MedCrave
Step into the World of Research

Biometrics & Biostatistics International Journal

Research Article

Open Access

CrossMark

A two–parameter Sujatha distribution

Abstract

This paper proposes a two–parameter Sujatha distribution (TPSD). This includes size–biased Lindley distribution and Sujatha distribution as particular cases. Its important statistical properties including shapes of density function for varying values of parameters, coefficient of variation, skewness, kurtosis, index of dispersion, hazard rate function, mean residual life function, stochastic ordering ,mean deviations, Bonferroni and Lorenz curves, and stress–strength reliability have been discussed. The estimation of parameters has been discussed using the method of moments and the method of maximum likelihood. Application of the distribution has been discussed with a real lifetime data.

Keywords: Sujatha distribution, moments, statistical properties, estimation of parameters, application

Volume 7 Issue 3 - 2018

Mussie Tesfay, Rama Shanker

Department of Statistics, Eritrea Institute of Technology Asmara, Eritrea

Correspondence: Rama Shanker, Department of Statistics, College of Science, Eritrea Institute of Technology, Asmara, Eritrea, Email shankerrama2009@gmail.com, mostes1106@gmail.com

Received: March 15, 2018 | **Published:** May 10, 2018

Introduction

The statistical analysis and modeling of lifetime data are crucial for statisticians working in various field of knowledge including medical science, engineering, social science, behavioral science, insurance, finance, among others. The classical one parameter lifetime distribution in statistics which were popular for modeling lifetime data are exponential distribution and Lindley distribution proposed by Lindley.[1] Shanker *et al*.[2] have detailed critical study on applications of exponential and Lindley distributions for modeling lifetime data from engineering and biomedical science and observed that exponential and Lindley distributions are not always suitable due to theoretical or applied point of view and presence of single parameter. In search for a lifetime distribution which gives a better fit than exponential and Lindley distributions, Shanker[3] has proposed a new lifetime distribution named Sujatha distribution defined by its probability density function (pdf) and cumulative distribution function (cdf).

$$f_1(x;\theta)=\frac{\theta^3}{\theta^2+\theta+2}\left(1+x+x^2\right)e^{-\theta x};\ x>0,\theta>0 \qquad (1.1)$$

$$F_1\left(x;\theta\right)=1-\left[1+\frac{\theta x\left(\theta x+\theta+2\right)}{\theta^2+\theta+2}\right]e^{-\theta x};\ x>0,\theta>0 \qquad (1.2)$$

where θ is a scale parameter. It has been shown by Shanker[3] that Sujatha distribution is a convex combination of exponential (θ) distribution, a gamma (2, θ) distribution and a gamma (3, θ) distribution. The first four moments about origin and central moments of Sujatha distribution obtained by Shanker[3] are

$$\mu_1'=\frac{\theta^2+2\theta+6}{\theta\left(\theta^2+\theta+2\right)} \qquad \mu_2'=\frac{2\left(\theta^2+3\theta+12\right)}{\theta^2\left(\theta^2+\theta+2\right)}$$

$$\mu_3'=\frac{6\left(\theta^2+4\theta+20\right)}{\theta^3\left(\theta^2+\theta+2\right)} \qquad \mu_4'=\frac{24\left(\theta^2+5\theta+30\right)}{\theta^4\left(\theta^2+\theta+2\right)}$$

$$\mu_2=\frac{\theta^4+4\theta^3+18\theta^2+12\theta+12}{\theta^2\left(\theta^2+\theta+2\right)^2}$$

$$\mu_3=\frac{2\left(\theta^6+6\theta^5+36\theta^4+44\theta^3+54\theta^2+36\theta+24\right)}{\theta^3\left(\theta^2+\theta+2\right)^3}$$

$$\mu_4=\frac{3\left(3\theta^8+24\theta^7+172\theta^6+376\theta^5+736\theta^4+864\theta^3+912\theta^2+480\theta+240\right)}{\theta^4\left(\theta^2+\theta+2\right)^4}$$

Shanker[3] has discussed its important properties including shapes of density function for varying values of parameter, hazard rate function, mean residual life function, stochastic ordering, mean deviations, Bonferroni and Lorenz curves, and stress–strength reliability. Shanker[3] discussed the maximum likelihood estimation of parameter and showed applications of Sujatha distribution to model lifetime data from biomedical science and engineering. Shanker[4] has introduced Poisson–Sujatha distribution (PSD), a Poisson mixture of Sujatha distribution, and studied its properties, estimation of parameter and applications to model count data. Shanker & Hagos[5] have discussed zero–truncated Poisson– Sujatha distribution (ZTPSD) and applications for modeling count data excluding zero counts. Shanker & Hagos[6] have also studied size–biased Poisson– Sujatha distribution and its applications for count data excluding zero counts.

The Lindley distribution and a size–biased Lindley distribution (SBLD) having parameter θ are defined by their pdf

$$f_2(x;\theta)=\frac{\theta^2}{\theta+1}\left(1+x\right)e^{-\theta x};\ x>0,\theta>0 \qquad (1.3)$$

$$f_3(x;\theta)=\frac{\theta^3}{\theta+2}\,x\left(1+x\right)e^{-\theta x};\ x>0,\theta>0 \qquad (1.4)$$

Ghitany *et al*.[7] have discussed various statistical and mathematical properties, estimation of parameter and application of Lindley distribution to model waiting time data in a bank and it has been showed that Lindley distribution provides better fit than exponential distribution.

In this paper, a two–parameter Sujatha distribution (TPSD), which includes size–biased Lindley distribution and Sujatha distribution as particular cases, has been proposed. Its important statistical properties including coefficient of variation, skewness, kurtosis, index of dispersion, hazard rate function, mean residual life function, stochastic ordering, mean deviations, Bonferroni and Lorenz curves, stress–strength reliability have been discussed. The estimation of the parameters has been discussed using method of moments and maximum likelihood estimation. A numerical example has been given to test the goodness of fit of TPSD over Lindley and Sujatha distributions.

A two–parameter Sujatha distribution

A Two parameter Sujatha distribution (TPSD) having parameters θ and α is defined by its pdf

$$f_4(x;\theta,\alpha)=\frac{\theta^3}{\alpha\theta^2+\theta+2}\left(\alpha+x+x^2\right)e^{-\theta x};x>0,\theta>0,\alpha\geq 0 \quad (2.1)$$

where θ is a scale parameter and is α is a shape parameter. It can be easily verified that (2.1) reduces to Sujatha distribution (1.1) and SBLD (1.4) for $\alpha=1$ and $\alpha=0$ respectively.

Like Sujatha distribution (1.1), TPSD (2.1) is also a convex combination of exponential (θ), gamma (2, θ) and gamma (3, θ) distributions. We have

$$f_4\left(x;\theta,\alpha\right)=p_1\,g_1\left(x,\theta\right)+p_2\,g_2\left(x,\theta\right)+\left(1-p_1-p_2\right)g_3\left(x,\theta\right) \quad (2.2)$$

where

$$p_1=\frac{\alpha\theta^2}{\alpha\theta^2+\theta+2},\ p_2=\frac{\theta}{\alpha\theta^2+\theta+2},\ g_1\left(x,\theta\right)=\theta e^{-\theta x};x>0,\theta>0$$

$$g_2\left(x,\theta\right)=\frac{\theta^2}{\Gamma(2)}e^{-\theta x}x^{2-1};x>0,\theta>0,\ g_3\left(x,\theta\right)=\frac{\theta^3}{\Gamma(3)}e^{-\theta x}x^{3-1};x>0,\theta>0.$$

The corresponding cdf of TPSD (2.1) can be obtained as

$$F_2\left(x;\theta,\alpha\right)=1-\left[1+\frac{\theta x\left(\theta x+\theta+2\right)}{\alpha\theta^2+\theta+2}\right]e^{-\theta x};x>0,\theta>0,\alpha\geq 0 \quad (2.3)$$

Behavior of the pdf and the cdf of TPSD for varying values of parameters θ and α shown in Figures 1 & 2 respectively.

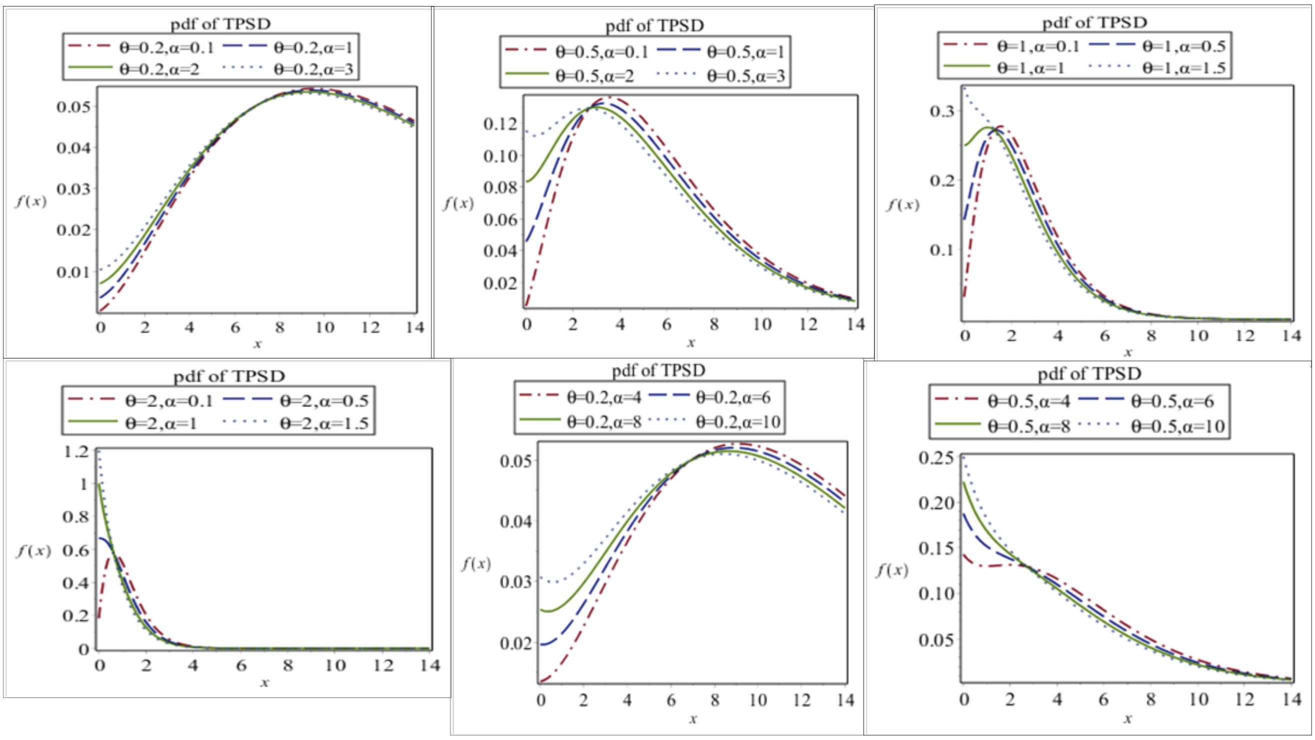

Figure 1 Behavior of the pdf of TPSD for varying values of parameters θ and α.

Moments and related measures

The moment generating function of TPSD (2.1) can be obtained as

$$M_X\left(t\right)=\frac{\theta^3}{\alpha\theta^2+\theta+2}\int_0^\infty e^{-(\theta-t)x}\left(\alpha+x+x^2\right)dx$$

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

$$=\frac{\theta^3}{\alpha\theta^2+\theta+2}\left[\frac{\alpha}{\theta}\sum_{k=0}^{\infty}\left(\frac{t}{\theta}\right)^k+\frac{1}{\theta^2}\sum_{k=0}^{\infty}\binom{k+1}{k}\left(\frac{t}{\theta}\right)^k+\frac{2}{\theta^3}\sum_{k=0}^{\infty}\binom{k+2}{k}\left(\frac{t}{\theta}\right)^k\right]$$

Citation: Tesfay M, Shanker R. A two–parameter Sujatha distribution. *Biom Biostat Int J.* 2018;7(3):188–197. DOI: 10.15406/bbij.2018.07.00208

$$= \sum_{k=0}^{\infty} \frac{\alpha\theta^2 + (k+1)\theta + (k+1)(k+2)}{\alpha\theta^2 + \theta + 2}\left(\frac{t}{\theta}\right)^k .$$

Thus, the r[th] moment about origin of TPSD (2.1), obtained as the coefficient of $\frac{t^r}{r!}$ in $M_X(t)$, is given by

$$\mu_r^{/} = \frac{r!\left\{\alpha\theta^2 + (r+1)\theta + (r+1)(r+2)\right\}}{\theta^r\left(\alpha\theta^2 + \theta + 2\right)} \quad ; r = 1, 2, 3, ... \qquad (3.1)$$

The first four moments about origin of TPSD are obtained as

$$\mu_1^{/} = \frac{\alpha\theta^2 + 2\theta + 6}{\theta\left(\alpha\theta^2 + \theta + 2\right)} \qquad \mu_2^{/} = \frac{2(\alpha\theta^2 + 3\theta + 12)}{\theta^2\left(\alpha\theta^2 + \theta + 2\right)}$$

$$\mu_3^{/} = \frac{6(\alpha\theta^2 + 4\theta + 20)}{\theta^3\left(\alpha\theta^2 + \theta + 2\right)} \qquad \mu_4^{/} = \frac{24(\alpha\theta^2 + 5\theta + 30)}{\theta^4\left(\alpha\theta^2 + \theta + 2\right)}$$

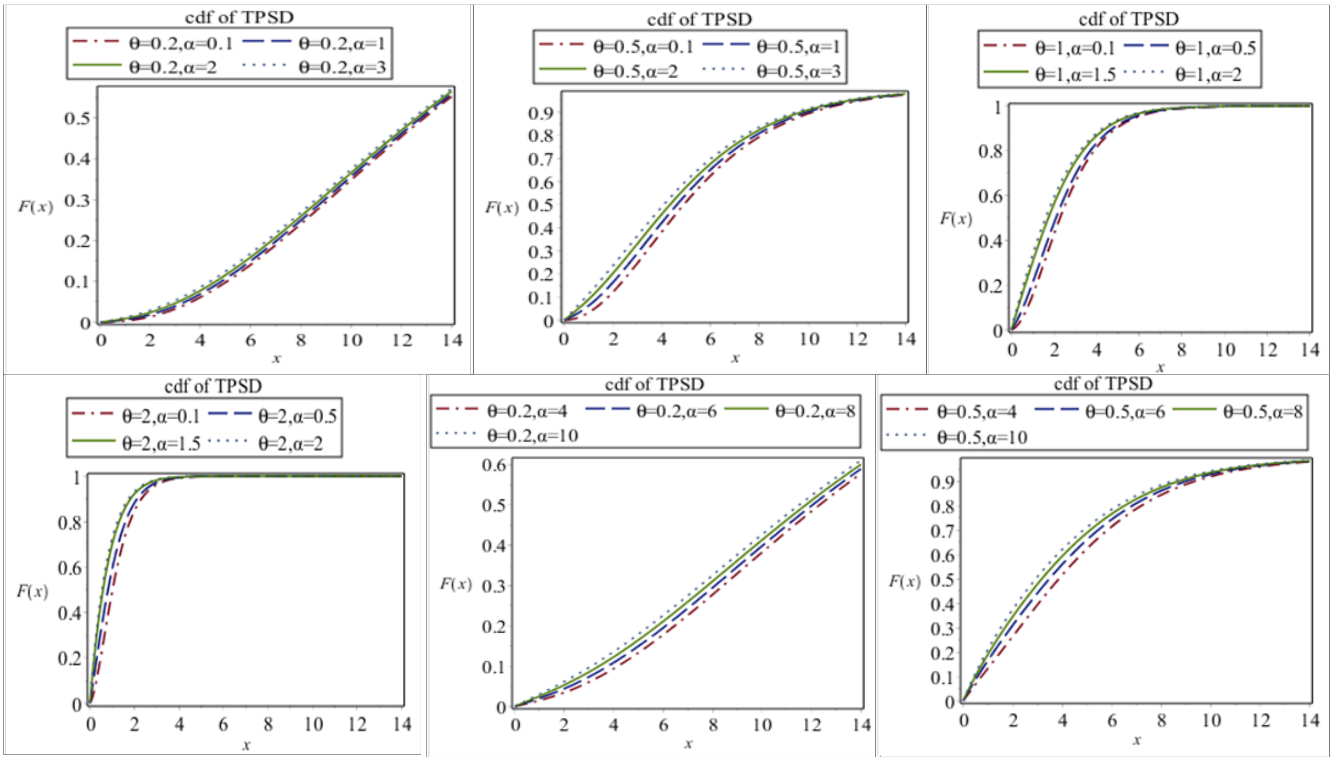

Figure 2 Behavior of the cdf of TPSD for varying values of parameter θ and α .

Using the relationship between moments about the mean and moments about the origin, the moments about mean of TPSD are obtained as

$$\mu_2 = \frac{\alpha^2\theta^4 + 4\alpha\theta^3 + 16\alpha\theta^2 + 2\theta^2 + 12\theta + 12}{\theta^2\left(\alpha\theta^2 + \theta + 2\right)^2}$$

$$\mu_3 = \frac{2\left(\alpha^3\theta^6 + 6\alpha^2\theta^5 + 30\alpha^2\theta^4 + 6\alpha\theta^4 + 42\alpha\theta^3 + 36\alpha\theta^2 + 2\theta^3 + 18\theta^2 + 36\theta + 24\right)}{\theta^3\left(\alpha\theta^2 + \theta + 2\right)^3}$$

$$\mu_4 = \frac{3\left(\begin{array}{l} 3\alpha^4\theta^8 + 24\alpha^3\theta^7 + 128\alpha^3\theta^6 + 44\alpha^2\theta^6 + 344\alpha^2\theta^5 + 408\alpha^2\theta^4 + 32\alpha\theta^5 + 320\alpha\theta^4 \\ +768\alpha\theta^3 + 8\theta^4 + 576\alpha\theta^2 + 96\theta^3 + 336\theta^2 + 480\theta + 240 \end{array}\right)}{\theta^4\left(\alpha\theta^2 + \theta + 2\right)^4}$$

The coefficient of variation (C.V), coefficient of skewness $\left(\sqrt{\beta_1}\right)$, coefficient of kurtosis $\left(\beta_2\right)$ and index of dispersion $\left(\gamma\right)$ of TPSD are given by

$$C.V = \frac{\sigma}{\mu_1^{'}} = \frac{\sqrt{\alpha^2\theta^4 + 4\alpha\theta^3 + 16\alpha\theta^2 + 2\theta^2 + 12\theta + 12}}{\alpha\theta^2 + 2\theta + 6}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2\left(\alpha^3\theta^6 + 6\alpha^2\theta^5 + 30\alpha^2\theta^4 + 6\alpha\theta^4 + 42\alpha\theta^3 + 36\alpha\theta^2 + 2\theta^3 + 18\theta^2 + 36\theta + 24\right)}{\left(\alpha^2\theta^4 + 4\alpha\theta^3 + 16\alpha\theta^2 + 2\theta^2 + 12\theta + 12\right)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3\left(\begin{array}{l} 3\alpha^4\theta^8 + 24\alpha^3\theta^7 + 128\alpha^3\theta^6 + 44\alpha^2\theta^6 + 344\alpha^2\theta^5 + 408\alpha^2\theta^4 + 32\alpha\theta^5 + 320\alpha\theta^4 \\ +768\alpha\theta^3 + 8\theta^4 + 576\alpha\theta^2 + 96\theta^3 + 336\theta^2 + 480\theta + 240 \end{array}\right)}{\left(\alpha^2\theta^4 + 4\alpha\theta^3 + 16\alpha\theta^2 + 2\theta^2 + 12\theta + 12\right)^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1^{'}} = \frac{\alpha^2\theta^4 + 4\alpha\theta^3 + 16\alpha\theta^2 + 2\theta^2 + 12\theta + 12}{\theta\left(\alpha\theta^2 + \theta + 2\right)\left(\alpha\theta^2 + 2\theta + 6\right)}$$

It can be easily verified that these statistical constants of TPSD reduce to the corresponding statistical constants of Sujatha distribution and SBLD at $\alpha = 1$ and $\alpha = 0$ respectively. To study the behavior of C.V., $\sqrt{\beta_1}$, β_2 and γ , their values for varying values of the parameters θ and α have been computed and presented in Tables 1–4.

Table 1 CV of TPSD for varying values of parameters θ and α

α \ θ	0.2	0.5	1	2	3	4	5
0.2	0.59658	0.624798	0.668399	0.739814	0.792609	0.8317	0.861102
0.5	0.599565	0.639569	0.708329	0.816497	0.882958	0.922627	0.946881
1	0.604466	0.662392	0.761739	0.892143	0.95119	0.977525	0.989835
2	0.614004	0.702377	0.83666	0.96225	0.996661	1.005655	1.007547
3	0.623205	0.736304	0.886072	0.991701	1.009814	1.011382	1.009973
4	0.632091	0.765466	0.920447	1.0059	1.014222	1.012415	1.009836
5	0.640678	0.790787	0.945247	1.013246	1.015576	1.012149	1.009163

For a given value of α , C.V increases as the value of θ increases .But for values $3 \le \alpha \le 5$, and $4 \le \theta \le 5$ CV decreases as the value of θ increases.

Table 2 Coefficient of skewness $\left(\sqrt{\beta_1}\right)$ of TPSD for varying values of parameters θ and α

α \ θ	0.2	0.5	1	2	3	4	5
0.2	1.156092	1.164414	1.193838	1.288579	1.40832	1.544566	1.694179
0.5	1.151692	1.153618	1.202728	1.394848	1.600302	1.785072	1.947347
1	1.145006	1.145839	1.247611	1.535588	1.733747	1.848046	1.912879
2	1.133828	1.153526	1.352316	1.647373	1.698737	1.653874	1.586127
3	1.125191	1.176753	1.43637	1.643895	1.562899	1.429794	1.310578
4	1.118703	1.206238	1.496066	1.59473	1.421347	1.244821	1.108
5	1.114041	1.237609	1.535958	1.528385	1.293862	1.097469	0.956984

Since $\sqrt{\beta_1} > 0$, TPSD is always positively skewed, and this means that TPSD is a suitable model for positively skewed lifetime data.

Table 3 Coefficient of kurtosis $\left(\beta_2\right)$ of TPSD for varying values of parameters θ and α

α \ θ	0.2	0.5	1	2	3	4	5
0.2	5.003116	5.022048	5.093943	5.346882	5.661781	5.979645	6.275987
0.5	4.991667	4.984856	5.082378	5.625	6.28542	6.865586	7.326691
1	4.973635	4.944566	5.170213	6.21499	7.193906	7.868405	8.297711
2	4.94128	4.924032	5.510204	7.2144	8.270528	8.774988	9.001011
3	4.913483	4.956867	5.903269	7.900925	8.799663	9.113171	9.206366
4	4.889821	5.022933	6.283795	8.3676	9.077558	9.253624	9.271966
5	4.869916	5.109996	6.633262	8.690336	9.230612	9.313262	9.289003

Since $\beta_2 > 3$, TPSD is always leptokurtic, and this means that TPSD is more peaked than the normal curve. Thus TPSD is suitable for lifetime data which are leptokurtic.

Table 4 Index of dispersion (γ) of TPSD for varying values of parameters θ and α

	θ						
α	0.2	0.5	1	2	3	4	5
0.2	5.164536	2.158531	1.144817	0.615741	0.424979	0.323306	0.259524
0.5	5.197861	2.220551	1.218487	0.666667	0.451356	0.334416	0.262078
1	5.252329	2.31348	1.305556	0.696429	0.452381	0.325758	0.251067
2	5.357375	2.466667	1.4	0.694444	0.431884	0.306064	0.235088
3	5.457478	2.585608	1.439394	0.676136	0.414263	0.293608	0.2264
4	5.552914	2.678571	1.452381	0.657692	0.401423	0.285531	0.221109
5	5.643939	2.751515	1.451923	0.641667	0.39193	0.279936	0.217569

As long as $0 \le \theta \le 1$ and $0 \le \alpha \le 5$, the nature of TPSD is over dispersed $\left(\sigma^2 > \mu_1'\right)$ and for $1 \le \theta \le 5$ and $0 \le \alpha \le 5$, the nature of TPSD is over dispersed $\left(\sigma^2 < \mu_1'\right)$.

The behavior of C.V., $\sqrt{\beta_1}$, β_2 and γ , for selected values of the parameters θ and α are shown in Figure 3.

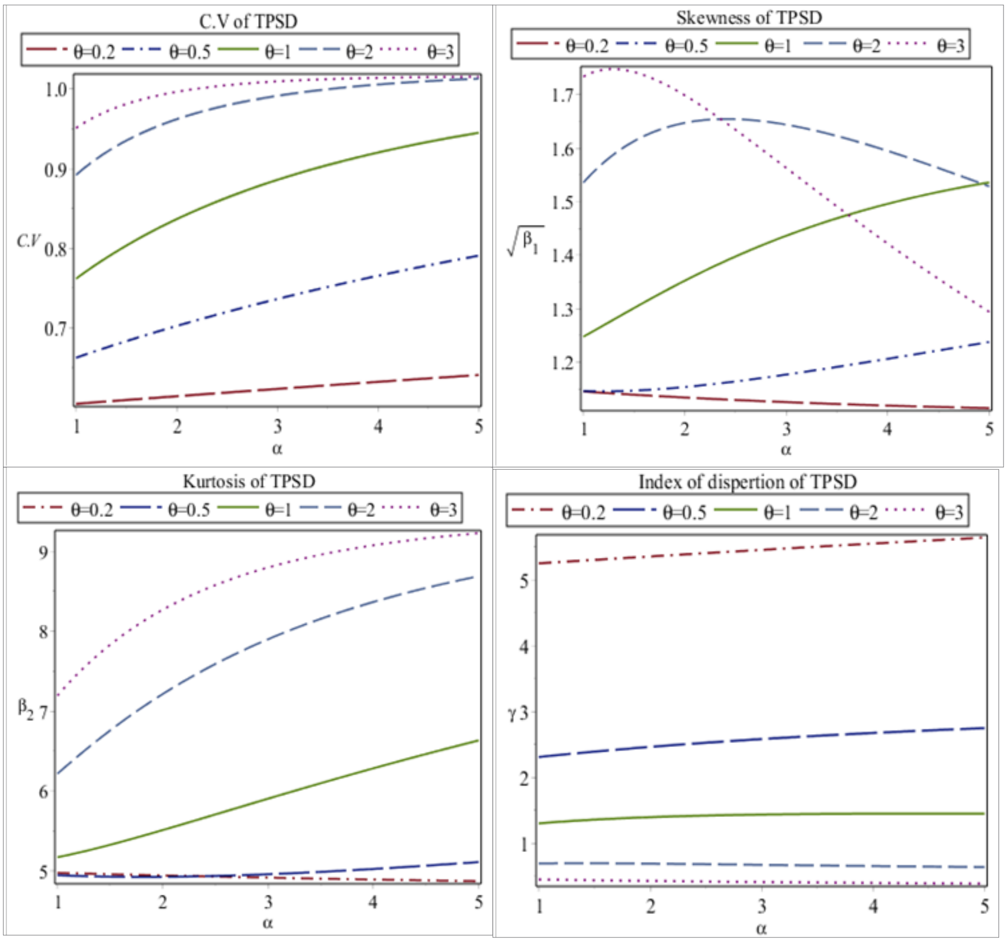

Figure 3 Behavior of C.V.,$\sqrt{\beta_1}$,β_2 and γ , for varying values of the parameters θ and α.

Statistical properties

In this section, statistical properties of TPSD including hazard rate function, mean residual life function, stochastic ordering, mean deviation, Bonferroni and Lorenz curves and stress–strength reliability have been discussed.

Hazard rate function and mean residual life function

Let X be a continuous random variable with pdf $f(x)$ and cdf $F(x)$ .The hazard rate function (also known as failure rate function), $h(x)$ and the mean residual function, $m(x)$ of X are respectively defined as

$$h(x) = \lim_{\Delta x \to 0} \frac{p(x \le X < x + \Delta x \mid X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)}$$

and $m(x) = E[X - x \mid X > x] = \dfrac{1}{1 - F(x)} \int_x^\infty [1 - F(t)]dt$

The corresponding hazard rate function, $h(x)$ and the mean residual function, $m(x)$ of TPSD (2.1) are thus obtained as

$$h(x) = \frac{\theta^3(\alpha + x + x^2)}{\theta^2(\alpha + x + x^2) + 2\theta x + \theta + 2}$$

And

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

$$= \frac{\theta^2(\alpha + x + x^2) + 2\theta(2x + 1) + 6}{\theta\left[(\alpha\theta^2 + \theta + 2) + \theta x(\theta x + \theta + 2)\right]}$$

$$= \frac{\theta^2 x^2 + \theta(\theta + 4)x + (\alpha\theta^2 + 2\theta + 6)}{\theta\left[\theta^2 x^2 + \theta(\theta + 2)x + (\alpha\theta^2 + \theta + 2)\right]}$$

It can be easily verified that $h(0) = \dfrac{\alpha\theta^3}{\alpha\theta^2 + \theta + 2} = f(0)$ and $m(0) = \dfrac{\alpha\theta^2 + 2\theta + 6}{\theta(\alpha\theta^2 + \theta + 2)} = \mu_1'$.

It can also be easily verified that the expression of $h(x)$ and $m(x)$ of TPSD reduce to the corresponding $h(x)$ and m(x) of Sujatha distribution at $\alpha = 1$.

The behavior of $h(x)$ and $m(x)$ of TPSD (2.1) for different values of its parameters are shown in Figures 4 & 5 respectively.

It is clearly seen from the graphs of $h(x)$ and $m(x)$ that $h(x)$ is monotonically increasing function of x, θ and α where as $m(x)$ is monotonically decreasing function of x, θ and α .

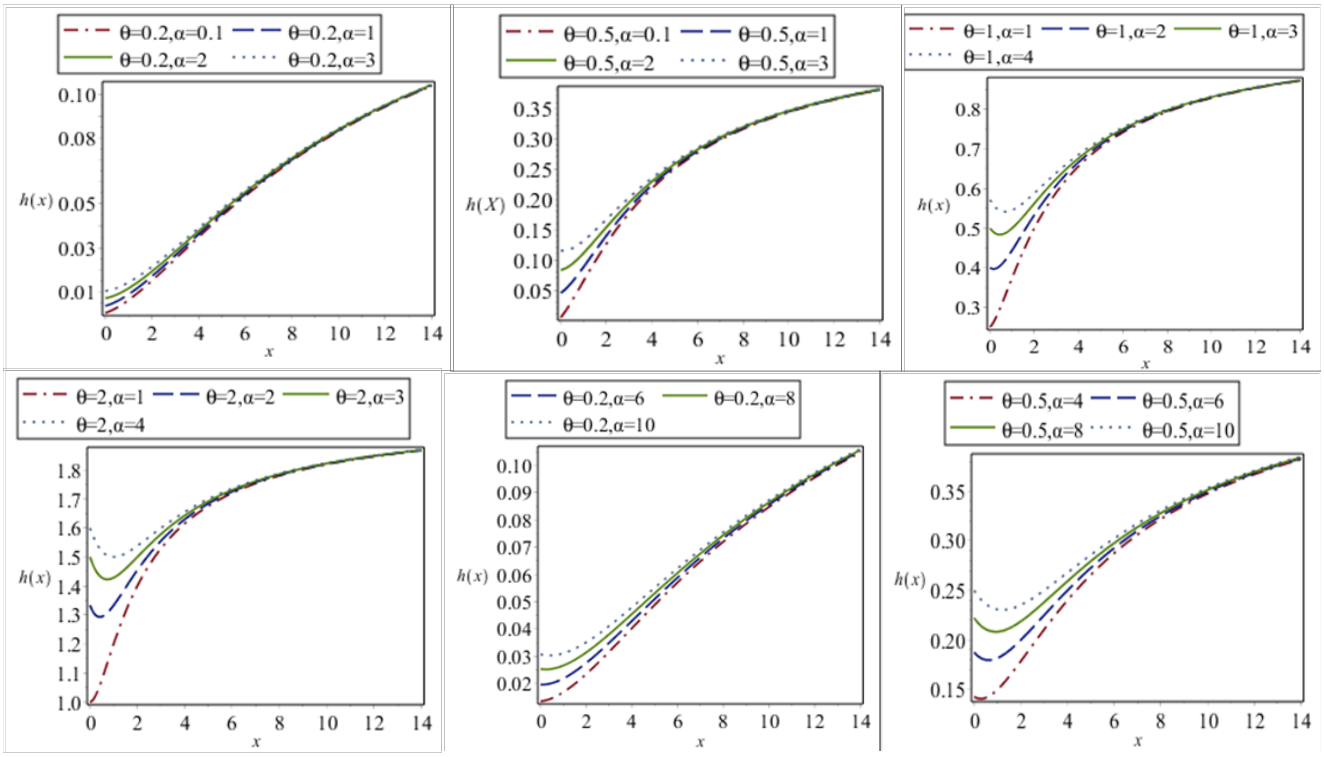

Figure 4 Behavior of $h(x)$ of TPSD for selected values of parameters θ and α.

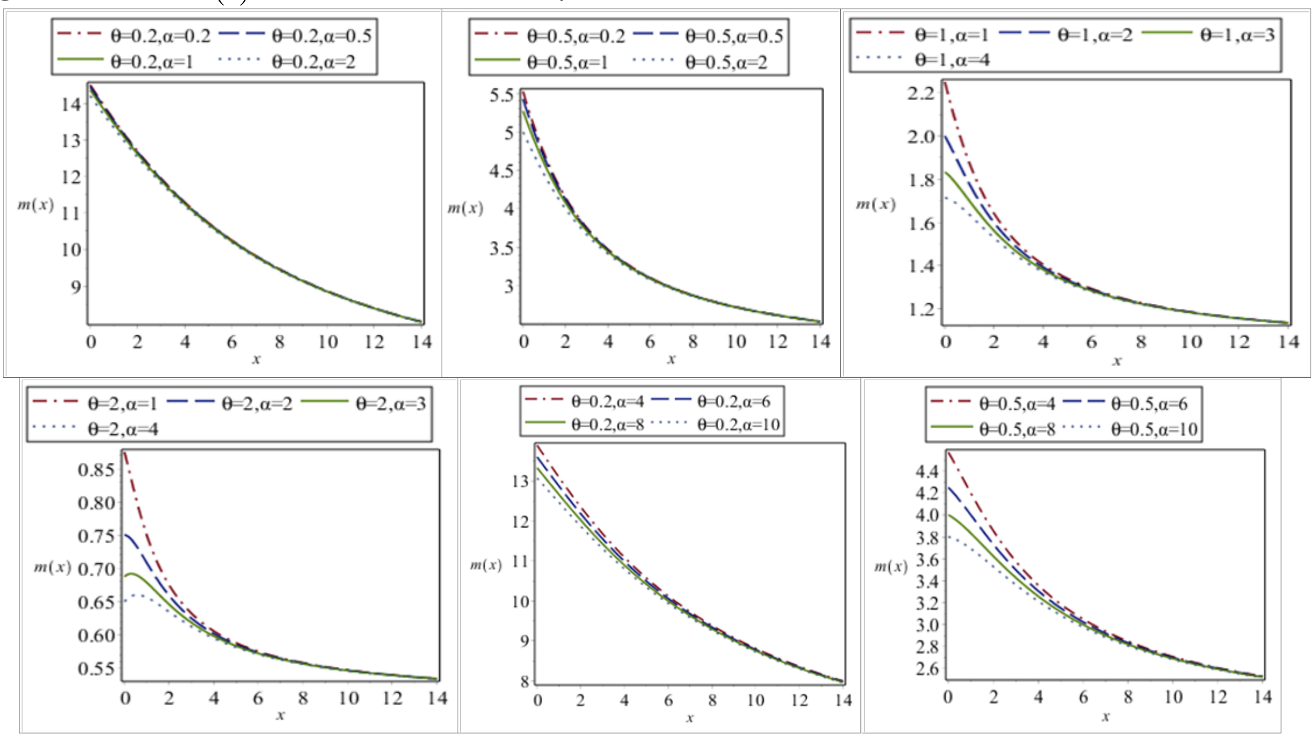

Figure 5 Behavior of $m(x)$ of TPSD for selected values of parameters θ and α.

Stochastic ordering

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

i. Stochastic order $\left(X \le_{st} Y\right)$ if $F_x\left(x\right) \ge F_y\left(x\right)$ for all x
ii. Hazard rate order $\left(X \le_{hr} Y\right)$ if $h_x\left(x\right) \ge h_y\left(x\right)$ for all x
iii. Mean residual life order $\left(X \le_{mrl} Y\right)$ if $m_x\left(x\right) \le m_y\left(x\right)$ for all x
iv. Likelihood ratio order $\left(X \le_{lr} Y\right)$ if $\frac{f_x\left(x\right)}{f_y\left(x\right)}$ decreases in x

The following results due to Shaked & Shanthikumar[8] are well known for establishing stochastic ordering of distributions

$$X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{mlr} Y .$$
$$\Downarrow$$
$$x \le_{st} y$$

The TPSD (2.1) is ordered with respect to the strongest "likelihood ratio" ordering as shown in the following theorem:

Theorem: Let $X \sim \text{TPSD}\left(\theta_1,\alpha_1\right)$ and $Y \sim \text{TPSD}\left(\theta_2,\alpha_2\right)$ .If $\theta_1 > \theta_2$ and $\alpha_1 = \alpha_2$ (or $\theta_1 = \theta_2$ and $\alpha_1 \ge \alpha_2$) then $X \le_{lr} Y$ and hence $X \le_{hr} Y, X \le_{mrl} Y$ and $X \le_{st} Y$.

Proof: We have

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

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

This gives

$$\frac{d}{dx}\ln\frac{f_X\left(x;\theta_1,\alpha_1\right)}{f_Y\left(x;\theta_2,\alpha_2\right)} = \frac{\left(\alpha_2-\alpha_1\right)+2\left(\alpha_2-\alpha_1\right)x}{\left(\alpha_1+x+x^2\right)\left(\alpha_2+x+x^2\right)}-\left(\theta_1-\theta_2\right).$$

Thus, for $\left(\theta_1 > \theta_2 \text{ and } \alpha_1 = \alpha_2\right)$ or $\left(\alpha_1 \ge \alpha_2 \text{ and } \theta_1 = \theta_2\right)$,

$$\frac{d}{dx}\ln\frac{f_X\left(x;\theta_1,\alpha_1\right)}{f_Y\left(x;\theta_2,\alpha_2\right)} < 0 .$$

This means that

$X \le_{lr} Y$ and hence $X \le_{hr} Y, X \le_{mrl} Y$ and $X \le_{st} Y$. This shows flexibility of TPSD over Sujatha distribution.

Mean deviations

The amount of scatter in a population is evidently measured to some extent by the totality of deviations from the mean and the median. These are known as the mean deviation about the mean and the mean deviation about the median and are defined as

$$\delta_1\left(x\right) = \int_0^\infty \left|x-\mu\right| f\left(x\right)dx \quad \text{and} \quad \delta_2\left(x\right) = \int_0^\infty \left|x-M\right| f\left(x\right)dx,$$

respectively,

where $\mu = E\left(X\right)$ and $M = Median\left(X\right)$. The measures $\delta_1\left(x\right)$ and $\delta_2\left(x\right)$ can be calculated using the following relationships

$$\begin{aligned}\delta_1\left(x\right) &= \int_0^\mu\left(\mu-x\right)f\left(x\right)dx+\int_\mu^\infty\left(x-\mu\right)f\left(x\right)dx \\ &= \mu F\left(\mu\right)-\int_0^\mu xf\left(x\right)dx-\mu\left[1-F\left(\mu\right)\right]+\int_\mu^\infty xf\left(x\right)dx \\ &= 2\mu F\left(\mu\right)-2\mu+2\int_\mu^\infty xf\left(x\right)dx \\ &= 2\mu F\left(\mu\right)-2\int_0^\mu xf\left(x\right)dx\end{aligned} \tag{4.3.1}$$

and

$$\begin{aligned}\delta_2\left(x\right) &= \int_0^M\left(M-x\right)f\left(x\right)dx+\int_M^\infty\left(x-M\right)f\left(x\right)dx \\ &= MF\left(M\right)-\int_0^M xf\left(x\right)dx-M\left(1-F\left(M\right)\right)+\int_M^\infty xf\left(x\right)dx \\ &= \mu+2\int_M^\infty xf\left(x\right)dx \\ &= \mu-2\int_0^M xf\left(x\right)dx\end{aligned} \tag{4.3.2}$$

Using the pdf (2.1) and expression for the mean of TPSD, we get

$$\int_0^\mu xf_4\left(x;\theta,\alpha\right)dx = \mu-\frac{\left[\theta^3\left(\mu^3+\mu^2+\alpha\mu\right)+\theta^2\left(3\mu^2+2\mu+\alpha\right)+2\theta\left(3\mu+1\right)+6\right]e^{-\theta\mu}}{\theta\left(\alpha\theta^2+\theta+2\right)} \tag{4.3.3}$$

$$\int_0^M xf_4\left(x;\theta,\alpha\right)dx = \mu-\frac{\left[\theta^3\left(M^3+M^2+\alpha M\right)+\theta^2\left(3M^2+2M+\alpha\right)+2\theta\left(3M+1\right)+6\right]e^{-\theta M}}{\theta\left(\alpha\theta^2+\theta+2\right)} \tag{4.3.4}$$

Using expressions from (4.3.1), (4.3.2), (4.3.3) and (4.3.4) and after some tedious algebraic simplifications, the mean deviation about the mean, $\delta_1\left(x\right)$ and the mean deviation about the median, $\delta_2\left(x\right)$ of TPSD are obtained as

$$\delta_1\left(x\right) = \frac{2\left[\theta^2\left(\mu^2+\mu+\alpha\right)+2\theta\left(2\mu+1\right)+6\right]e^{-\theta\mu}}{\theta\left(\alpha\theta^2+\theta+2\right)} \tag{4.3.5}$$

and

$$\delta_2\left(x\right) = \frac{2\left[\theta^3\left(M^3+M^2+\alpha M\right)+\theta^2\left(3M^2+2M+\alpha\right)+2\theta\left(3M+1\right)+6\right]e^{-\theta M}}{\theta\left(\alpha\theta^2+\theta+2\right)}-\mu \tag{4.3.6}$$

Bonferroni and Lorenz curves and indices

The Bonferroni and Lorenz curves and Bonferroni[9] and Gini indices have applications not only in economics to study income and poverty, but also in other fields like reliability, demography and medical science. The Bonferroni and Lorenz curves are defined as

$$B\left(p\right) = \frac{1}{p\mu}\int_0^q xf\left(x\right)dx = \frac{1}{p\mu}\left[\int_0^\infty xf\left(x\right)dx-\int_q^\infty xf\left(x\right)dx\right] = \frac{1}{p\mu}\left[\mu-\int_q^\infty xf\left(x\right)dx\right] \tag{4.4.1}$$

and

$$L(p)=\frac{1}{\mu}\int_0^q xf(x)\,dx=\frac{1}{\mu}\left[\int_0^\infty xf(x)\,dx-\int_q^\infty xf(x)\,dx\right]=\frac{1}{\mu}\left[\mu-\int_q^\infty xf(x)\,dx\right] \quad (4.4.2)$$

respectively or equivalently.

$$B(p)=\frac{1}{p\mu}\int_0^p F^{-1}(x)\,dx \quad (4.4.3)$$

and $$L(p)=\frac{1}{\mu}\int_0^p F^{-1}(x)\,dx \quad (4.4.4)$$

respectively, where $\mu=E(x)$ and $q=F^{-1}(p)$.

The Bonferroni and Gini indices are thus defined as

$$B=1-\int_0^1 B(p)\,dp \quad (4.4.5)$$

and $$G=1-2\int_0^1 L(p)\,dp \quad (4.4.6)$$

respectively.

Using pdf of TPSD (2.1), we get

$$\int_q^\infty xf(x)\,dx=\frac{\{\theta^3\left(q^3+q^2+\alpha q\right)+\theta^2\left(3q^2+2q+\alpha\right)+2\theta(3q+1)+6\}e^{-\theta q}}{\theta\left(\alpha\theta^2+2\theta+6\right)} \quad (4.4.7)$$

Now using equation (4.4.7), (4.4.1) and (4.4.2), we get

$$B(p)=\frac{1}{p}\left[1-\frac{\{\theta^3\left(q^3+q^2+\alpha q\right)+\theta^2\left(3q^2+2q+\alpha\right)+2\theta(3q+1)+6\}e^{-\theta q}}{\alpha\theta^2+2\theta+6}\right] \quad (4.4.8)$$

and

$$L(p)=1-\frac{\{\theta^3\left(q^3+q^2+\alpha q\right)+\theta^2\left(3q^2+2q+\alpha\right)+2\theta(3q+1)+6\}e^{-\theta q}}{\alpha\theta^2+2\theta+6} \quad (4.4.9)$$

Now using the equations (4.4.8) and (4.4.9) in (4.4.5) and (4.4.6), the Bonferroni and Gini indices of TPSD (2.1) are obtained as

$$B=1-\frac{\{\theta^3\left(q^3+q^2+\alpha q\right)+\theta^2\left(3q^2+2q+\alpha\right)+2\theta(3q+1)+6\}e^{-\theta q}}{\alpha\theta^2+2\theta+6} \quad (4.4.10)$$

$$G=-1+\frac{2\{\theta^3\left(q^3+q^2+\alpha q\right)+\theta^2\left(3q^2+2q+\alpha\right)+2\theta(3q+1)+6\}e^{-\theta q}}{\alpha\theta^2+2\theta+6} \quad (4.4.11)$$

Stress–strength reliability

The stress–strength reliability of a component illustrates the life of the component which has random strength that is subjected to random stress. When the stress of the component Y applied to it exceeds the strength of the component X, the component fails instantly and the component will function satisfactorily till $X>Y$. Therefore, $R=P(Y<X)$ is a measure of the component reliability and is known as stress–strength reliability in statistical literature. It has extensive application in almost all areas of knowledge especially in engineering such as structure, deterioration of rocket motor, static fatigue of ceramic component, aging of concrete pressure vessels etc.

Let X and Y be independent strength and stress random variables having TPSD (2.1) with parameter (θ_1,α_1) and (θ_2,α_2) respectively. Then the stress–strength reliability R of TPSD can be obtained as

$$R=P(Y<X)=\int_0^\infty P(Y<X\mid X=x)f_x(x)\,dx$$

$$=\int_0^\infty f(x;\theta_1,\alpha_1)F(x;\theta_2,\alpha_2)\,dx$$

$$=1-\frac{\theta_1^3\left[\begin{array}{l}(\alpha_1\alpha_2)\theta_2^6+\left(2\alpha_1+\alpha_2+4\alpha_1\alpha_2\theta_1\right)\theta_2^5+\left(7\alpha_1\theta_1+3\alpha_2\theta_1+6\alpha_1+2\alpha_2+6\alpha_1\alpha_2\theta_1^2+3\right)\theta_2^4\\+\left(9\alpha_1\theta_1^2+3\alpha_2\theta_1^2+18\alpha_1\theta_1+4\alpha_2\theta_1+7\theta_1+4\alpha_1\alpha_2\theta_1^3+20\right)\theta_2^3\\+\left(5\alpha_1\theta_1^3+\alpha_2\theta_1^3+20\alpha_1\theta_1^2+2\alpha_2\theta_1^2+5\theta_1^2+30\theta_1+\alpha_1\alpha_2\theta_1^4+40\right)\theta_2^2\\+\left(\alpha_1\theta_1^3+10\alpha_1\theta_1^2+\theta_1^2+12\theta_1+20\right)\theta_1\theta_2+2\left(\alpha_1\theta_1^2+\theta_1+2\right)\theta_1^2\end{array}\right]}{\left(\alpha_1\theta_1^2+\theta_1+2\right)\left(\alpha_2\theta_2^2+\theta_2+2\right)\left(\theta_1+\theta_2\right)^5}$$

It can be verified that the stress–strength reliability of Sujatha distribution is a particular case of stress–strength reliability of TPSD at $\alpha_1=\alpha_2=1$.

Estimation of parameters

In this section, the estimation of parameters of TPSD using the method of moments and the method of maximum likelihood have been discussed.

Method of moment estimates (MOME)

Since TPSD (2.1) has two parameters to be estimated, the first two moments about the origin are required to estimate its parameters using method of moments. Equating the population mean to the sample mean, we have

$$\bar{x}=\frac{\alpha\theta^2+2\theta+6}{\theta\left(\alpha\theta^2+\theta+2\right)}=\frac{\alpha\theta^2+\theta+2}{\theta\left(\alpha\theta^2+\theta+2\right)}+\frac{\theta+4}{\theta\left(\alpha\theta^2+\theta+2\right)}$$

$$\bar{x}=\frac{1}{\theta}+\frac{\theta+4}{\theta\left(\alpha\theta^2+\theta+2\right)}$$

$$\left(\alpha\theta^2+\theta+2\right)=\frac{\theta+4}{\theta\bar{x}-1} \quad (5.1.1)$$

Again equating the second population moment with the corresponding sample moment, we have

$$m_2' = \frac{2\left(\alpha\theta^2+3\theta+12\right)}{\theta^2\left(\alpha\theta^2+\theta+2\right)} = \frac{2\left(\alpha\theta^2+\theta+2\right)}{\theta^2\left(\alpha\theta^2+\theta+2\right)} + \frac{4\left(\theta+5\right)}{\theta^2\left(\alpha\theta^2+\theta+2\right)}$$

$$m_2' = \frac{2\left(\alpha\theta^2+3\theta+12\right)}{\theta^2\left(\alpha\theta^2+\theta+2\right)} = \frac{2}{\theta^2} + \frac{4\left(\theta+5\right)}{\theta^2\left(\alpha\theta^2+\theta+2\right)}$$

$$\alpha\theta^2+\theta+2 = \frac{4\left(\theta+5\right)}{m_2'\theta^2-2} \quad (5.1.2)$$

Equations (5.1.1) and (5.1.2) give the following cubic equation in θ

$$m_2'\theta^3 + 4\left(m_2' - \bar{X}\right)\theta^2 - 2\left(10\bar{X}-1\right)\theta + 12 = 0 \quad (5.1.3)$$

Solving equation (5.1.3) using any iterative method such as Newton–Raphson method, Regula–Falsi method or Bisection method, method of moment estimation (MOME) $\tilde{\theta}$ of θ can be obtained and substituting the value of $\tilde{\theta}$ in equation (5.1.1), MOME $\tilde{\alpha}$ of α can be obtained as

$$\tilde{\alpha} = \frac{-\bar{x}\tilde{\theta}^2 - 2(\bar{x}-1)\tilde{\theta} + 6}{\tilde{\theta}^2(\tilde{\theta}\bar{x}-1)} \quad (5.1.4)$$

Maximum likelihood estimates (MLE)

Let $\left(x_1, x_2, x_3, ...x_n\right)$ be random sample from TPSD (2.1). The likelihood function L is given by

$$L = \left(\frac{\theta^3}{\alpha\theta^2+\theta+2}\right)^n \prod_{i=1}^{n}\left(\alpha + x_i + x_i^2\right)e^{-n\theta\bar{x}},$$

where $\bar{x}$ is the sample mean.
The natural log likelihood function is thus obtained as

$$\ln L = n\left[3\ln\theta - \ln\left(\alpha\theta^2+\theta+2\right)\right] + \sum_{i=1}^{n}\ln\left(\alpha + x_i + x_i^2\right) - n\theta\bar{x}$$

The maximum likelihood estimate (MLE's) $(\hat{\theta}, \hat{\alpha})$ of (θ, α) are then the solutions of the following non–linear equations

$$\frac{\partial \ln L}{\partial\theta} = \frac{3n}{\theta} - \frac{n\left(2\alpha\theta+1\right)}{\alpha\theta^2+\theta+2} - n\bar{x} = 0 \quad (5.2.1)$$

$$\frac{\partial \ln L}{\partial\alpha} = -\frac{n\theta^2}{\alpha\theta^2+\theta+2} + \sum_{i=1}^{n}\frac{1}{\alpha + x_i + x_i^2} = 0 \quad (5.2.2)$$

These two natural log likelihood equations do not seem to be solved directly, because they cannot be expressed in closed forms. The (MLE's) $(\hat{\theta}, \hat{\alpha})$ of (θ, α) can be computed directly by solving the natural log likelihood equations using Newton–Raphson iteration method using R–software till sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained. The initial values of parameters θ and α are the MOME $(\tilde{\theta}, \tilde{\alpha})$ of the parameters (θ, α).

A numerical example

In this section an application of TPSD using maximum likelihood estimates has been discussed with a real lifetime data set. The data set regarding vinyl chloride obtained from clean up gradient monitoring wells in mg/l, available in Bhaumik *et al.*[10] has been considered. The data set is

5.1	1.2	1.3	0.6	0.5	2.4	0.5	1.1	8	0.8	0.4	0.6	0.9
0.4	2	0.5	5.3	3.2	2.7	2.9	2.5	2.3	1	0.2	0.1	0.1
1.8	0.9	2.4	6.8	1.2	0.4	0.2						

In order to compare lifetime distributions, values of $-2\ln L$, AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and K–S Statistic (Kolmogorov–Smirnov Statistic) for the above data set has been computed. The formulae for computing AIC, BIC, and K–S Statistics are as follows:

$AIC = -2\ln L + 2k$, $BIC = -2\ln L + k\ln n$ and $D = \operatorname{Sup}_x \left|F_n\left(x\right) - F_0\left(x\right)\right|$, where k = the number of parameters, n = the sample size, and the $F_n\left(x\right)$ = empirical distribution function. The best distribution is the distribution which corresponds to lower values of $-2\ln L$, AIC, and K–S statistic and higher p–value. The MLE $\left(\hat{\theta}, \hat{\alpha}\right)$ along with their standard errors, $-2\ln L$, AIC, BIC, K–S Statistic and p–value of the fitted distributions are presented in the Table 5.

It is obvious that TPSD gives much closer fit than Sujatha and Lindley distributions. Therefore, TPSD can be considered as an important two–parameter lifetime distribution. In order to see the closeness of the fit given by Lindley, Sujatha and TPSD, the fitted pdf plots of these distributions for the given dataset have been shown in Figure 6. It is also obvious from the fitted plots of the distribution along with the histogram of the original dataset that TPSD gives much closer fit than Lindley and Sujatha distributions.

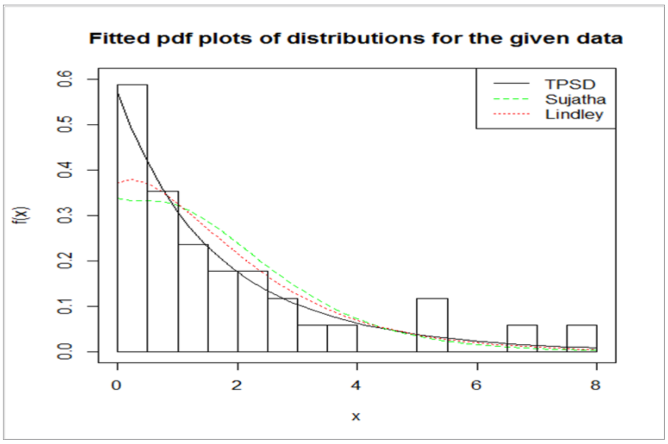

Figure 6 Fitted pdf plots of distributions for the given dataset.

Table 5 MLE's, S.E$\left(\hat{\theta},\hat{\alpha}\right)$, $-2\ln L$, AIC, BIC, and K–S Statistics of the fitted distributions of the given data set

Distribution	MLE's	S.E	–2lnL	AIC	BIC	K– S	P–Value
TPSD	$\hat{\theta}=0.6919$	0.2781	110.72	114.72	117.77	0.084	0.9693
	$\hat{\alpha}=27.017$	73.095					
Sujatha	$\hat{\theta}=1.1461$	0.116	115.54	117.54	119.07	0.164	0.3196
Lindley	$\hat{\theta}=0.8238$	0.1054	112.61	114.61	116.14	0.1326	0.5881

Conclusion

A two–parameter Sujatha distribution (TPSD) has been introduced which includes size–biased Lindley distribution and Sujatha distribution, proposed by Shanker[3] as particular cases. Moments about origin and moments about mean have been obtained and nature of coefficient of variation, coefficient of skewness, coefficient of kurtosis and index of dispersion of TPSD have been studied with varying values of the parameters. The nature of probability density function, cumulative distribution function, hazard rate function and mean residual life function have been discussed with varying values of the parameters. The stochastic ordering, mean deviations, Bonferroni and Lorenz curves, and stress–strength reliability have also been discussed. The method of moments and method of maximum likelihood have been discussed for estimating parameters. A numerical example of real lifetime data have been presented to show the application of TPSD and the goodness of fit of TPSD gives much closer fit over Sujatha and Lindley distributions.

Acknowledgements

Authors are grateful to the editor–in–chief of the journal and the anonymous reviewer for constructive comments on the paper.

Conflict of interest

Authors declare that there is no conflict of interest.

References

1. Lindley DV. Fiducial distributions and Bayes'theorem. *Journal of the Royal Statistical Society. Series B*. 1958;20(1):102–107.
2. Shanker R, Hagos F, Sujatha S. On modeling of lifetimes data using exponential and Lindley distribution. *BBIJ*. 2015;2(5):1–9.
3. Shanker R. Sujatha distribution and its applications. *Statistics in Transition New series*. 2016a;17(3):1–20.
4. Shanker R. The discrete Poisson–Sujatha distribution. *International Journal of Probability and Statistics*. 2016 b;5(1):1–9.
5. Shanker R, Hagos F. Zero–Truncated Poisson–Sujatha distribution with Applications. *Journal of Ethiopian Statistical Association*. 2015;24:55–63.
6. Shanker R, Hagos F. Size–biased Poisson–Sujatha distribution with Applications. *American Journal of Mathematics and Statistics*. 2016;6(4):145–154.
7. Ghitany ME, Atieh B, Nadarajah S. Lindley distribution and its Application. *Mathematics Computing and Simulation*. 2008;78(4):493–506.
8. Shaked M, Shanthikumar JG. *Stochastic Orders and Their Applications*. New York: Academic Press; 1994.
9. Bonferroni CE. Elementi di Statistca Generale. Firenze: Seeber;1930.
10. Bhaumik DK, Kapur K, Gibbson RD. Testing of parameters of a Gamma distribution for small samples. *Technometrics*. 2009;51(3):326–334.