x;\theta)$ is defined as

$$h_r(x)=\frac{f(x;\theta)}{F(x;\theta)}$$

Thus, the reverse hazard rate function of Uma distribution can be obtained as

$$h_r(x)=\frac{\theta^4(1+x+x^3)e^{-\theta x}}{(\theta^3+\theta^2+6)-\left[(\theta^3+\theta^2+6)+\theta x(\theta^2x^2+3\theta x+\theta^2+6)\right]e^{-\theta x}}$$

Mill's ratio of a random variable X having pdf $f(x;\theta)$ and cdf $F(x;\theta)$ is defined as

$$\text{Mill's ratio}=\frac{1}{h(x)}=\frac{1-F(x;\theta)}{f(x;\theta)}$$

Thus, the Mill's ratio of Uma distribution can be obtained as

$$\frac{1}{h(x)}=\frac{\theta^3(x^3+x+1)+\theta^2(3x^2+1)+6\theta x+6}{\theta^4(1+x+x^3)}$$

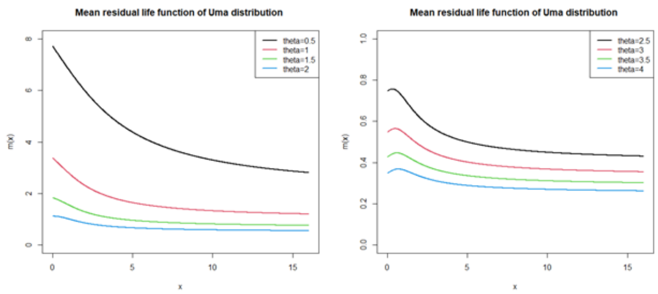

Figure 4 Graphs of the mean residual life function of Uma distribution for selected values of the parameter.

Stochastic ordering

In Probability theory and statistics, a stochastic order quantifies the concept of one random variable being bigger than another. A random variable X is said to be smaller than a random variable Y in the

i. Stochastic order $(X\le_{st}Y)$ if $F_X(x)\ge F_Y(y)$ for all x

ii. Hazard rate order $(X\le_{hr}Y)$ if $h_X(x)\ge h_Y(y)$ for all x

iii. Mean residual life order $(X\le_{mrl}Y)$ if $m_X(x)\ge m_Y(y)$ for all x

iv. Likelihood ratio order $(X\le_{lr}Y)$ if $\frac{f_X(x)}{f_Y(y)}$ decrease in x

The following results due to Shaked and Shantikumar,[8] are well known for establishing stochastic ordering of distributions

$$X<_{lr}Y\Rightarrow X<_{hr}Y\Rightarrow X<_{mrl}Y$$
$$\Downarrow$$
$$X<_{st}Y$$

Theorem: Let $X\sim$ Uma distribution (θ_1) and $Y\sim$ Uma(θ_2). If $\theta_1>\theta_2$, then $X<_{lr}Y$ hence $X<_{hr}Y$, $X<_{mrl}Y$ and $X<_{st}Y$.

Proof: We have

$$\frac{f_X(x)}{f_Y(x)}==\frac{\theta_1^4(\theta_2^3+\theta_2^2+6)}{\theta_2^4(\theta_1^3+\theta_1^2+6)}e^{-(\theta_1-\theta_2)x}$$

We have, $\log\left[\frac{f_X(x)}{f_Y(x)}\right]=\log\left[\frac{\theta_1^4(\theta_2^3+\theta_2^2+6)}{\theta_2^4(\theta_1^3+\theta_1^2+6)}\right]-(\theta_1-\theta_2)x$

Therefore, $\frac{d}{dx}\log\left[\frac{f_X(x)}{f_Y(x)}\right]=-(\theta_1-\theta_2)$

Thus, for $\theta_1>\theta_2$, $\frac{d}{dx}\log\left[\frac{f_X(x)}{f_Y(x)}\right]<0$. this means that $X<_{lr}Y$ hence $X<_{hr}Y$, $X<_{mrl}Y$ and $X<_{st}Y$.

Moments based descriptive measures

The r th moment about origin μ_r' of Uma distribution can be obtained as

$$\mu_r' = E\left(X^r\right) = \frac{\theta^4}{\theta^3+\theta^2+6}\int_0^\infty x^r\left(1+x+x^2\right)e^{-\theta x}dx$$

$$= \frac{r!\left\{\theta^3+(r+1)\theta^2+(r+1)(r+2)(r+3)\right\}}{\theta^r\left(\theta^3+\theta^2+6\right)};r=1,2,3,\cdots$$

Substituting $r=1,2,3,4$ in the above equation, the first four moments about origin of Uma distribution can be obtained as

$$\mu_1' = \frac{\theta^3+2\theta^2+24}{\theta\left(\theta^3+\theta^2+6\right)},\quad \mu_2' = \frac{2\left(\theta^3+3\theta^2+60\right)}{\theta^2\left(\theta^3+\theta^2+6\right)}$$

$$\mu_3' = \frac{6\left(\theta^3+4\theta^2+120\right)}{\theta^3\left(\theta^3+\theta^2+6\right)},\quad \mu_4' = \frac{24\left(\theta^3+5\theta^2+210\right)}{\theta^4\left(\theta^3+\theta^2+6\right)}.$$

The moments about the mean, using relationship between moments about the mean and the moments about the origin, can thus be obtained as

$$\mu_2 = \frac{\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144}{\theta^2\left(\theta^3+\theta^2+6\right)^2}$$

$$\mu_3 = \frac{2\left(\theta^9+6\theta^8+6\theta^7+200\theta^6+270\theta^5+108\theta^4+324\theta^3+432\theta^2+864\right)}{\theta^3\left(\theta^3+\theta^2+6\right)^3}$$

$$\mu_4 = \frac{3\left(\begin{array}{l}3\theta^{12}+24\theta^{11}+44\theta^{10}+968\theta^9+2336\theta^8+2016\theta^7+7488\theta^6+13248\theta^5\\+5760\theta^4+31104\theta^3+24192\theta^2+31104\end{array}\right)}{\theta^4\left(\theta^3+\theta^2+6\right)^4}.$$

The descriptive constants including coefficient of variation (CV), coefficient of skewness (CS), coefficient of kurtosis (CK) and the index of dispersion (ID) of Uma distribution are thus obtained as

$$CV = \frac{\sqrt{\mu_2}}{\mu_1'} = \frac{\sqrt{\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144}}{\theta^3+2\theta^2+24}$$

$$CS = \frac{\mu_3^2}{\mu_2^3} = \frac{4\left(\theta^9+6\theta^8+6\theta^7+200\theta^6+270\theta^5+108\theta^4+324\theta^3+432\theta^2+864\right)^2}{\left(\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144\right)^3}$$

$$CK = \frac{\mu_4}{\mu_2^2} = \frac{3\left(\begin{array}{l}3\theta^{12}+24\theta^{11}+44\theta^{10}+968\theta^9+2336\theta^8+2016\theta^7+7488\theta^6+13248\theta^5\\+5760\theta^4+31104\theta^3+24192\theta^2+31104\end{array}\right)}{\left(\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144\right)^2}$$

$$ID = \frac{\mu_2}{\mu_1'} = \frac{\theta^6+4\theta^5+2\theta^4+84\theta^3+60\theta^2+144}{\theta\left(\theta^3+\theta^2+6\right)\left(\theta^3+2\theta^2+24\right)}.$$

Behaviour of coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion for changing values of parameter are shown in the Figure 5.

Deviations from mean and median

Mean deviation about the mean and the mean deviation about median of a random variable X having pdf $f(x)$ and cdf $F(x)$ are defined by

$$\delta_1(x) = \int_0^\infty |x-\mu|\, f(x)dx = 2\mu F(\mu) - 2\int_0^\mu x\, f(x)dx$$

and $\delta_2(x) = \int_0^\infty |x-M|\, f(x)dx = -\mu + 2\int_M^\infty x\, f(x)dx$ respectively,

where $\mu = E(X)$ and $M = Median(X)$. Using pdf and expressions for the mean of Uma distribution, we get

$$\int_0^\mu x\, f(x;\theta)dx = \mu - \frac{\left[\theta^4\left(\mu^4+\mu^2+\mu\right)+\theta^3\left(4\mu^3+2\mu+1\right)+2\theta^2\left(6\mu^2+1\right)+24(\theta\mu+1)\right]e^{-\theta\mu}}{\theta\left(\theta^3+\theta^2+6\right)}$$

$$\int_0^M x\, f(x;\theta)dx = \mu - \frac{\left[\theta^4\left(M^4+M^2+M\right)+\theta^3\left(4M^3+2M+1\right)+2\theta^2\left(6M^2+1\right)+24(\theta M+1)\right]e^{-\theta M}}{\theta\left(\theta^3+\theta^2+6\right)}$$

Using above expressions some algebraic simplifications, the mean deviation about the mean $\delta_1(x)$, and the mean deviation about the median $\delta_2(x)$ of Uma distribution are obtained as

$$\delta_1(x) = \frac{2\left[\theta^3\mu^3+6\theta^2\mu^2+\theta^3\mu+18\theta\mu+\left(\theta^3+2\theta^2+24\right)\right]e^{-\theta\mu}}{\theta\left(\theta^3+\theta^2+6\right)}$$

$$\delta_2(x) = \frac{2\left[\theta^4\left(M^4+M^2+M\right)+\theta^3\left(4M^3+2M+1\right)+2\theta^2\left(6M^2+1\right)+24(\theta M+1)\right]e^{-\theta M}}{\theta\left(\theta^3+\theta^2+6\right)} - \mu$$

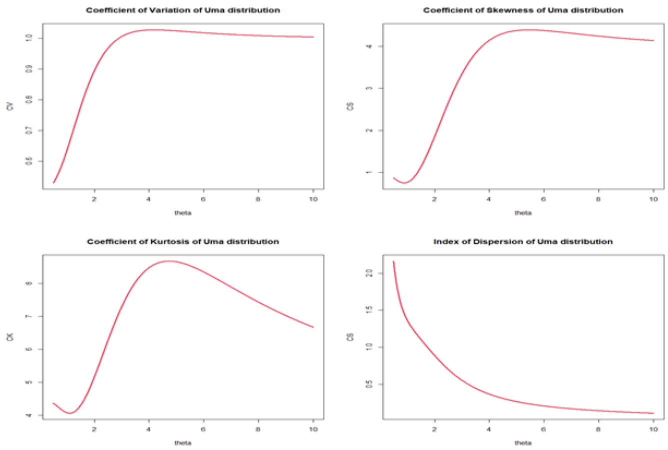

Figure 5 Graph of CV, CS, CK and ID of Uma distribution for different values of the parameter.

Parameter estimation of Uma distribution

Suppose $(x_1, x_2, x_3, \ldots, x_n)$ be a random sample of size n from Uma distribution. The log likelihood function, L of Uma distribution is given by

$$\log L = \sum_{i=1}^n \log f(x_i;\theta) = n\left\{4\log\theta - \log\left(\theta^3+\theta^2+6\right)\right\} + \sum_{i=1}^n \log\left(1+x_i+x_i^3\right) - n\theta\bar{x}$$

The maximum likelihood estimate (MLE) $(\hat{\theta})$ of the parameters (θ) of Uma distribution is the solution of the following log likelihood equation

$$\frac{d\log L}{d\theta} = \frac{4n}{\theta} - \frac{\left(3\theta^2+2\theta\right)n}{\theta^3+\theta^2+6} - n\,\bar{x} = 0$$

This gives

$$\bar{x}\theta^4 + (\bar{x}-1)\theta^3 - 2\theta^2 + 6\bar{x}\theta - 24 = 0.$$

This is a fourth degree polynomial equation in θ. It should be noted that the method of moment estimate is also the same as that of the MLE. The above equation can easily be solved using Newton-Raphson method, taking the initial value of the parameter $\theta = 0.5$.

Applications and goodness of fit

The applications and the goodness of fit of Uma distribution has been discussed with three datasets. Keeping in mind the flexibility and tractability of the distribution with the dataset following three datasets have been considered.

Data set 1: This data set represents the lifetime data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by Gross and Clark.[9]

1.1, 1.4, 1.3, 1.7, 1.9, 1.8, 1.6, 2.2, 1.7, 2.7, 4.1, 1.8, 1.5, 1.2, 1.4, 3, 1.7, 2.3, 1.6, 2.0

Data Set 2: This data set is the strength data of glass of the aircraft window reported by Fuller et al.[10]:

18.83, 20.80, 21.657, 23.03, 23.23, 24.05, 24.321, 25.5, 25.52, 25.80, 26.69, 26.77, 26.78, 27.05, 27.67, 29.90, 31.11, 33.2, 33.73, 33.76, 33.89, 34.76, 35.75, 35.91, 36.98, 37.08, 37.09, 39.58, 44.045, 45.29, 45.381

Data Set 3: The following data represent the tensile strength, measured in GPa, of 69 carbon fibers tested under tension at gauge lengths of 20 mm(Bader and Priest, 1982)[11]:

1.312, 1.314, 1.479, 1.552, 1.700, 1.803, 1.861, 1.865, 1.944, 1.958, 1.966, 1.997, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.140, 2.179, 2.224, 2.240, 2.253, 2.270, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382, 2.382, 2.426, 2.434, 2.435, 2.478, 2.490, 2.511, 2.514, 2.535, 2.554, 2.566, 2.570, 2.586, 2.629, 2.633, 2.642, 2.648, 2.684, 2.697, 2.726, 2.770, 2.773, 2.800, 2.809, 2.818, 2.821, 2.848, 2.880, 2.954, 3.012, 3.067, 3.084, 3.090, 3.096, 3.128, 3.233, 3.433, 3.585, 3.585 .

The values ML estimates of parameter, $-2\log L$, AIC (Akaike Information Criterion), AICC (Akaike Information Criterion corrected), BIC (Bayesian Information criterion), K-S (Kolmogorov-Smirnov) for the considered distributions for the given datasets have been computed and presented in Tables 1–3 respectively.

It is clear from the goodness of fit in the Tables 1 to 3 that Uma distribution gives much better fit over exponential, Lindley, Shanker, Akash and Sujatha distributions.

Table 1 ML estimates, $-2\log L$, AIC, AICC, BIC, K-S of the distribution for the dataset-1

Sl. No	Distributions	$\hat{\theta}$	$-2\log L$	AIC	AICC	BIC	K-S
1	Uma	1.6024	38.61	40.61	40.83	41.60	0.238
2	Sujatha	1.1367	57.50	59.50	59.72	60.49	0.309
3	Akash	1.1569	59.52	61.52	61.74	62.51	0.320
4	Shanker	0.8038	59.78	61.78	61.22	62.51	0.315
5	Lindley	0.8161	60.50	62.50	62.72	63.49	0.341
6	Exponential	0.5263	65.67	67.67	67.90	68.67	0.389

Table 2 ML estimates, $-2\log L$, AIC, AICC, BIC, K-S of the distributions for the dataset-2

Sl. No	Distributions	$\hat{\theta}$	$-2\log L$	AIC	AICC	BIC	K-S
1	Uma	0.1299	232.54	234.54	234.67	235.97	0.233
2	Sujatha	0.0956	241.50	243.50	243.64	244.94	0.27
3	Akash	0.0971	240.68	242.68	242.82	244.11	0.266
4	Shanker	0.0647	252.35	254.35	254.49	255.78	0.326
5	Lindley	0.0629	253.99	255.99	256.13	257.42	0.333
6	Exponential	0.0325	274.53	276.53	276.67	277.96	0.426

Table 3 ML estimates, $-2\log L$, AIC, AICC, BIC, K-S of the distributions for the dataset-3

Sl. No	Distributions	$\hat{\theta}$	$-2\log L$	AIC	AICC	BIC	K-S
1	Uma	1.3828	156.41	158.41	158.47	160.64	0.312
2	Sujatha	0.9361	221.61	223.61	223.67	225.84	0.348
3	Akash	0.9647	224.28	226.28	226.34	228.51	0.348
4	Shanker	0.6580	233.01	235.01	235.06	237.24	0.355
5	Lindley	0.6590	238.38	240.38	240.44	242.61	0.390
6	Exponential	0.4079	261.74	263.74	263.80	265.97	0.434

Conclusion and future works

A new lifetime distribution named Uma distribution has been suggested. Statistical properties, estimation of parameter and applications of the distribution has been presented. As the distribution is new one, it is expected and hoped that it will be of great use to statisticians working in the field of modeling lifetime data from different fields of knowledge.

Being a new lifetime distribution with flexibility, tractability and practicability, a lot of future works can be done on Uma distribution.

Acknowledgements

Author is really grateful to the Editor-In-Chief of the Journal and the anonymous reviewer for quick and valuable comments on the paper.

Conflicts of interests

The authors declare that there are no conflicts of interest.

Funding

None.

References

1. Lindley DV. Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B*. 1958;20:102–107.
2. Shanker R. Shanker distribution and its applications. *International Journal of Statistics and Applications*. 2015a;5(6):338–348.
3. Shanker R. Akash Distribution and its applications. *International Journal of Probability and Statistics*. 2015b;4(3):65–75.

4. Shanker R. Sujatha Distribution and its applications. *Statistics in Transition-New series*. 2016a;17(3):391–410.
5. Shanker R, Hagos F, Sujatha S. On Modeling of Lifetimes data using Exponential and Lindley distributions. *Biom Biostat Int J*. 2015;2(5):140–147.
6. Shanker R, Hagos F, Sujatha S. On modeling of lifetime data using one parameter Akash, Lindley and exponential distributions. *Biom Biostat Int J*. 2016;3(2):54–62.
7. Shanker R, Hagos F. On modeling of lifetime data using Akash, Shanker, Lindley and exponential distributions**.** *Biom & Biostat Int J*. 2016;3(6):214–224.
8. Shaked M, Shanthikumar JG. *Stochastic Orders and Their Applications*. Academic Press, New York: 1994.
9. Gross AJ, Clark VA. Survival Distributions: Reliability Applications in the Biometrical Sciences. New York: John Wiley. 1975.
10. Fuller EJ, Frieman S, Quinn J, et al. Fracture mechanics approach to the design of glass aircraft windows: A case study, SPIE Proc 1994;2286:419–430.
11. Bader MG, Priest AM. Statistical aspects of fiber and bundle strength in hybrid composites. In: Hayashi T, Kawata K, Umekawa S, editors. *Progress in Science in Engineering Composites, ICCM-IV*. Tokyo; 1982:1129–1136.