ISSN 0971-0388

Aligarh Journal of Statistics
Vol. 42 (2022), 91-104

Truncated Prakaamy distribution: Properties and Applications

Kamlesh Kumar Shukla and Rama Shanker
[Received on March, 2021. Accepted on October, 2022]

ABSTRACT

A truncated version of Prakaamy distribution has been introduced. The behaviors of its probability density function and survival function have been studied. The first two raw moments and the variance of the distribution have been given. Estimation of parameter has been discussed with maximum likelihood method. The Simulation study of proposed distribution has been presented. Goodness of fit of the proposed distribution has been illustrated with two real datasets.

1. Introduction

The statistical modeling of lifetime data using one parameter lifetime distribution and their truncated version has emerged as a fascination for recent researchers in distribution theory and related fields of knowledge. During recent decades several one parameter and two parameter lifetime distributions have been introduced in statistics. It has been observed that each new lifetime distribution introduced in statistics has advantages and disadvantages due to their theoretical and applied point of view. Further, the real lifetime data has finite range whereas each lifetime distribution has range from zero to infinity and this is completely illogical to fit a finite range real lifetime data with a distribution having lower limit zero and upper limit infinity. Due to this truncated lifetime distribution gaining momentum in statistics among researchers. Some important one parameter lifetime distributions introduced recently including Lindley distribution of Lindley (1958) and studied by Ghitany *et al.* (2008), Akash distribution introduced by Shanker (2015), Ishita distribution proposed by Shanker and Shukla (2017), Prakaamy distribution by Shukla (2018), are some among others.

✉ : Kamlesh Kumar Shukla
Email: kkshukla22@gmail.com

Extended authors information available after reference list of the article.

Truncated type of distributions, as observed by different researchers, are more useful for modeling lifetime data because its limits used as bound either upper or lower or both can be adjusted according to the range of the data. For example, truncated normal distribution proposed by Johnson *et al*. (1994) has wide application in economics and statistics. Several truncated distributions have been proposed in statistics literature in the recent years and their applications in different areas of statistics, especially in censor data such as truncated Weibull distribution of Zange and Xie (2010), truncated Lomax distribution of Aryuyuen and Bodhisuwan (2018), truncated Pareto distribution of Janinetti and Ferraro (2008), truncated Lindley distribution (TLD) of Singh *et al.* (2014), are some among others.

The probability density function (pdf), the cumulative distribution function (cdf) and the rth moment about origin μ_r' of Prakaamy distribution obtained by Shukla (2018) are given by

$$g(x;\theta)=\frac{\theta^6}{\theta^5+120}\left(1+x^5\right)e^{-\theta x} \quad ;x>0,\ \theta>0, \tag{1.1}$$

$$G(x;\theta)=1-\left[1+\frac{\theta x\left(x^4\theta^4+5x^3\theta^3+20x^2\theta^2+60x\theta+120\right)}{\theta^5+120}\right]e^{-\theta x} \quad ;x>0,\theta>0, \tag{1.2}$$

$$\mu_r'=\frac{r!\left\{\theta^5+(r+1)(r+2)(r+3)(r+4)(r+5)\right\}}{\theta^r\left(\theta^5+120\right)} \quad ;r=1,2,3,... \tag{1.3}$$

Statistical properties, parameter estimation and applications of Prakaamy distribution are available in Shukla (2018).

The truncation of a continuous distribution from left and right can be defined as:

Definition1. Let a continuous random variable X has pdf $g(x;\theta)$ and cdf $G(x;\theta)$, where θ may be a vector of parameter. Let X defined in the interval $[a,b]$, where $-\infty<a\le x\le b<\infty$, then X, conditional on $a\le x\le b$ is known as doubly truncated distribution. The pdf and the cdf of X is defined as

$$f(x;\theta)=g(x\,|\,a\le x\le b;\theta)=\frac{g(x;\theta)}{G(b;\theta)-G(a;\theta)}, \tag{1.4}$$

$$F(x;\theta)=\int_a^x f(x;\theta)dx=\frac{G(x;\theta)-G(a;\theta)}{G(b;\theta)-G(a;\theta)}. \tag{1.5}$$

where $f(x;\theta)=g(x;\theta)$ for all $a\le x\le b$ and $f(x;\theta)=0$ elsewhere. Remember that $f(x;\theta)$, in fact, is a pdf of X on interval $a\le x\le b$

The main objective of the paper is to study a doubly truncated version of Prakaamy distribution, known as Truncated Prakaamy distribution. This study has been divided and presented into eight sections.

2. Truncated Prakaamy Distribution

This section deals with the pdf, cdf and survival function of truncated Prakaamy distribution.

Definition 2: Let X follows Truncated Prakaamy distribution (TPD) with location parameters a, b and scale parameter θ, denoted by TPD(a,b,θ). Using (1.1), (1.2), (1.4) and (1.5), the pdf, cdf and the survival function of TPD are obtained as:

$$f(x;\theta)=\frac{\theta^6\left(x^5+1\right)e^{-\theta x}}{\left[\begin{array}{l}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right)\end{array}\right]} \tag{2.1}$$

$$F(x;\theta)=\frac{\left[\begin{array}{l}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}-x\theta\left(x^4\theta^4+5x^3\theta^3+20x^2\theta^2+60x\theta+120\right)e^{-\theta x}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta x}\right)\end{array}\right]}{\left[\begin{array}{l}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right)\end{array}\right]} \tag{2.2}$$

$$S(x)=1-F(x)=\frac{\left[\begin{array}{l}x\theta\left(x^4\theta^4+x^3\theta^3+x^2\theta^2+x\theta+120\right)e^{-\theta x}\\-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta x}-e^{-\theta b}\right)\end{array}\right]}{\left[\begin{array}{l}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}\\-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right)\end{array}\right]} \tag{2.3}$$

where $-\infty < a \le x \le b < \infty$, and $\theta > 0$.

Behaviors of the pdf and the survival function of TPD have been illustrated in the figures 1 and 2. It is obvious from the figures that as the values of b increases from 1, the nature of the pdf of TPD changes from leptokurtic to platykurtic. Also as the values of θ increases from 0.5 the shape of the pdf of TPD changes drastically from smooth to zigzag, which shows that TPD can be used for data of any natures.

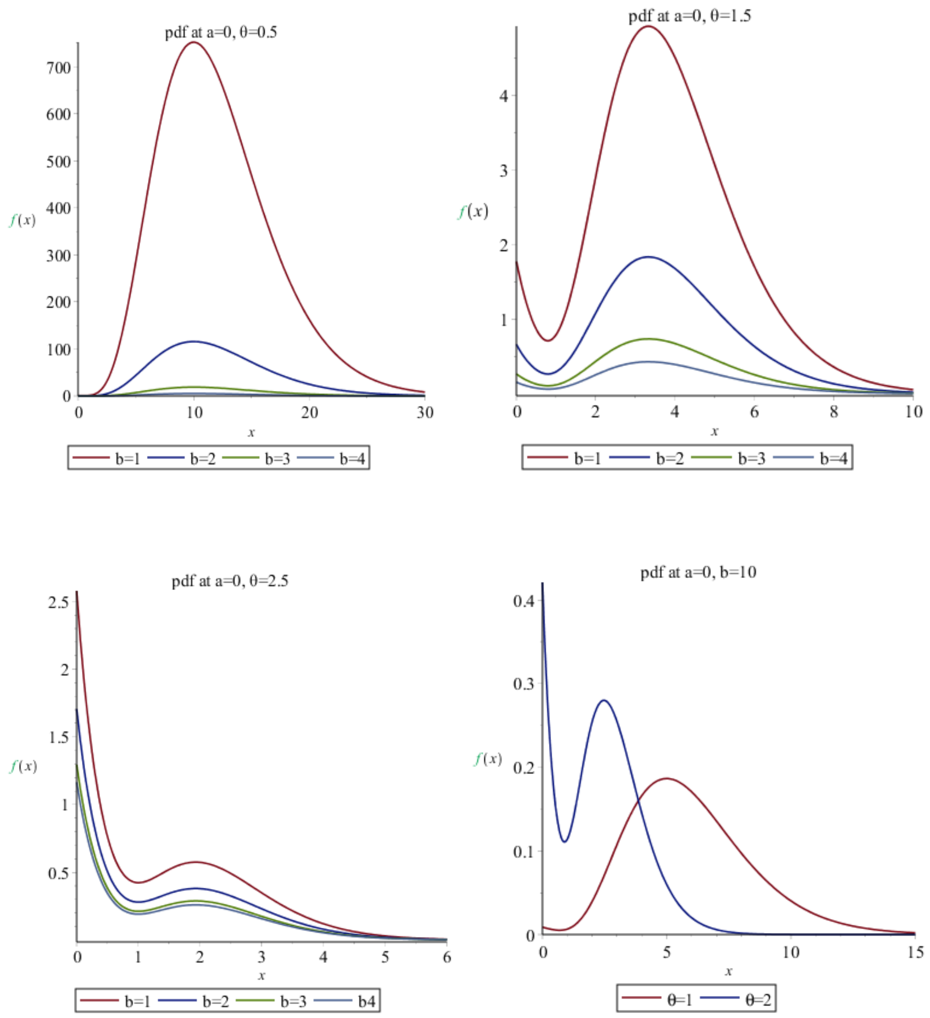
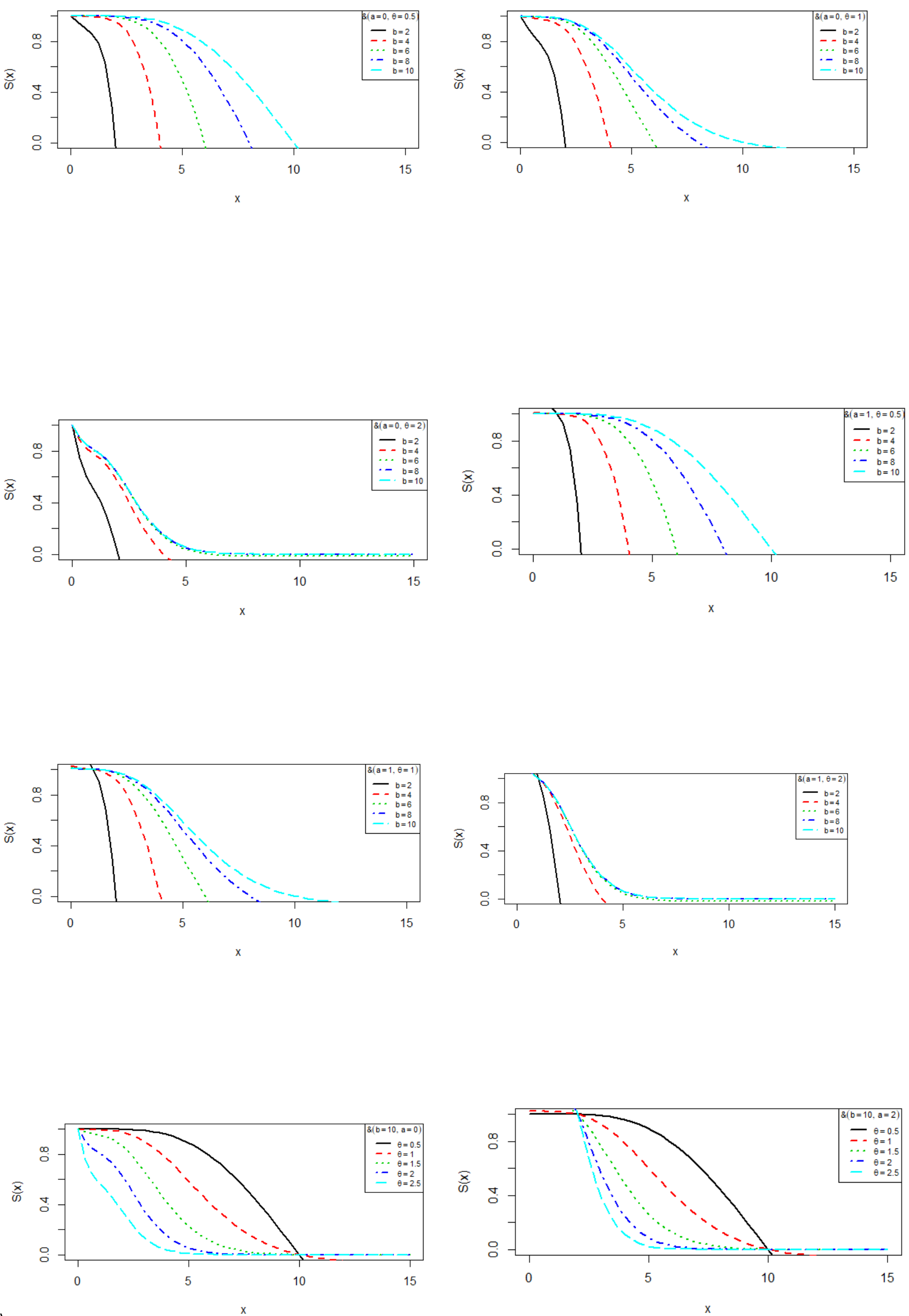


Fig. 1: pdf plots of TPD for varying values of parameters.



\

Fig 2: S(x) plots of TPD for varying values of parameter.

3. Hazard Rate Function

The hazard rate function $h(x)$ of TPD are defined as

$$h(x)=\frac{f(x)}{S(x)}=\frac{\theta^{6}\left(1+x^{5}\right)e^{-\theta x}}{\begin{bmatrix} x\theta\left(x^{4}\theta^{4}+5x^{3}\theta^{3}+20x^{2}\theta^{2}+60x\theta+120\right)e^{-\theta x} \\ -b\theta\left(b^{4}\theta^{4}+5b^{3}\theta^{3}+20b^{2}\theta^{2}+60b\theta+120\right)e^{-\theta b} \\ +\left(\theta^{5}+120\right)\left(e^{-\theta x}-e^{-\theta b}\right) \end{bmatrix}}$$

Here $h(x)$ is independent from location parameter a. Behavior of $h(x)$ of TPD has been shown in figure 3. From the graphs of $h(x)$, it is quite obvious that it changes very much from increasing to decreasing, upside bathtub etc., which shows the flexibility of TPD over un-truncated Prakaamy distribution.

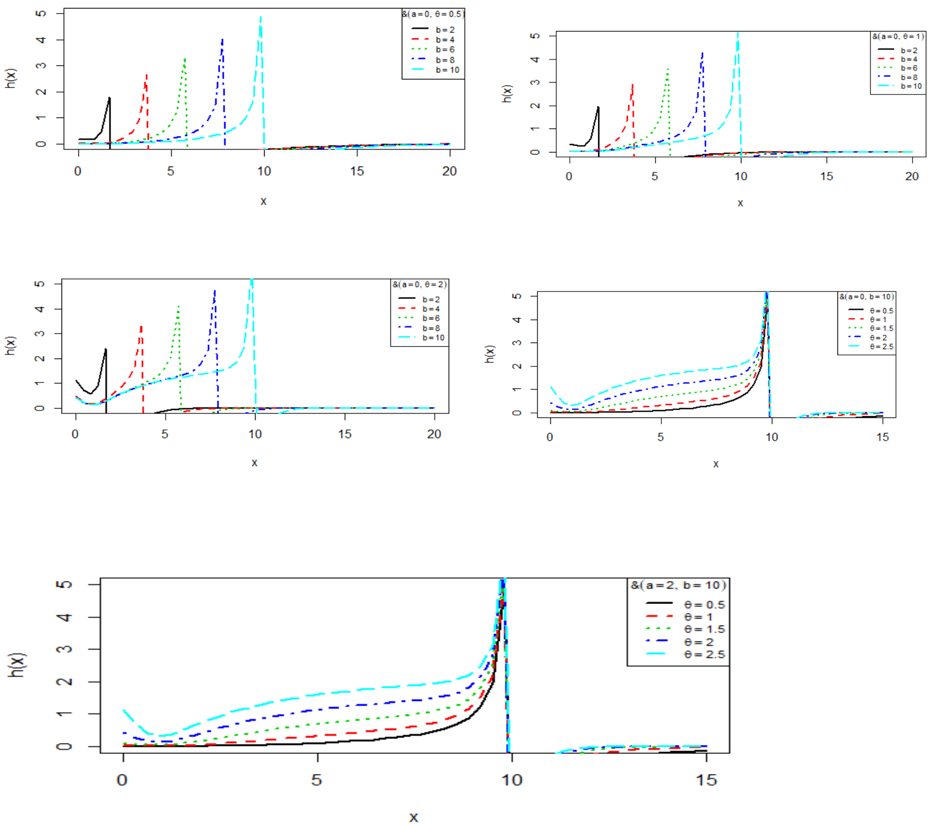


Fig 3: h(x) plots of TPD.

4. Moments based statistical constants

Moments of a distribution are used to study the most important characteristics of the distribution including mean, variance, skewness, kurtosis, etc. The rth moment about origin μ_r' of TPD can be expressed explicitly in terms of incomplete gamma functions.

Theorem: Suppose X follows doubly TPD (θ, a, b). Then the rth moment about origin μ_r' of TPD is

$$\mu_r' = \frac{\theta^5 \left\{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)\right\}+\left\{\gamma(r+6,\theta b)-\gamma(r+6,\theta a)\right\}}{\theta^r \begin{pmatrix} a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a} \\ -b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b} \\ +\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right) \end{pmatrix}} \; ; r = 1, 2, 3, ...$$

Proof: Considering $K = \begin{Bmatrix} a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a} \\ -b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b} \\ +\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right) \end{Bmatrix}$

in (2.1), we have

$$\mu_r' = \frac{\theta^6}{K}\int_a^b x^r\left(1+x^5\right)e^{-\theta x}dx$$

$$= \frac{\theta^6}{K}\left[\int_a^b e^{-\theta x}x^r dx + \int_a^b e^{-\theta x}x^{r+5}dx\right]$$

Taking $u = \theta x, x = \frac{u}{\theta}$, we get

$$= \mu_r' = \frac{\theta^6}{K}\left[\begin{matrix} \frac{1}{\theta^{r+1}}\left\{\int_0^{\theta b} e^{-u}x^r dx - \int_0^{\theta a} e^{-u}x^r du\right\} + \\ \frac{1}{\theta^{r+6}}\left\{\int_0^{\theta b} e^{-u}u^{r+5}du - \int_0^{\theta a} e^{-u}x^{r+5}du\right\} \end{matrix}\right]$$

Where $\gamma(\alpha,z)=\int_0^z e^{-x}x^{\alpha-1}dx, \alpha>0, x>0$ is the lower incomplete gamma function. Thus, we have

$$\mu_r{}'=\frac{\theta^6}{K}\left[\frac{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)}{\theta^{r+1}}+\frac{\gamma(r+6,\theta b)-\gamma(r+6,\theta a)}{\theta^{r+6}}\right]$$

$$=$$

$$\frac{1}{K}\left[\frac{\theta^5\left\{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)\right\}+\left\{\gamma(r+6,\theta b)-\gamma(r+6,\theta a)\right\}}{\theta^r}\right]$$

$$=$$

$$\frac{\theta^5\left\{\gamma(r+1,\theta b)-\gamma(r+1,\theta a)\right\}+\left\{\gamma(r+6,\theta b)-\gamma(r+6,\theta a)\right\}}{\theta^r\begin{pmatrix}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}\\-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right)\end{pmatrix}}$$

Now taking $r=1$ and 2, the mean and the variance of TPD can be obtained as

$$\mu_1^{'}=\frac{\theta^5\left\{\gamma(2,\theta b)-\gamma(2,\theta a)\right\}+\left\{\gamma(7,\theta b)-\gamma(7,\theta a)\right\}}{\theta\begin{pmatrix}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}\\-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right)\end{pmatrix}}$$

$$\mu_2^{'}=\frac{\theta^5\left\{\gamma(3,\theta b)-\gamma(3,\theta a)\right\}+\left\{\gamma(8,\theta b)-\gamma(8,\theta a)\right\}}{\theta^2\begin{pmatrix}a\theta\left(a^4\theta^4+5a^3\theta^3+20a^2\theta^2+60a\theta+120\right)e^{-\theta a}\\-b\theta\left(b^4\theta^4+5b^3\theta^3+20b^2\theta^2+60b\theta+120\right)e^{-\theta b}\\+\left(\theta^5+120\right)\left(e^{-\theta a}-e^{-\theta b}\right)\end{pmatrix}}$$

$$\mu_2=\mu_2^{'}-(\mu_1^{'})^2$$

5. Estimation of parameter

Assuming $(x_1, x_2, x_3, \ldots, x_n)$ a random sample from (2.1), the likelihood function, L and the log-likelihood function $\ln L$ of TPD are

$$L = \frac{\theta^{6n}}{\begin{bmatrix} a\theta\left(a^4\theta^4 + 5a^3\theta^3 + 20a^2\theta^2 + 60a\theta + 120\right)e^{-\theta a} \\ -b\theta\left(b^4\theta^4 + 5b^3\theta^3 + 20b^2\theta^2 + 60b\theta + 120\right)e^{-\theta b} \\ +\left(\theta^5 + 120\right)\left(e^{-\theta a} - e^{-\theta b}\right) \end{bmatrix}^n} \prod_{i=1}^{n}\left(1 + x_i^{\,5}\right)e^{-n\theta \bar{x}}$$

$$\ln L = n \ln \frac{\theta^6}{\begin{bmatrix} a\theta\left(a^4\theta^4 + 5a^3\theta^3 + 20a^2\theta^2 + 60a\theta + 120\right)e^{-\theta a} \\ -b\theta\left(b^4\theta^4 + 5b^3\theta^3 + 20b^2\theta^2 + 60b\theta + 120\right)e^{-\theta b} \\ +\left(\theta^5 + 120\right)\left(e^{-\theta a} - e^{-\theta b}\right) \end{bmatrix}} + \sum_{i=1}^{n}\ln\left(1 + x_i^{\,5}\right) - n\theta\bar{x}$$

Suppose $\hat{a} = \min(x_1, x_2, x_3, \ldots, x_n)$, $\hat{b} = \max(x_1, x_2, x_3, \ldots, x_n)$. Then the maximum likelihood estimate (MLE) $\hat{\theta}$ of parameter θ is the solution of $\frac{\partial \ln L}{\partial \theta} = 0$. Obviously $\frac{\partial \ln L}{\partial \theta} = 0$ is not in closed form and hence some mathematical optimization method can be applied for MLE of θ. Here, nonlinear optimization method available in R software (stats4 package) has been used to find the MLE of θ.

6. A Simulation study

The simulation study of TPD has been studied based on random number generated using Acceptance - Rejection method to know performance and behavior of proposed distribution. To conduct simulation study, following steps have been followed: generate 10000 samples of size n = 40, 60, 80, 100 from the TPD (0.1, 0.5, 1, and 1.5) compute the Average Bias Error (ABE) and MSE using MLE(θ) and proposed value of parameter (θ). The simulation results are presented in tables 1 and 2 respectively.

Table1: ABE and MSE for the varying values parameter θ with fixed value of a=0 and b=10.

Sample Size (n)	Parameter (θ)	ABE	MSE
40	0.1 0.5 1.0 1.5	0.04115 0.031459 0.015093 0.008977	0.06776 0.03958 0.00911 0.00322
60	0.1 0.5 1.0 1.5	0.027237 0.018923 0.014446 0.002703	0.04451 0.02148 0.01252 0.00043
80	0.1 0.5 1.0 1.5	0.02094 0.01469 0.00605 0.00214	0.03509 0.01728 0.002936 0.000368
100	0.1 0.5 1.0 1.5	0.01539 0.011612 0.006078 0.000874	0.023688 0.013484 0.003694 0.000076

Table2. ABE and MSE for the varying values parameter θ with fixed value of a=2 and b=10.

Sample Size (n)	Parameter (θ)	ABE	MSE
40	0.1 0.5 1.0 1.5	0.0415202 0.0315202 0.0190202 0.0065202	0.0689572 0.0397410 0.0144707 0.0017005
60	0.1 0.5 1.0 1.5	0.0266648 0.0199982 0.0116648 0.0033315	0.04266101 0.02399575 0.00816419 0.00066595
80	0.1 0.5 1.0 1.5	0.0191329 0.0141329 0.0078829 0.0016329	0.02928558 0.01597922 0.00497127 0.00021332
100	0.1 0.5 1.0 1.5	0.014699778 0.01069977 0.005699778 0.000699778	0.0216083 0.0114485 0.003248747 0.00004896

It is obvious from tables 1 and 2 that the Average Bias Error and the Mean square error are decreasing with increased sample size and parameter θ respectively.

7. Applications

The applications of TPD have been discussed with two datasets using maximum likelihood estimates. The first dataset given by Birnbaum and Saunders (1969). The second dataset is reported by Fuller *et al* (1994). Parameter θ is estimated using MLE whereas another parameters a, and b are considered as lowest and highest values of data. i. e. $\hat{a} = \min(x_1, x_2, x_3, \ldots, x_n)$ and $\hat{b} = \max(x_1, x_2, x_3, \ldots, x_n)$. Goodness of fit has been decided using Akaike information criteria (AIC), Bayesian Information criteria (BIC) and Kolmogorov Simonov test (KS), respectively.

The MLE's- 2ln L, SE, AIC and K-S test of the fitted distributions for datasets 1 & 2 have been presented in table 3 and 4 respectively. The plots of the fitted distributions for datasets 1& 2 have been presented in figures 4 and 5. It is obvious from the goodness of fit of the considered distributions from the tables 3 and 4 and the plots of the fitted distributions as well as from figures 4 & 5 that TPD can be the best model among all considered distributions.

Table 3: MLE's, Standard Errors, - 2ln L, AIC, BIC, K-S and p-values for first dataset.

Distributions	ML Estimates	Standard Errors	$-2\ln L$	AIC	BIC	K-S	p-value
TPD	$\theta = 0.08625$	0.00380	916.24	918.24	920.16	0.105	0.217
TLD	$\theta = 0.02199$	0.00273	958.88	960.88	962.31	0.186	0.001
Prakaamy	$\theta = 0.08781$	0.00358	918.72	920.72	922.63	0.116	0.132
Akash	$\theta = 0.04387$	0.00253	950.97	952.97	954.40	0.194	0.001
Ishita	$\theta = 0.04390$	0.002533	950.92	9952.92	954.35	0.194	0.001
Lindley	$\theta = 0.02886$	0.002038	983.10	985.10	986.54	0.252	0.000
Exponential	$\theta = 0.01463$	0.001457	1044.87	1046.87	1048.30	0.336	0.000

Table 4: MLE's, Standard Errors, - 2ln L, AIC,BIC K-S and p-values for second dataset.

Distributions	ML Estimates	Standard Errors	$-2\ln L$	AIC	BIC	K-S	p-value
TPD	$\theta = 0.18629$	0.02519	201.63	203.63	205.54	0.103	0.862
TLD	$\theta = 0.05392$	0.023917	202.18	204.18	205.61	0.117	0.738
Prakaamy	$\theta = 0.19473$	0.01427	223.07	225.07	227.31	0.197	0.154
Akash	$\theta = 0.09706$	0.01004	240.68	242.68	242.67	0.298	0.005
Ishita	$\theta = 0.097328$	0.01008	240.48	242.48	243.48	0.297	0.006
Lindley	$\theta = 0.06299$	0.00800	253.98	255.98	256.98	0.365	0.000
Exponential	$\theta = 0.032452$	0.00582	274.52	276.52	277.52	0.458	0.000

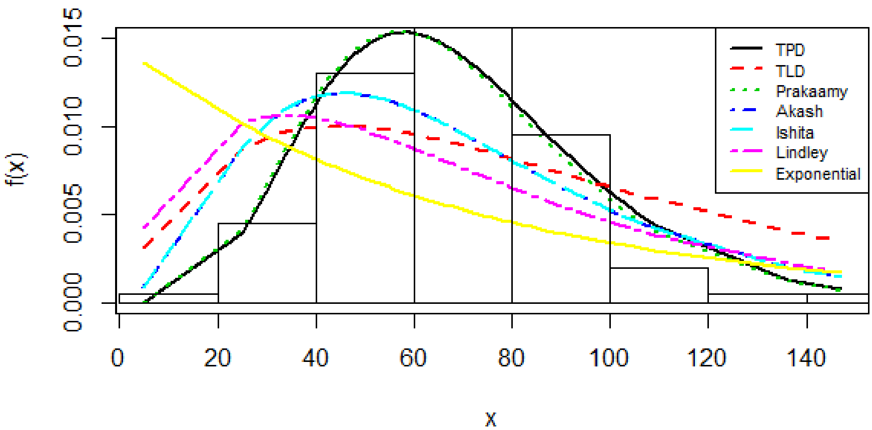


Fig. 4: Plot of the fitted distributions for first dataset

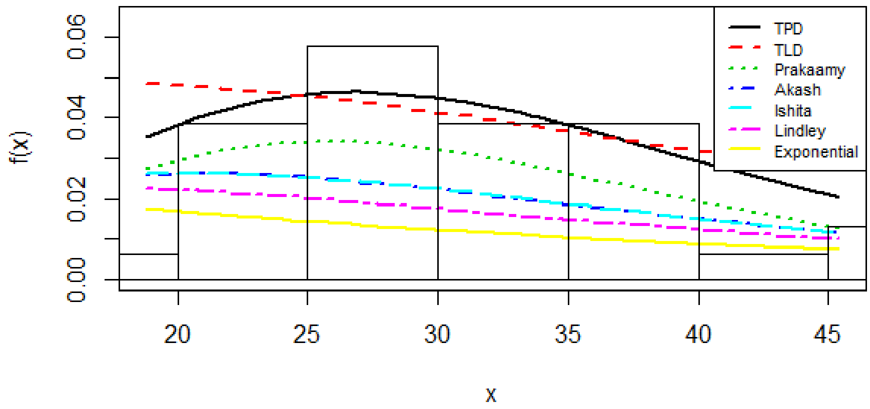


Fig. 5: Plot of the fitted distributions for second dataset

6. Concluding Remarks

A truncated Prakaamy distribution (TPD) has been proposed. Behaviors of proposed distribution have been studied. Mean and variance of the TPD have been obtained. Maximum likelihood estimation has been explained for estimation of parameter. Goodness of fit of TPD has been discussed with two lifetime datasets and compared with TLD other one parameter lifetime distribution including exponential distributions and it is obvious that it provides better fit for the considered dataset.

Acknowledgements

Authors are grateful to the anonymous reviewers for their constructive and valuable suggestions for improving the quality and the presentation of the paper.

References

Aryuyuen,S. and Bodhisuwan, W. (2018): The truncated power Lomax distribution: Properties and applications, *Walailak Journal of Science and Technology*, **16**, 655-668.

Birnbaum, Z.W., & Saunders, S.C. (1969): Estimation for a family of life distribution with applications to fatigue. *Journal of applied probability*, **6(2)**, 328-347

Fuller, E.J., Friedman, S., Quinn, J., Quinn, G., and Carter, W.(1994): Fracture mechanics approach to the design of glass aircraft windows: A case study, SPIE Proc. **2286**, 419-430.

Ghitany ME, Atieh B, Nadarajah S. (2008): Lindley distribution and its applications. *Mathematics Computing and Simulation*. **78,** 493– 506.

Johnson, N. L., Kotz, S. and Balakrishnan, N. (1994): Continuous Univariate Distributions, Wiley, New York (1994).

Lindley, D.V. (1958): Fiducial distributions and Bayes' Theorem, *Journal of the Royal Statistical Society*, Series B, **20,** 102 – 107.

Shanker, R. (2015): Akash distribution and its application, *International Journal of Probability and Statistics*. **4(3)**, 65–75.

Shanker,R. and Shukla, K.K. (2017): Ishita distribution and its applications, *Biometrics & Biostatistics International Journal,* **5,** 1-9.

Shukla, K. K. (2018). Prakaamy Distribution with properties and Applications, *Journal of Applied Quantitative Methods* **13(3),** 30- 38.

Singh,S. K., Singh,U. and Sharma, V.K. (2014): The truncated Lindley distribution: Inference and application, *Journal of Statistics Applications & Probability*, **3(2),** 219-228.

Zange,Z., and Xie, M., (2011): On the upper truncated Weibull distribution and its reliability implications, *Reliability Engineering & System Safety*, **96,** 194-200.

Zaninetti,L. and Ferraro,M. (2008): On the truncated Pareto distribution with applications, *Central European Journal of Physics*, **6**, 1-6.

Authors and Affiliations

Kamlesh Kumar Shukla[1] and Rama Shanker[2]

Rama Shanker
Email:shankerrama2009@gmail.com

[1]Department of Community Medicine, Noida International Institute of Medical Sciences, Noida International University, Gautam Budh Nagar
[2]Department of Statistics, Assam University, Silchar, Assam