Original Article

Alpha-power power distribution

Malek Fathizadeh*, Abbas Mahdavi, Leila Jabbari

Department of Statistics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran

ARTICLE INFO

ABSTRACT

Background & Aim: Due to the applicability of the statistical distributions in many areas of sciences, adding parameters to an existing distribution for developing more flexible models have been overlooked in the statistical literatures.

Methods & Materials: A new generalization of power distribution is proposed using alpha power transformation method. The new distribution is more flexible than the power distribution and contains distributions that can be unimodal or right skewed.

Results: We study some statistical properties of the new distribution, including mean residual lifetime, quantiles, mode, moments, moment generating function, order statistics, some entropies and maximum likelihood estimators.

Conclusion: We fit the APP and some competitive models to one real data set and show that the new model has a superior performance among the compared distributions as evidenced by some goodness-of fit statistics.

Introduction

In recent years, many impressive families of new statistical distributions have been generated by statisticians. The necessity to develop an extended class of classical distributions is even more in areas such as survival data analysis, finance and risk modeling, insurance, biomedicine, modeling rare events etc.

The power distribution is used to analysis of environmental policy and public health Boyce et al. (1) and also in financial engineering domain Van Dorp and Kotz (2). The PDF and the cumulative distribution function (CDF) of power distribution is given by

$$f(x; a, b) = ab(bx)^{b-1}, \quad 0 < x < \frac{1}{b},$$

$$F(x; a, b) = (bx)^a,$$

respectively, where $a > 0$ is shape

---

* Corresponding Author: Malek Fathizadeh, Postal Address: Department of Statistics, Vali-e-Asr University of Rafsanjan, Rafsanjan, Iran.
Email: m.fathi@vru.ac.ir

parameter and $b > 0$ is scale parameter. Cordeiro et al. (3) introduced the beta power distribution (BP) which extends the power distribution defined by Balakrishnan and Nevzorov (4). The probability density function (PDF) of BP is

$$f(x; a, b, \alpha, \beta)$$
$$= \frac{ab(bx)^{\alpha a-1}[1 - (bx)^a]^{\beta-1}}{B(\alpha, \beta)}, 0 < x < \frac{1}{b},$$

where $B(\alpha, \beta)$ is the beta function. $\alpha > 0$, $\beta > 0$ and $a > 0$ are shape parameters and $b > 0$ is the scale parameter. It is observed that this four parameter BP distribution has several desirable properties and it can be produce better fit for the data.

Recently, Mahdavi and Kundu (5) introduced a new method to expand family of distributions by adding an extra parameter α, called α-power transformation method (APT). The aim of this paper is to introduce an extra parameter to the power distribution to bring more flexibility. We have used the APT method to the power distribution and generated a three-parameter α-power power (APP) distribution. Several properties of APP distribution have been established. Such that, the random samples

Please cite this article in press as: FathizadehM, Mahdavi A, Jabbari L. Alpha-power power distribution. *J B i o s t a t Epidemiol. 2018; 4(3): 129-135*

http://jbe.tums.ac.ir

generation of the APP is straight forward. It is shown that the distribution function, survival function, hazard rate function, moment generating function and the n-th moment can be obtained in closed form. The maximum likelihood estimators (MLEs) of unknown parameters can be obtained by solving three nonlinear equations. Finally, one real data set is used to illustrate the usefulness and applicability of the APP distribution.

Methods

Let X a continuous random variable with CDF and PDF, $F(x, \xi)$ and $f(x, \xi)$, respectively, and parameter vector ξ. Then the α-power transformation of X has the following CDF and PDF,

$$F_{APT}(x, \alpha, \xi) = \begin{cases} \dfrac{\alpha^{F(x,\xi)} - 1}{\alpha - 1} & if \quad \alpha > 1, \alpha \neq 1 \\ F(x, \xi) & if \qquad \alpha = 1, \end{cases} \quad (1)$$

and

$$f_{APT}(x, \alpha, \xi) = \begin{cases} \dfrac{\log \alpha}{\alpha - 1} f(x, \xi)\alpha^{F(x,\xi)} & if \quad \alpha > 1, \alpha \neq 1 \\ f(x, \xi) & if \qquad \alpha = 1, \end{cases}$$

In order to generate APP distribution, we take $F(x)$ as a CDF of power distribution in (1). Thus, the CDF of APP for $0 < y < 1/b$ yields as:

$$F(y; a, b, \alpha) = \begin{cases} \dfrac{\alpha^{(by)^a} - 1}{\alpha - 1} & if \quad \alpha \neq 1 \\ (by)^a & if \quad \alpha = 1, \end{cases}$$

and 0 otherwise. The PDF of APP for $0 < y < 1/b$, becomes:

$$f(y; a, b, \alpha) = \begin{cases} \dfrac{\log \alpha}{\alpha - 1} ab(by)^{a-1}\alpha^{(by)^a} & if \qquad , \alpha \neq 1 \\ ab(by)^{a-1} & if \qquad \alpha = 1, \end{cases}$$

and 0 otherwise. We denoted the random variable Y that follows APP distribution by $Y \sim APP(a, b, \alpha)$, where $\alpha > 0$ and $a > 0$ are the shape parameters and $b > 0$ is the scale parameter.
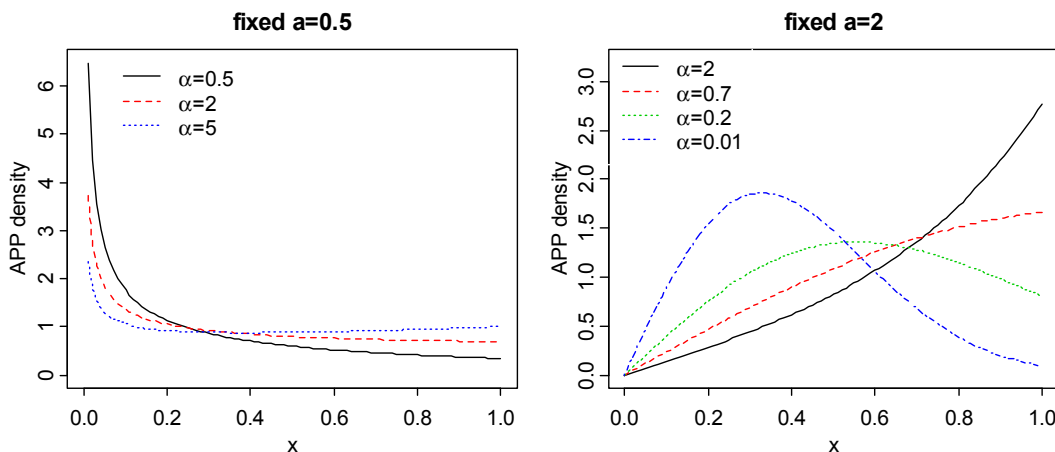
Results
Plots of the APP

Plots of the APP density function for selected choices of the parameters a, α and fixed scale parameter $b = 1$ are given in Figure 1. The PDF of the APP distribution can be either unimodal or right skewed. If $\alpha \leq 1$ and $0 < a \leq 1$, then it is a decreasing function of y, and for $0 < \alpha < 1$ and $a > 1$, it is an unimodal function.

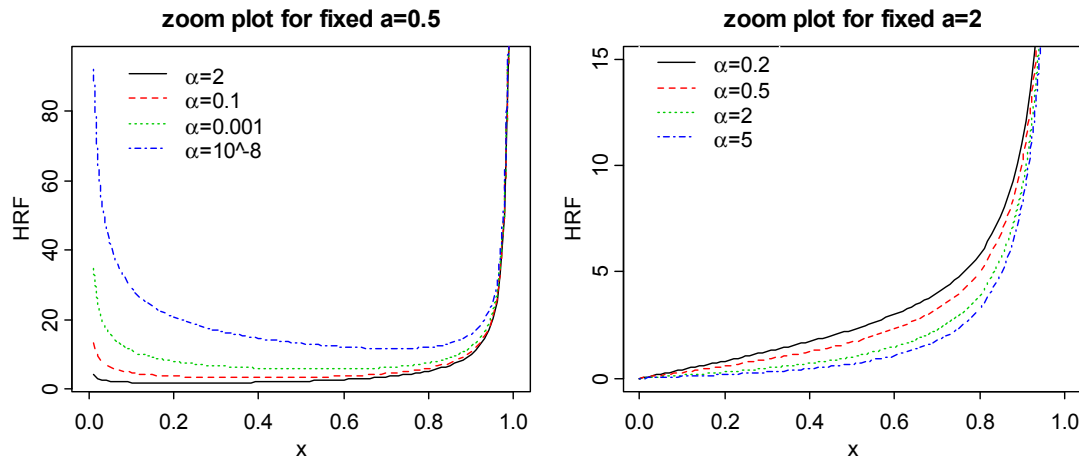The survival function, $S(y)$, and the HRF, $h(y)$, for APP distribution, are in the following forms for $0 < y < 1/b$

$$S(y; a, b, \alpha) = \begin{cases} \dfrac{\log \alpha}{\alpha - 1}(1 - \alpha^{(by)^a - 1}) & if \qquad , \alpha \neq 1 \\ 1 - (by)^a & if \qquad \alpha = 1, \end{cases}$$

and

$$h(y; a, b, \alpha) = \begin{cases} \left(\dfrac{\log \alpha}{\alpha}\right)\dfrac{ab(by)^{a-1}\alpha^{(by)^a}}{1 - \alpha^{(by)^a - 1}} & if \qquad , \alpha \neq 1 \\ \dfrac{ab(by)^{a-1}}{1 - (by)^a} & if \qquad \alpha = 1, \end{cases}$$

Figure 1. Density plots of the APP distribution with various shape parameters and fixed scale parameter $b = 1$.

Figure 2. The HRF of APP distribution with various shape parameters and fixed scale parameter $b = 1$.

Plots of the HRF for selected choices of the parameters a, α and fixed scale parameter $b = 1$ are given in Figure 2. If $a \geq 1$, then $h\ (y; a, b, \alpha)$ is an increasing function of y, and $h\ (y; a, b, \alpha)$ can be a bathtub shaped for $a < 1$, as α tends to 0. The failure rate function is an important quantity characterizing life phenomena. Almost all of the standard distributions in statistics do not exhibit a bathtub shape for hazard rate function. Even the traditional Weibull distribution does not exhibit a bathtub shape for hazard rate function. Thus, it is important that one knows which distributions exhibit this shape and most real-life systems exhibit bathtub shapes for their hazard rate functions.

Mean residual lifetime function

The mean residual lifetime (MRL) function of APP distribution is given by

$$m(x) = E(X - x | X > x)$$

$$= \frac{1}{b(\alpha - \alpha^{(bx)^a})} \left\{ \alpha(1 - bx) \right.$$

$$\left. - \sum_{k=0}^{\infty} \frac{(\log \alpha)^k \left(1 - (bx)^{ak+1}\right)}{k!\,(ak+1)} \right\}.$$

For $0 < \alpha < 1$ the MLR function can be written as

$$m(x)$$
$$= \frac{\gamma\left(\frac{1}{a}, -\log \alpha\right) - \gamma\left(\frac{1}{a}, -(bx)^a \log \alpha\right) - \alpha(1 - bx)}{b(\alpha^{(bx)^a} - \alpha)},$$

where $\gamma(s, t) = \int_0^t x^{s-1} e^{-x}\, dx$ is the incomplete gamma function. If $a \geq 1$, then the APP distribution has increasing hazard rate function and hence, decreasing MRL function. For $a < 1$ and α close to 0, the APP distribution has bathtub-shape hazard rate function and hence, upside-down bathtub-shape MRL function.

Mode, quantile and simulation

The Mode of APP distribution can be obtained using the following simple formula:

$$M = \frac{1}{b}\left(\frac{1 - a}{a \log \alpha}\right)^{\frac{1}{a}}.$$

The $p - th$ quantile of APP distribution is given by

$$y_p = \frac{1}{b}\left(\frac{\log\{(\alpha - 1)p + 1\}}{\log \alpha}\right)^{\frac{1}{a}}. \quad (2)$$

The median can be derive from (2) by considering $p = 0.5$ as

$$m = \frac{1}{b}\left(\frac{\log(\alpha + 1) - \log 2}{\log \alpha}\right)^{\frac{1}{a}}.$$

One of the advantages of the APP distribution is that its CDF has a closed form which helps us to generate random variables by using the following simple formula

$$X = \frac{1}{b}\left(\frac{\log\{(\alpha - 1)U + 1\}}{\log \alpha}\right)^{\frac{1}{a}},$$

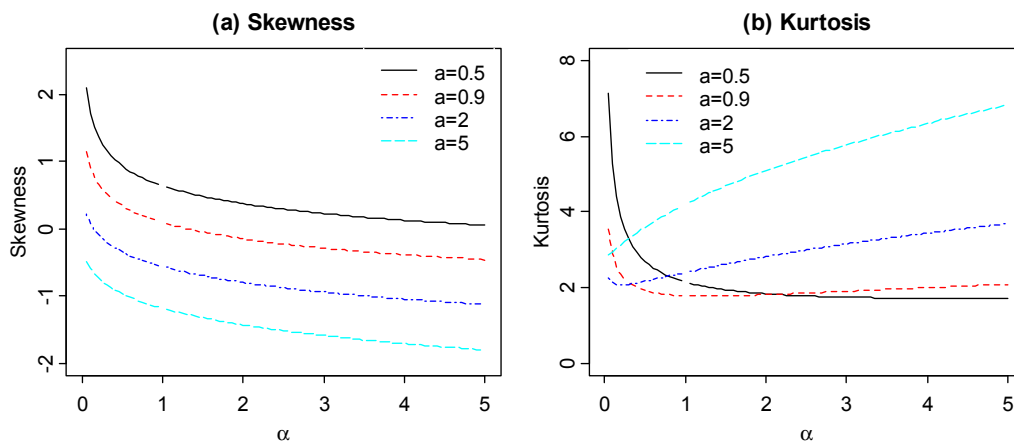where U is a uniformly distributed random variable on $(0, 1)$.

131

Figure 3. The skewness and kurtosis of APP distribution as a function of α for some values of a and fixed $b = 1$.

Moments and moment generating function

Let X a random variable follows $APP(a, b, \alpha)$. We derive infinite expansions for the n-th ordinary moment of X as

$$\mu_n = E(X^n)$$
$$= \frac{a}{b^n(\alpha - 1)} \sum_{k=0}^{\infty} \frac{(\log \alpha)^{k+1}}{k!\,(ak + a + n)}. \qquad (3)$$

If $0 < \alpha < 1$ then (3) can be expressed with a simpler formula as

$$\mu_n = E(X^n) = \frac{\gamma\left(\frac{n}{a} + 1, -\log \alpha\right)}{(1 - \alpha)b^n(-\log \alpha)^{\frac{n}{a}}}.$$

Based on the first four ordinary moments of the APP distribution, the measures of skewness (δ_1) and kurtosis (κ_1) of the APP distribution can obtained as

$$\delta_1 = \frac{\mu_3 - 3\mu_1\mu_2 + 2\mu_1^3}{(\mu_1 - \mu_1^2)^{\frac{3}{2}}},$$

and

$$\kappa_1 = \frac{\mu_4 - 4\mu_1\mu_3 + 6\mu_1^2\mu_2 - 3\mu_1^4}{(\mu_1 - \mu_1^2)^2}.$$

Figures 3 (a) and (b) show the skewness and kurtosis of APP distribution as a function of parameter α for different values of a and fixed scale parameter $b = 1$. It is observed that the skewness is a decreasing function of α and kurtosis first decrease as α increase and then start increasing.

The moment generating function (MGF) can be obtained by the following series expansion

$$M_X(t) = \sum_{n=0}^{\infty} \frac{t^n E(X^n)}{n!}.$$

So, the MGF of APP distribution is given by

$$M_X(t) = \frac{a}{\alpha - 1} \sum_{n=0}^{\infty} \sum_{k=0}^{\infty} \frac{(\log \alpha)^{k+1}}{b^n k!\,(ak + a + n)}.$$

Also, if $0 < \alpha < 1$ then the MGF can be obtained using the following formula:

$$M_X(t)$$
$$= \frac{1}{\alpha - 1} \sum_{k=0}^{\infty} \frac{\gamma\left(\frac{k}{a} + 1, -\log \alpha\right)}{k!} \left(\frac{t}{b(-\log \alpha)^{\frac{1}{a}}}\right)^k.$$

Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. We now give the PDF of the k-th order statistic $Y = X_{k:n}$ in a random sample of size n from the APP distribution as follows

$$f_Y(y) = \frac{n!}{(k-1)!\,(n-k)!} F^{k-1}(y)\{1 - F(y)\}^{n-k} f(y)$$
$$= \frac{n!\,ab \log \alpha}{(\alpha - 1)^n} (by)^{a-1} \sum_{i=0}^{k-1} \sum_{j=0}^{n-k} \frac{(-1)^{n-i-j-1}\alpha^j \alpha^{(1+n+i-k-j)}(by)^a}{(k-i-1)!\,(n-k-j)!\,i!\,j!}$$

Entropies

Entropy has been used in various situations in Science and Engineering. The entropy of a random variable X with PDF $f(x)$ is a measure of variation of the uncertainty. There exist many entropy definitions and they are not equally good for all applications. While the most famous (and most liberal) Shannon (6) Entropy, which quantifies the encoding length, is extremely useful in information theory. Shannon showed important applications of this entropy in communication theory and many applications have been used in different areas such as Engineering, Physics, Biology and Economics. The Shannon entropy for APP distribution is obtained as

132

$$E(-\log f(x)) = -\log\left(\frac{ab^a \log \alpha}{\alpha - 1}\right) +$$
$$\frac{1}{\alpha - 1}\sum_{k=0}^{\infty}(\log \alpha)^{k+1}\{$$
$$\frac{a - 1}{a(k + 1)(k + 1)!} + \frac{(a - 1)\log b}{(k + 1)!}$$
$$-\frac{a \log \alpha}{k!\,(ka + 2a)}\}.\quad (4)$$

A generalized definition of entropy that stems from modifying the additivity postulate and results in a class of information measures that contain Shannons definitions as special cases is Rényi (7) entropy. If X has the PDF f(x) then Rényi entropy is defined by

$$I_R(\rho) = \frac{1}{1 - \rho}\log\left\{\int f(x)^\rho dx\right\}$$

where $\rho > 0$ and $\rho \neq 1$. Using (4), the integral in $I_R(\rho)$ for the APP distribution can be reduced to

$$\int_0^\infty f(x)^\rho dx$$
$$= \left(\frac{ab \log \alpha}{\alpha - 1}\right)^\rho \sum_{k=0}^{\infty}\frac{(\rho \log \alpha)^k}{k!\,b(ka + a\rho - \rho + 1)}.$$

So, one obtains the Rényi entropy as
$$I_R(\rho)$$
$$= \frac{1}{1 - \rho}\left\{\rho \log\left(\frac{ab \log \alpha}{\alpha - 1}\right)\right.$$
$$\left. + \log\left(\sum_{k=0}^{\infty}\frac{(\rho \log \alpha)^k}{k!\,b(ka + a\rho - \rho + 1)}\right)\right\}.$$

Discussion
Maximum likelihood estimation

We now determine the MLEs of the parameters of the APP distribution from complete samples only. Let $x_1, x_2, \ldots, x_n$ be a sample from $APP(a, b, \alpha)$ distribution. Then, the log-likelihood function is

$$\log L = n \log a + na \log b + n \log\left(\frac{\log \alpha}{\alpha - 1}\right)$$
$$+ (a - 1)\sum_{i=1}^{n}\log x_i$$
$$+ b^a \log \alpha \sum_{i=1}^{n}x_i^a.$$
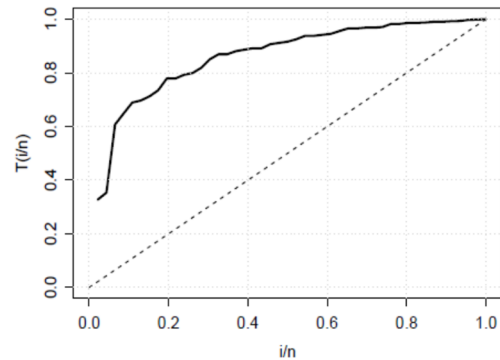
The first order derivatives of log L are

Figure 4. Empirical TTT plot for the glass fibres data

$$\frac{\partial \log L}{\partial a} = \frac{n}{a} + n \log b + \sum_{i=1}^{n}\log x_i$$
$$+ b^a \log \alpha\left(\log b \sum_{i=1}^{n}x_i^a\right.$$
$$\left. + \sum_{i=1}^{n}x_i^a \log x_i\right),$$
$$\frac{\partial \log L}{\partial b} = \frac{na}{b} + ab^{a-1} + \log \alpha \sum_{i=1}^{n}x_i^a,$$

and

$$\frac{\partial \log L}{\partial \alpha} = \frac{b^a}{\alpha}\sum_{i=1}^{n}x_i^a + n\frac{\alpha - 1 - \alpha \log \alpha}{\alpha(\alpha - 1)\log \alpha}.$$

The MLEs of the unknown parameters cannot be obtained explicitly. They have to be obtained by solving some numerical methods, like Newton-Raphson or Gauss Newton methods or their variants. In this paper we use the optim function from the statistical software R (R Core Team, (8)) to maximize the logarithm of the likelihood function. This Fisher information matrix is important because it can be used to construct asymptotic confidence intervals for the parameters from its estimates. Let $\theta = (a, b, \alpha)'$ and $\hat{\theta} = (\hat{a}, \hat{b}, \hat{\alpha})'$ denote the vector of parameters and its respective estimates. Under regularity conditions, the asymptotic distribution of $\hat{\theta}$ is given by

$$\sqrt{n}(\hat{\theta} - \theta) \sim N_3(0, \mathbf{I}^{-1}(\theta)),$$

where $\mathbf{I}(\theta)$ is expected information matrix. This asymptotic behavior is valid if $\mathbf{I}(\theta)$ is replaced by $\mathbf{J}(\hat{\theta})$ where $\mathbf{J}(\hat{\theta})$ is the observed information matrix evaluated at $\hat{\theta}$. The asymptotic multivariate normal $N_3(0, \mathbf{J}^{-1}(\theta))$ distribution can be used to construct approximate confidence intervals for the individual parameters and for the hazard rate and survival functions. Package numDeriv of R language can

133

be used to compute the Hessian matrix and its inverse, standard errors and asymptotic confidence intervals.

Application: glass fibres data

The following 46 data points represent the strengths of 15 cm glass fibres, originally obtained by workers at the UK National Physical Laboratory. The data set is obtained from Smith and Naylor (9). The sorted data are given as follows: 0.37 0.40 0.70 0.75 0.80 0.81 0.83 0.86 0.92 0.92 0.94 0.95 0.98 1.03 1.06 1.06 1.08 1.09 1.10 1.10 1.13 1.14 1.15 1.17 1.20 1.20 1.21 1.22 1.25 1.28 1.28 1.29 1.29 1.30 1.35 1.35 1.37 1.37 1.38 1.40 1.40 1.42 1.43 1.51 1.53 1.61.

In order to identify the shape of the hazard function, we shall consider a graphical method based on the Total Time on Test (TTT) plot. In its empirical version the TTT plot is given by

$$T(r/n) = \frac{\sum_{i=1}^{r} Y_{i:n} + (n-r)Y_{r:n}}{\sum_{i=1}^{r} Y_{i:n}}$$

where $r = 1, 2, \cdots, n$ and $Y_{i:n}$ represents the i-th order statistic of the sample. If the empirical

TTT transform is convex, concave, first convex then concave, and first concave then convex, the shape of the corresponding hazard rate function is, respectively, decreasing, increasing, bathtub, and unimodal. (For more details, see Aarset, (10)). Figure 4 shows the empirical TTT plots for the glass fibres data, which is concave indicating an increasing failure rate function, which can be properly accommodated by APP distribution. We compare APP distribution with power distribution, BP distribution and Transmuted Power Function distribution (T-Ps) introduced by ul Haq et al.

11) with PDF

$$f(x) = \frac{\alpha x^{\alpha}}{\beta^{\alpha}} \left\{ 1 + \theta - 2\theta \left(\frac{x}{\beta}\right)^{\alpha} \right\},$$
$$0 < x < \beta, \alpha > 0, \beta > 0, -1 < \theta < 1.$$

To see which one of these models is more appropriate to fit data. The MLEs of parameters, Akaike Information Criterion (AIC) value, Bayesian Information Criterion (BIC) value, Kolmogorov-Smirnov (K-S) statistic and its p-value are given in Table 1. The APP distribution
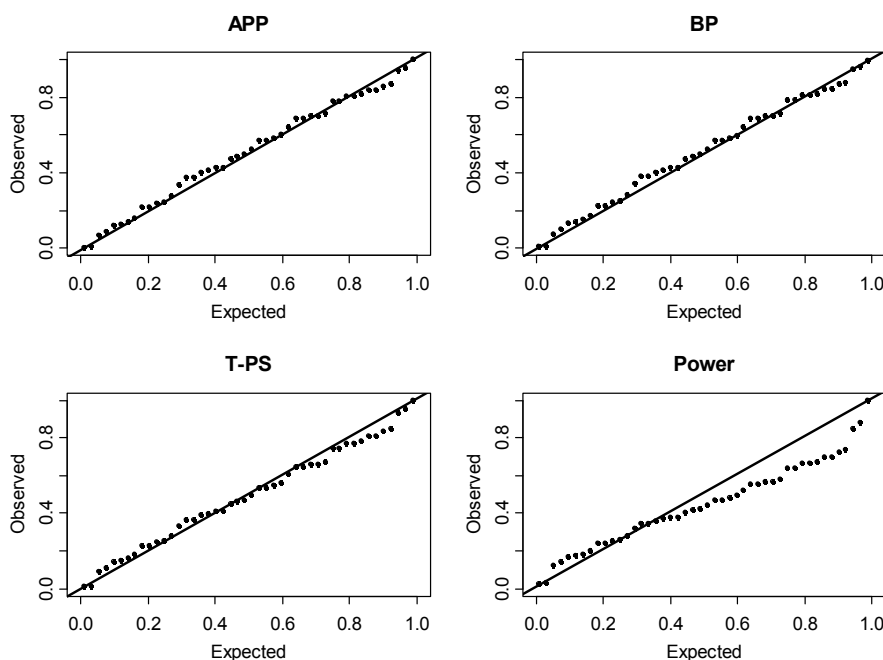

Figure 5. Probability plots for the fitted distributions

Table 1. The maximum likelihood estimates, AIC, BIC and K-S and its p-values for glass fibers data.

The model	MLEs of the parameters	AIC	BIC	K-S statistic	p-value
APP	$\hat{a} = 4.6523, \hat{b} = 0.6211, \hat{\alpha} = 0.0455$	10.1961	15.68202	0.0699	0.9872
BP	$\hat{a} = 3.5953, \hat{b} = 0.5996, \hat{\alpha} = 1.17, \hat{\beta} = 2.6908$	12.7105	20.0251	0.0701	0.9770
T-Ps	$\hat{a} = 3.6454, \hat{b} = 1.61, \hat{\theta} = 0.8614$	11.0284	16.5143	0.0895	0.8548
Power	$\hat{a} = 2.5565, \hat{b} = 0.6211$	17.4726	21.1299	0.1963	0.0578

gives the smallest AIC, BIC and K-S statistics. The probability plots for the fitted distributions are plotted in Figure 5. It is clear from Table 1 and Figure 4 that the APP model provides the best fits to the data.

Conclusion

In this paper, we proposed a new family of distributions called alpha-power power distribution (APP), by applying the alpha-power transformation method initially proposed by Mahdavi and Kundu (5) to the classical power distribution. The HRF of the APP distribution can be a bathtub shape that is important quantity characterizing life phenomena. Various properties of the new distribution are obtained. These properties include moments, quantiles, mode, moment generating function, order statistics and some entropies. We dis-cuss maximum likelihood estimation of the model parameters and derive the observed information matrix. An application of the APP distribution is demonstrated in a real data set. It provides a better fit than other competing models.

Conflicts of interests

The authors declare that there is no conflict of interest regarding the publication of this article.

References

1. Boyce JK, Klemer AR, Templet PH, Willis CE. Power distribution, the environment, and public health: A state-level analysis. Ecologic Econom. 1999;29(1):127–140.

2. Van Dorp JR, Kotz S. The standard two-sided power distribution and its properties: with applications in financial engineering. Am Stat. 2002;56(2):90–99.

3. Cordeiro GM, dos Santos Brito R. The beta power distribution. Braz J Probabil Stat. 2012; 26(1):88–112.

4. Balakrishnan N, Nevzorov V. A primer on statistical distributions. Hoboken, New Jersey: A John Wiley & Sons; 2003.

5. Mahdavi A, Kundu D. A new method for generating distributions with an application to exponential distribution. Commun Stat Theory Method. 2017;46(13):6543–6557.

6. Shannon CE. Prediction and entropy of printed English. Bell System Technic J. 1951;30(1):50–64.

7. Rényi A. On measures of entropy and information. In Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics. The Regents of the University of California; 1961.

8. R Core, Team R. A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria; 2017.

9. Smith RL, Naylor JA. Comparison of maximum likelihood and bayesian estimators for the three-parameter weibull distribution. Appl Stat. 1987; 358–369.

10. Aarset MV. How to identify a bathtub hazard rate. IEEE Transact Reliabil. 1987;36(1):106–108.

11. Ul Haq MA, Butt NS, Usman RM, Fattah AA. Transmuted power function distribution. Gazi Uni J Sci. 2016;29(1):177–185.