

POWER AKASH DISTRIBUTION AND ITS APPLICATION

Rama SHANKER¹

PhD, University Professor, Department of Statistics,
College of Science, Eritrea Institute of Technology, Asmara, Eritrea

E-mail: shankerrama2009@gmail.com



Kamlesh Kumar SHUKLA^{2,*}

PhD, Associate Professor, Department of Statistics,
Eritrea Institute of Technology, Asmara, Eritrea

E-mail: kkshukla22@gmail.com

* Corresponding author



Abstract

A two-parameter power Akash distribution (PAD), which includes Akash distribution introduced by Shanker (2015) as a particular case, has been proposed and its important statistical properties including shapes of the density, hazard rate function, moments, skewness and kurtosis measures, and stochastic ordering have been discussed. The maximum likelihood estimation has been discussed for estimating its parameters. Finally, the goodness of fit PAD has been discussed with a real lifetime data set from engineering and the fit has been found better as compared with two-parameter power Lindley distribution (PLD) and one parameter Akash, Lindley and exponential distributions.

Key words: Akash distribution; Hazard rate function; Moments; Stochastic ordering; Maximum likelihood estimation; Goodness of fit

1. Introduction

The probability density function (pdf) of Akash distribution introduced by Shanker (2015) is given by

$$f(y; \theta) = \frac{\theta^3}{\theta^2 + 2} (1 + y^2) e^{-\theta y} ; y > 0, \theta > 0 \quad (1.1)$$

$$= p f_1(y; \theta) + (1 - p) f_2(y; \theta)$$

where

$$p = \frac{\theta^2}{\theta^2 + 2}$$

$$f_1(y; \theta) = \theta e^{-\theta y} ; y > 0$$

$$f_2(y; \theta) = \frac{\theta^3 y^2 e^{-\theta y}}{2} ; y > 0$$

The pdf in (1.1) reveals that the Akash distribution is a two – component mixture of an exponential distribution (with scale parameter θ) and a gamma distribution (with shape parameter 2 and scale parameter θ), with mixing proportion $p = \frac{\theta^2}{\theta^2 + 2}$. Shanker (2015)

has discussed some of its mathematical and statistical properties including its shape, moments, skewness, kurtosis, hazard rate function, mean residual life function, stochastic ordering, mean deviations, order statistics, Bonferroni and Lorenz curves, entropy measure, stress-strength reliability, and applications of Akash distribution for modeling lifetime data from engineering and biomedical sciences. However, there are some situations where the Akash distribution may not be suitable from either theoretical or applied point of view. Recently, Shanker et al (2016) have done a critical study on the applications of one parameter Akash, Lindley and exponential distributions and observed that each of these distributions has some advantage over the others and there are some situations where these distributions do not provide satisfactory fit. Shanker (2017) has also obtained a Poisson mixture of Akash distribution and named it Poisson-Akash distribution and showed its superiority over Poisson-Lindley distribution and Poisson distribution for modeling count data.

The corresponding cumulative distribution function (cdf) of (1.1) is given by

$$F(y; \theta) = 1 - \left[1 + \frac{\theta y (\theta y + 2)}{\theta^2 + 2} \right] e^{-\theta y} ; y > 0, \theta > 0 \quad (1.2)$$

In this paper, a power Akash distribution (PAD) has been introduced and its various properties including shapes of density function for varying values of parameters, survival function, hazard rate function, moments and stochastic ordering has been studied. The maximum likelihood estimation for estimating its parameters has been discussed. Finally an application of PAD has been illustrated with a real lifetime data from engineering and the PAD shows satisfactory fit over PLD, Akash, Lindley and exponential distributions.

2. Power Akash distribution

Assuming the power transformation $X = Y^{1/\alpha}$ in (1.1), the pdf of the random variable X can be obtained as

$$f_1(x; \theta, \alpha) = \frac{\alpha \theta^3}{(\theta^2 + 2)} (1 + x^{2\alpha}) x^{\alpha-1} e^{-\theta x^\alpha} ; x > 0, \theta > 0, \alpha > 0 \quad (2.1)$$

$$= p g_1(x; \theta, \alpha) + (1 - p) g_2(x; \theta, \alpha) \quad (2.2)$$

where

$$p = \frac{\theta^2}{\theta^2 + 2}$$

$$g_1(x; \theta, \alpha) = \alpha \theta x^{\alpha-1} e^{-\theta x^\alpha}; x > 0, \theta > 0, \alpha > 0$$

$$g_2(x; \theta, \alpha) = \frac{\alpha \theta^3 x^{3\alpha-1} e^{-\theta x^\alpha}}{2}; x > 0, \theta > 0, \alpha > 0$$

We would call the density in (2.1) a power Akash distribution (PAD). It is obvious that the PAD is also a two-component mixture of Weibull distribution (with shape parameter α and scale parameter θ), and a generalized gamma distribution (with shape parameters

$3, \alpha$ and scale parameter θ) with mixing proportion $p = \frac{\theta^2}{\theta^2 + 2}$. The corresponding cdf of (2.1) can be obtained as

$$F_1(x; \theta, \alpha) = 1 - \left[1 + \frac{\theta x^\alpha (\theta x^\alpha + 2)}{\theta^2 + 2} \right] e^{-\theta x^\alpha}; x > 0, \theta > 0, \alpha > 0 \quad (2.3)$$

To study the nature and behavior of the pdf of PAD, various graphs of the pdf of PAD for varying values of parameters have been drawn and presented in figure 1. From the graphs of the pdf of PAD, it is clear that it takes different shapes for varying values of parameters. As the values of α increases, the shapes of PAD become normal. Further, if the value of $\alpha > 3$, then the shapes of PAD become leptokurtic.

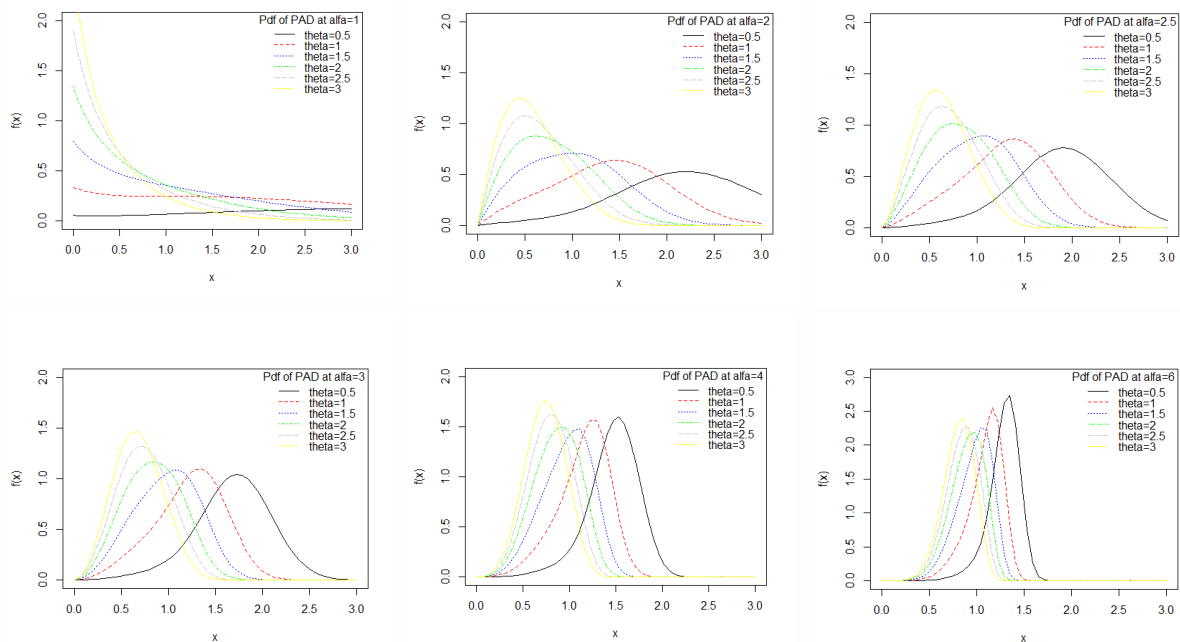


Figure 1. Graphs of the pdf of PAD for varying values of the parameters θ and α

To study the nature and behavior of the cdf of PAD, various graphs of cdf of PAD for varying values of parameters have been drawn and presented in figure 2.

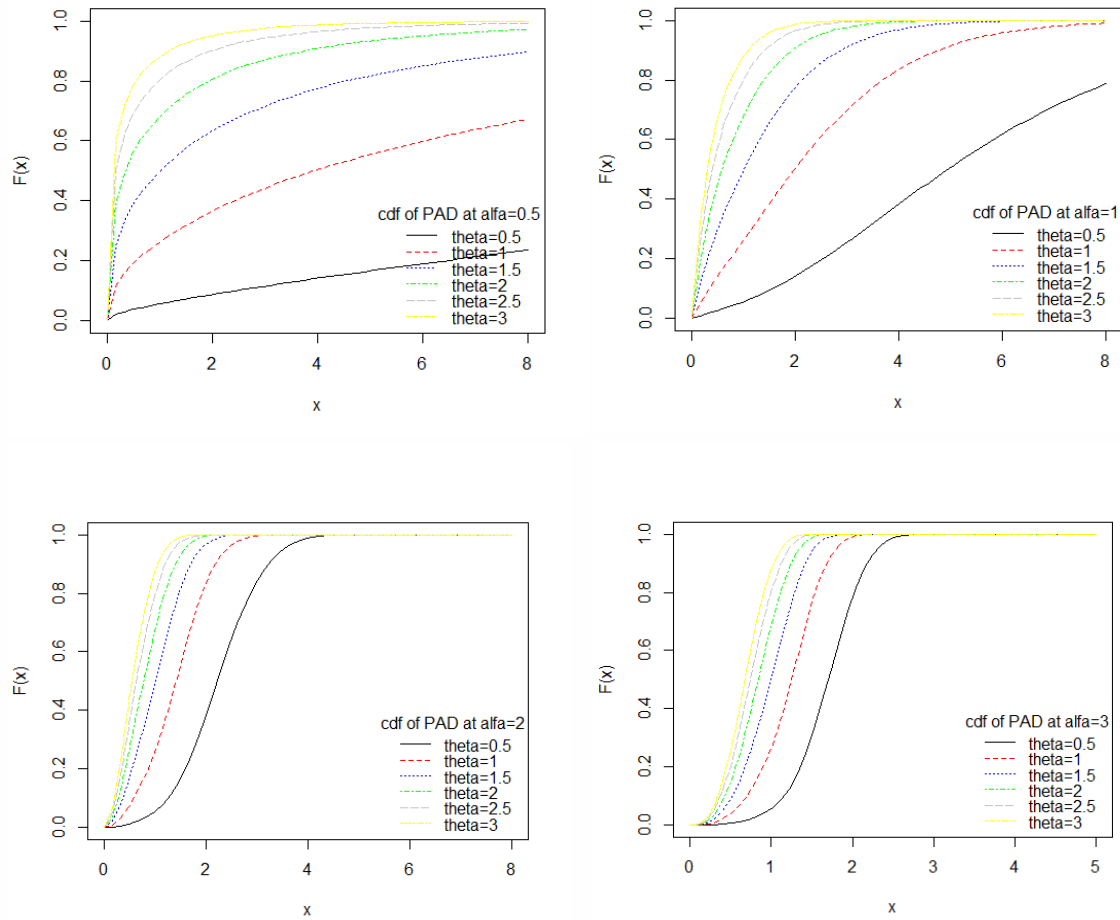


Figure 2. Graphs of the cdf of PAD for varying values of the parameters θ and α

Recall that Ghitany *et al* (2013) obtained the power Lindley distribution (PLD) having pdf and cdf given by

$$f_2(x; \theta, \alpha) = \frac{\alpha \theta^2}{(\theta + 1)} (1 + x^\alpha) x^{\alpha-1} e^{-\theta x^\alpha}; x > 0, \theta > 0, \alpha > 0 \quad (2.4)$$

$$F_2(x; \theta, \alpha) = 1 - \left[1 + \frac{\theta x^\alpha}{\theta + 1} \right] e^{-\theta x^\alpha}; x > 0, \theta > 0, \alpha > 0 \quad (2.5)$$

Ghitany *et al* (2013) have discussed its mathematical and statistical properties and established its goodness of fit over other distributions.

3. Survival and hazard rate functions

The survival function, $S(x)$, and hazard rate function, $h(x)$, of PAD can be obtained a

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

$$h(x) = \frac{f(x)}{S(x)} = \frac{\alpha \theta^3 (1 + x^{2\alpha}) x^{\alpha-1}}{\theta x^\alpha (\theta x^\alpha + 2) + (\theta^2 + 2)}; x > 0, \theta > 0, \alpha > 0 \quad (3.1)$$

The behavior of $h(x)$ at $x=0$ and $x=\infty$, respectively, are given by

$$h(0) = \begin{cases} \infty & \text{if } 0 < \alpha < 1 \\ \frac{\theta^3}{\theta^2 + 2} & \text{if } \alpha = 1 \\ 0 & \text{if } \alpha > 1 \end{cases}, \text{ and } h(\infty) = \begin{cases} 0 & \text{if } \alpha < 1 \\ \theta^3 & \text{if } \alpha = 1 \\ \infty & \text{if } \alpha > 1 \end{cases}$$

The shapes of hazard rate function of PAD for varying values of the parameters are shown in the figure 3. The graphs of hazard rate function of PAD shows that it takes different shapes including monotonically increasing and decreasing for varying values of parameters.

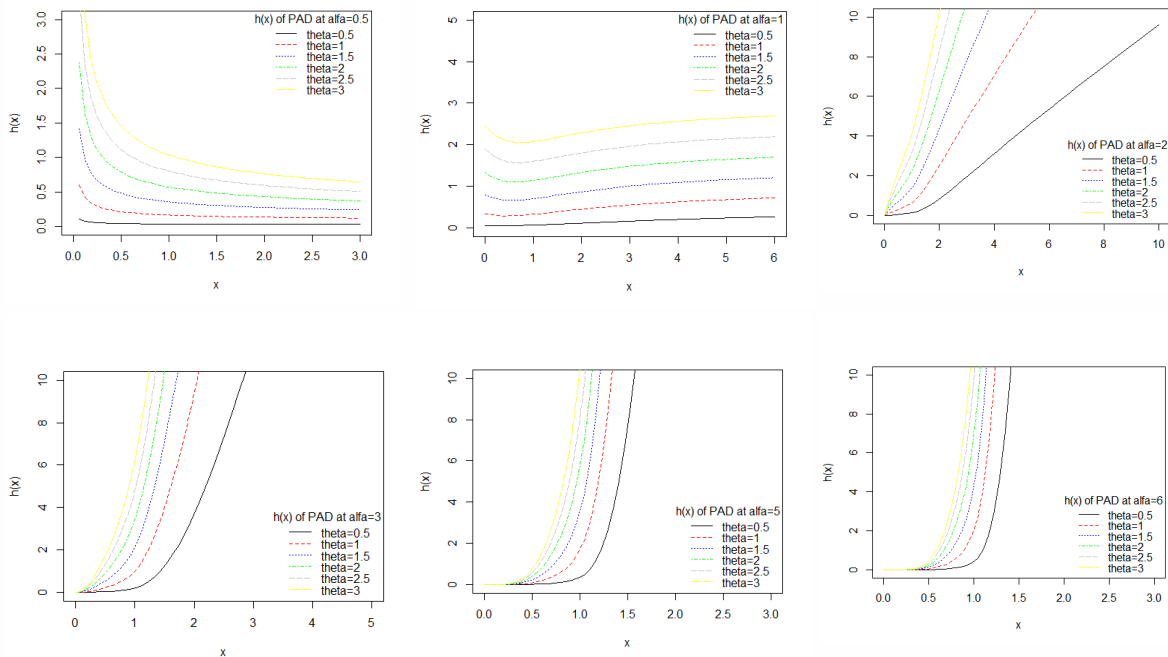


Figure 3. Graphs of $h(x)$ of PAD for varying values of the parameters θ and α

4. Moments and related measures

Using the mixture representation (2.2), the r th moment about origin of PAD can be obtained as

$$\mu_r' = E(X^r) = p \int_0^\infty x^r g_1(x; \theta, \alpha) dx + (1-p) \int_0^\infty x^r g_2(x; \theta, \alpha) dx$$

$$= \frac{r\Gamma\left(\frac{r}{\alpha}\right)\left[\alpha^2\theta^2 + (r+\alpha)(r+2\alpha)\right]}{\alpha^3\theta^{r/\alpha}(\theta^2+2)}; r=1,2,3,\dots \quad (4.1)$$

It should be noted that at $\alpha = 1$, the above expression will reduce to the r th moment about origin of Akash distribution and is given by

$$\mu_r' = \frac{r!\left[\theta^2 + (r+1)(r+2)\right]}{\theta^r(\theta^2+2)}; r=1,2,3,\dots$$

Therefore, the mean and the variance of PAD, respectively, are obtained as

$$\mu = \frac{\Gamma\left(\frac{1}{\alpha}\right)\left[\alpha^2\theta^2 + (\alpha+1)(2\alpha+1)\right]}{\alpha^3\theta^{1/\alpha}(\theta^2+2)}$$

$$\sigma^2 = \frac{2\Gamma\left(\frac{2}{\alpha}\right)\left[\alpha^2\theta^2 + 2(\alpha+1)(\alpha+2)\right]\alpha^3(\theta^2+2) - \left(\Gamma\left(\frac{1}{\alpha}\right)\right)^2\left[\alpha^2\theta^2 + (\alpha+1)(2\alpha+1)\right]^2}{\alpha^6\theta^{2/\alpha}(\theta^2+2)^2}$$

The skewness and kurtosis measures of PAD, upon substituting for the raw moments, can be obtained using the expressions

$$\text{Skewness} = \frac{\mu_3' - 3\mu_2'\mu + 2\mu^3}{\sigma^3} \quad \text{and} \quad \text{Kurtosis} = \frac{\mu_4' - 4\mu_3'\mu + 6\mu_2'\mu^2 - 3\mu^4}{\sigma^4}.$$

5. Stochastic ordering

Stochastic ordering of positive continuous random variables is an important tool for judging their comparative behavior. A random variable X is said to be smaller than a random variable Y in the

- (i) stochastic order ($X \leq_{st} Y$) if $F_X(x) \geq F_Y(x)$ for all x
- (ii) hazard rate order ($X \leq_{hr} Y$) if $h_X(x) \geq h_Y(x)$ for all x
- (iii) mean residual life order ($X \leq_{mrl} Y$) if $m_X(x) \leq m_Y(x)$ for all x
- (iv) likelihood ratio order ($X \leq_{lr} Y$) if $\frac{f_X(x)}{f_Y(x)}$ decreases in x .

The following important interrelationships due to Shaked and Shanthikumar (1994) are well known for establishing stochastic ordering of distributions

$$X \leq_{lr} Y \Rightarrow X \leq_{hr} Y \Rightarrow X \leq_{mrl} Y$$

$$\Downarrow$$

$$X \leq_{st} Y$$

The PAD is ordered with respect to the strongest 'likelihood ratio ordering' as shown in the following theorem:

Theorem: Let $X \sim \text{PAD}(\theta_1, \alpha_1)$ and $Y \sim \text{PAD}(\theta_2, \alpha_2)$. If $\theta_1 > \theta_2$ and $\alpha_1 = \alpha_2$ (or $\alpha_1 < \alpha_2$)

and $\theta_1 = \theta_2$) then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Proof: We have

$$\frac{f_X(x)}{f_Y(x)} = \frac{\alpha_1 \theta_1^3 (\theta_2^2 + 2)}{\alpha_2 \theta_2^3 (\theta_1^2 + 2)} x^{\alpha_1 - \alpha_2} \left(\frac{1 + x^{2\alpha_1}}{1 + x^{2\alpha_2}} \right) e^{-(\theta_1 x^{\alpha_1} - \theta_2 x^{\alpha_2})}; \quad x > 0$$

Now

$$\log \frac{f_X(x)}{f_Y(x)} = \log \left[\frac{\alpha_1 \theta_1^3 (\theta_2^2 + 2)}{\alpha_2 \theta_2^3 (\theta_1^2 + 2)} \right] + (\alpha_1 - \alpha_2) \log x + \log \left[\frac{1 + x^{2\alpha_1}}{1 + x^{2\alpha_2}} \right] - (\theta_1 x^{\alpha_1} - \theta_2 x^{\alpha_2})$$

This gives

$$\begin{aligned} \frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} &= \frac{\alpha_1 - \alpha_2}{x} + \frac{4(\alpha_1 x^{2\alpha_1 - 1} - \alpha_2 x^{2\alpha_2 - 1}) + 4(\alpha_1 x^{2(\alpha_1 + \alpha_2) - 1} - \alpha_2 x^{2(\alpha_1 + \alpha_2) - 1})}{(1 + x^{2\alpha_1})(1 + x^{2\alpha_2})} \\ &\quad - (\alpha_1 \theta_1 x^{\alpha_1 - 1} - \alpha_2 \theta_2 x^{\alpha_2 - 1}) \end{aligned}$$

Thus for $\theta_1 > \theta_2$ and $\alpha_1 = \alpha_2$ (or $\alpha_1 < \alpha_2$ and $\theta_1 = \theta_2$), $\frac{d}{dx} \log \frac{f_X(x)}{f_Y(x)} < 0$. This means that $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

6. Maximum likelihood estimation of parameters

Let $(x_1, x_2, x_3, \dots, x_n)$ be a random sample of size n from PAD (θ, α) . Then log-likelihood function is given by

$$\begin{aligned} \ln L &= \sum_{i=1}^n \ln f(x_i) \\ &= n \left[\ln \alpha + 3 \ln \theta - \ln(\theta^2 + 2) \right] + \sum_{i=1}^n \ln(1 + x_i^{2\alpha}) + (\alpha - 1) \sum_{i=1}^n \ln(x_i) - \theta \sum_{i=1}^n x_i^\alpha. \end{aligned}$$

The maximum likelihood estimate (MLE) $(\hat{\theta}, \hat{\alpha})$ of (θ, α) of PAD (2.1) are the solutions of the following equations

$$\begin{aligned} \frac{\partial \ln L}{\partial \theta} &= \frac{3n}{\theta} - \frac{2n\theta}{\theta^2 + 2} - \sum_{i=1}^n x_i^\alpha = 0 \\ \frac{\partial \ln L}{\partial \alpha} &= \frac{n}{\alpha} + 2 \sum_{i=1}^n \frac{x_i^{2\alpha} \ln(x_i)}{1 + x_i^{2\alpha}} + \sum_{i=1}^n \ln(x_i) - \theta \sum_{i=1}^n x_i^\alpha \ln(x_i) = 0 \end{aligned}$$

These two likelihood equations do not seem to be solved directly because these cannot be expressed in closed form. However, Fisher's scoring method can be applied to solve these equations iteratively. For, we have

$$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{3n}{\theta^2} + \frac{2n(\theta^2 - 2)}{(\theta^2 + 2)^2}$$

$$\frac{\partial^2 \ln L}{\partial \alpha^2} = -\frac{n}{\alpha^2} + 4 \sum_{i=1}^n \frac{(x_i \ln(x_i))^2}{(1+x_i^{2\alpha})^2} - \theta \sum_{i=1}^n x_i^\alpha (\ln(x_i))^2$$

$$\frac{\partial^2 \log L}{\partial \theta \partial \alpha} = -\sum_{i=1}^n x_i^\alpha \ln(x_i)$$

The MLE $(\hat{\theta}, \hat{\alpha})$ of (θ, α) of PAD (2.1) are the solution of the following equations

$$\begin{bmatrix} \frac{\partial^2 \ln L}{\partial \theta^2} & \frac{\partial^2 \ln L}{\partial \theta \partial \alpha} \\ \frac{\partial^2 \ln L}{\partial \theta \partial \alpha} & \frac{\partial^2 \ln L}{\partial \alpha^2} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}} \begin{bmatrix} \hat{\theta} = \theta_0 \\ \hat{\alpha} = \alpha_0 \end{bmatrix} = \begin{bmatrix} \frac{\partial \ln L}{\partial \theta} \\ \frac{\partial \ln L}{\partial \alpha} \end{bmatrix}_{\substack{\hat{\theta}=\theta_0 \\ \hat{\alpha}=\alpha_0}}$$

where θ_0 and α_0 are initial values of θ and α . These equations are solved iteratively till sufficiently close estimates of $\hat{\theta}$ and $\hat{\alpha}$ are obtained. In this paper, R-software has been used to estimate the parameters θ and α for the given data sets.

7. Data analysis

In this section, we present the goodness of fit of PAD using maximum likelihood estimates of parameters to a real data set from engineering and compare its fit with the one parameter exponential, Lindley and Akash distributions and two-parameter PLD. The following real lifetime data have been considered for the goodness of fit of the proposed distribution.

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

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								

In order to compare these distributions, $-2 \ln L$, AIC (Akaike Information Criterion), K-S Statistic (Kolmogorov-Smirnov Statistic) for the real data set have been computed using maximum likelihood estimates and presented in table 1. The formulae for computing AIC and K-S Statistics are as follows:

$$AIC = -2 \ln L + 2k \quad \text{and} \quad K-S = \sup_x |F_n(x) - F_0(x)|, \quad \text{where } k = \text{the number of}$$

parameters, n = the sample size, $F_n(x)$ is the empirical (sample) cumulative distribution function and $F_0(x)$ is the theoretical cumulative distribution function.

The best distribution is the distribution corresponding to lower values of $-2\ln L$, AIC, and K-S statistics and higher p-value.

Table 1. MLE's, $-2\ln L$, AIC, K-S Statistic and p -value of the fitted distributions of the data set

Model	ML Estimates	-2ln L	AIC	K-S Statistic	p-value
PAD	$\hat{\theta} = 0.169$ $\hat{\alpha} = 3.061$	98.02	102.02	0.038	0.999
PLD	$\hat{\theta} = 0.050$ $\hat{\alpha} = 3.868$	98.12	102.12	0.044	0.998
Akash	$\hat{\theta} = 0.96473$	224.28	226.28	0.348	0.001
Lindley	$\hat{\theta} = 0.65900$	238.38	240.38	0.390	0.000
Exponential	$\hat{\theta} = 0.40794$	261.74	263.74	0.434	0.000

From the goodness of fit of two-parameter PAD and PLD and one parameter exponential, Lindley and Akash distributions in table 1, it is obvious that PAD is well competing with PLD and gives better fit and thus it can be used for modeling lifetime data from engineering over PLD and other one parameter lifetime distributions.

The variance-covariance matrix and 95% confidence intervals (CI's) for the parameters $\hat{\theta}$ and $\hat{\alpha}$ of PAD for data set has been presented in table 2.

Table 2. Variance-covariance matrix and 95% confidence intervals (CI's) for the parameters $\hat{\theta}$ and $\hat{\alpha}$ of PAD

Parameters	Variance-Covariance Matrix		95% CI	
	$\hat{\theta}$	$\hat{\alpha}$	Lower	Upper
$\hat{\theta}$	0.001799	- 0.010092	0.10096	0.26873
$\hat{\alpha}$	- 0.010092	0.061291	2.60037	3.57013

The profile of likelihood estimation for parameters of PAD for parameters $\hat{\theta}$ and $\hat{\alpha}$ for data has been presented in figure 4.

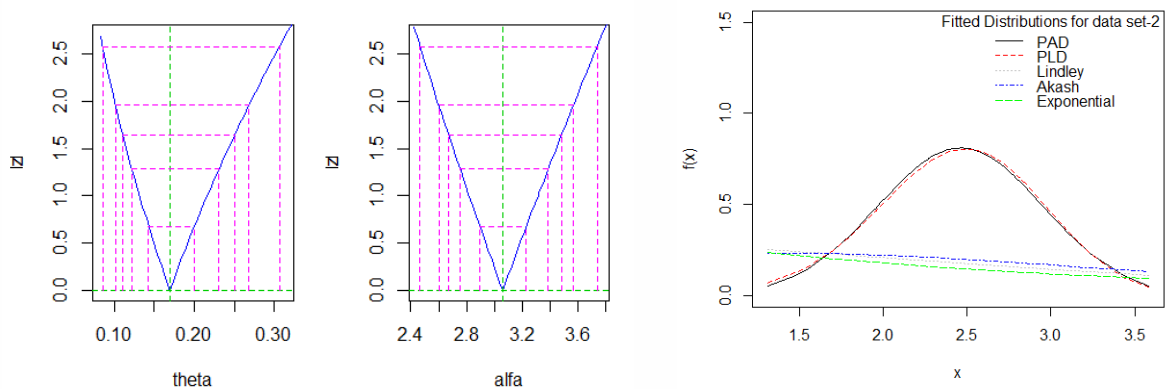


Figure 4. Likelihood estimate for parameters of PAD and Fitted distributions plots for the given data set

8. Concluding remarks

In this paper a two-parameter power Akash distribution (PAD), of which one parameter Akash distribution introduced by Shanker (2015) is a special case, has been proposed. Its important statistical properties including shapes of the density for varying values of parameters, hazard rate function, moments, skewness and kurtosis measures have been discussed. The maximum likelihood estimation has been discussed for estimating its parameters. The goodness of fit of the proposed distribution for a real lifetime data set from engineering has been discussed and it shows quite satisfactory over two-parameter PLD and one parameter Akash, Lindley and exponential distributions. Therefore, PAD can be considered an important lifetime distribution for modeling lifetime data from engineering.

9. References

1. Bader, M.G. and Priest, A. M. **Statistical aspects of fiber and bundle strength in hybrid composites**, in Hayashi, T., Kawata, K. and Umekawa, S. (Eds), "Progress in Science in Engineering Composites", ICCM-IV, Tokyo, 1982, pp. 1129–1136.
2. Ghitany, M.E., Al-Mutairi, D.K., Balakrishnan, N., and Al-Enezi, L.J. **Power Lindley distribution and Associated Inference**, Computational Statistics and Data Analysis, Vol. 64, 2013, pp. 20–33.
3. Lindley, D.V. **Fiducial distributions and Bayes' Theorem**, Journal of the Royal Statistical Society, Series B, Vol. 20, 1958, pp. 102–107.
4. Shanker, R. **Akash Distribution and Its Applications**, International Journal of Probability and Statistics, Vol. 4, No. 3, 2015, pp. 65–75.
5. Shanker, R. **The Discrete Poisson-Akash Distribution**, International Journal of Probability and Statistics, Vol. 6, No. 1, 2017, pp. 1–10.
6. Shanker, R., Hagos, F. and Sujatha, S. **On Modeling of Lifetime Data Using One Parameter Akash, Lindley and Exponential Distributions**, Biometrics & Biostatistics International Journal, Vol. 3, No. 2, 2016, pp. 1–10.

¹ Prof. Rama Shanker has completed his Bachelor, Master and Ph.D in Statistics from Department of Statistics, Patna University, Patna, India. He has more than twenty years of teaching and research experience. Presently, he is working as Professor and Head, Department of Statistics, College of Science, Eritrea Institute of Technology, Asmara, Eritrea. He is the founding head of Department of Statistics under College of Science at Eritrea Institute of technology, Asmara, Eritrea. He was also the founding Editor-in-Chief of "Eritrean Journal of Science and Engineering (EJSE)", a biannual Science and Engineering Journal, published from Eritrea Institute of Technology, Eritrea. His research interests include Distribution Theory, Modeling of Lifetime Data, Statistical Inference, Mathematical Demography, Transportation and Assignment Problems. He has more than 90 research papers published in national and international journals of Statistics, Mathematics and Operations research. He is the international advisory Board members of research journals published from India, Singapore, Turkey and USA and reviewers of research papers in distribution theory of many international journals of Statistics. He has been working as international advisory committee members of conferences/seminars in Statistics. He has delivered invited/ plenary talks in international/national conferences/seminars in Statistics and Statistics related disciplines. He is the external examiner of Ph.D thesis in Statistics for different universities.

² Dr. Kamlesh Kumar Shukla is presently working as Associate Professor at Department of Statistics, Eritrea Institute of Technology, Eritrea since February, 2016. He has been awarded Ph.D. in Statistics from Banaras Hindu University, Varanasi, India, and Masters Degree in Statistics (Gold Medalist) in 1997. Dr. Shukla has worked as Assistant Professor in Adama Science and Technology University, Ethiopia and having more than 14 years of teaching experience of colleges/ Universities including International experience. Author has worked on six projects in International and National organization Viz: IIPS, WHO, DST, New Delhi and presented many papers in International and National conferences. He has published more than 42 research papers in International and National Journal including one book entitled "Elements of Statistics" for BCA Students.