

A GENERALIZATION OF SUJATHA DISTRIBUTION AND ITS APPLICATIONS WITH REAL LIFETIME DATA

**Rama Shanker, Kamlesh Kumar Shukla and
Hagos Fesshaye**

Journal of Institute of Science and Technology

Volume 22, Issue 1, July 2017

ISSN: 2469-9062 (print), 2467-9240 (e)

Editors:

Prof. Dr. Kumar Sapkota

Prof. Dr. Armila Rajbhandari

Assoc. Prof. Dr. Gopi Chandra Kaphle

JIST, 22 (1): 66-83 (2017)

Published by:

Institute of Science and Technology

Tribhuvan University

Kirtipur, Kathmandu, Nepal



A GENERALIZATION OF SUJATHA DISTRIBUTION AND ITS APPLICATIONS WITH REAL LIFETIME DATA

Rama Shanker¹, Kamlesh Kumar Shukla^{1,*} and Hagos Fesshaye²

¹Department of Statistics, Eritrea Institute of Technology, Asmara, Eritrea

²Department of Economics, College of Business and Economics, Halhale, Eritrea

* Corresponding E-mail: kkshukla22@gmail.com

ABSTRACT

A two-parameter generalization of Sujatha distribution (AGSD), which includes Lindley distribution and Sujatha distribution as particular cases, has been proposed. Its important mathematical and statistical properties including its shape for varying values of parameters, moments, 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. Maximum likelihood estimation method has been discussed for estimating its parameters. AGSD provides better fit than Sujatha, Aradhana, Lindley and exponential distributions for modeling real lifetime data.

Keywords: Lindley distribution, Hazard rate function, Mean residual life function, Bonferroni and Lorenz curves, Stress-strength reliability.

INTRODUCTION

The probability density function (PDF) and the cumulative distribution function (CDF) of Lindley (1958) distribution are given by

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

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

The density (1.1) is a two-component mixture of an exponential distribution having scale parameter θ and a gamma distribution having shape parameter 2 and scale parameter θ with their mixing proportions $\frac{\theta}{\theta + 1}$ and $\frac{1}{\theta + 1}$ respectively. Detailed study on its various mathematical properties, estimation of parameter and application showing the superiority of

Lindley distribution over exponential distribution for the waiting times before service of the bank customers has been done by Ghitany *et al.* (2008). Shanker *et al.* (2015) had critically studied the modeling of lifetime data using exponential and Lindley distributions and concluded that there are several lifetime data where these distributions are not suitable from theoretical or applied point of view. In recent years much work have been done on Lindley distribution, its mixture with other distributions, extensions, and generalizations by many researchers including Zakerzadeh and Dolati (2009), Nadarajah *et al.* (2011), Deniz and Ojeda (2011), Bakouch *et al.* (2012), Shanker and Mishra (2013 a, 2013 b), Shanker and Amanuel (2013), Shanker *et al.* (2015), Ghitany *et al.* (2013), Shanker *et al.* (2016 a, 2016 b, 2016 c), Shanker (2016 a), Shanker *et al.* (2017).

The PDF and the CDF of Aradhana distribution introduced by Shanker (2016 b) are given by

$$f_2(x; \theta) = \frac{\theta^3}{\theta^2 + 2\theta + 2} (1 + x)^2 e^{-\theta x} ; x > 0, \theta > 0 \quad (1.3)$$

$$F_2(x; \theta) = 1 - \left[1 + \frac{\theta x (\theta x + 2\theta + 2)}{\theta^2 + 2\theta + 2} \right] e^{-\theta x} ; x > 0, \theta > 0 \quad (1.4)$$

Shanker (2016 b) has shown that the Aradhana distribution is a three-component mixture of an exponential distribution having scale parameter θ , a gamma distribution having shape parameter 2 and scale parameter θ , and a gamma distribution with shape parameter 3 and scale parameter θ with their mixing proportions $\frac{\theta^2}{\theta^2 + \theta + 2}$, $\frac{\theta}{\theta^2 + \theta + 2}$ and $\frac{2}{\theta^2 + \theta + 2}$. Shanker (2016 b) has discussed its important properties

$$f_3(x; \theta) = \frac{\theta^3}{\theta^2 + \theta + 2} (1 + x + x^2) e^{-\theta x} ; x > 0, \theta > 0 \tag{1.5}$$

$$F_3(x, \theta) = 1 - \left[1 + \frac{\theta x (\theta x + \theta + 2)}{\theta^2 + \theta + 2} \right] e^{-\theta x} ; x > 0, \theta > 0 \tag{1.6}$$

Shanker (2016 c) has shown that the Sujatha distribution is a three-component mixture of an exponential distribution having scale parameter θ , a gamma distribution having shape parameter 2 and scale parameter θ , and a gamma distribution with shape parameter 3 and scale parameter θ with their mixing proportions $\frac{\theta^2}{\theta^2 + 2\theta + 2}$, $\frac{2\theta}{\theta^2 + 2\theta + 2}$ and $\frac{2}{\theta^2 + 2\theta + 2}$, respectively. Shanker (2016 c) has discussed its important properties including its shape for varying values of parameters, moments, 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. Shanker (2016 c) has discussed the maximum likelihood estimation of parameter and showed the applications of Sujatha distribution to model lifetime data from biomedical science and engineer. Shanker (2016 d) has also obtained a Poisson mixture of Sujatha distribution namely Poisson-Sujatha distribution (PSD) and studied its properties, estimation of parameter and applications for count data. Shanker and Hagos (2016, 2015) have obtained and discussed the size-

including its shape for varying values of parameters, moments, 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. Shanker (2016 b) has discussed the applications of Aradhana distribution for modeling lifetime data using maximum likelihood estimation. The PDF and the CDF of Sujatha distribution introduced by Shanker (2016 c) are given by

biased and zero-truncated Poisson-Sujatha distribution and their various statistical and mathematical properties, estimation of parameter and applications to model count data which structurally exclude zero-counts. Recently, Shanker (2016 e) has introduced a quasi Sujatha distribution (QSD), for modeling lifetime data from biomedical science and engineering. In this paper, a generalization of Sujatha distribution (AGSD), of which one parameter Lindley (1958) distribution and Sujatha distribution introduced by Shanker (2016 c) are particular cases, has been proposed. It's important properties including hazard rate function, mean residual life function, stochastic ordering, mean deviations, Bonferroni and Lorenz curves, stress-strength reliability have also been discussed. The estimation of the parameters has been discussed using maximum likelihood estimation. Some numerical examples have been given to test the goodness of fit of AGSD and the fit has been compared with one parameter Sujatha, Aradhana, Lindley and exponential distributions.

A GENERALIZATION OF SUJATHA DISTRIBUTION

The probability density function of a generalization of Sujatha distribution (AGSD) can be introduced as:

$$f_4(x; \theta, \alpha) = \frac{\theta^3}{\theta^2 + \theta + 2\alpha} (1 + x + \alpha x^2) e^{-\theta x} ; x > 0, \theta > 0, \alpha > 0 \tag{2.1}$$

where θ is a scale parameter and α is a shape parameter. It can be easily verified that (2.1) reduces to the one parameter Lindley (1958) distribution (1.1) and Sujatha distribution (1.3), introduced by Shanker (2016 c) for $\alpha = 0$ and

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

Where,

$$p_1 = \frac{\theta^2}{\theta^2 + \theta + 2\alpha}$$

$$p_2 = \frac{\theta}{\theta^2 + \theta + 2\alpha}$$

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

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

$$g_3(x; \theta) = \frac{\theta^3 x^2 e^{-\theta x}}{2}; x > 0, \theta > 0$$

Since AGSD includes Lindley and Sujatha distributions as particular cases, it is expected to give better fit than both Lindley and Sujatha distributions for modeling lifetime data from biological sciences and engineering.

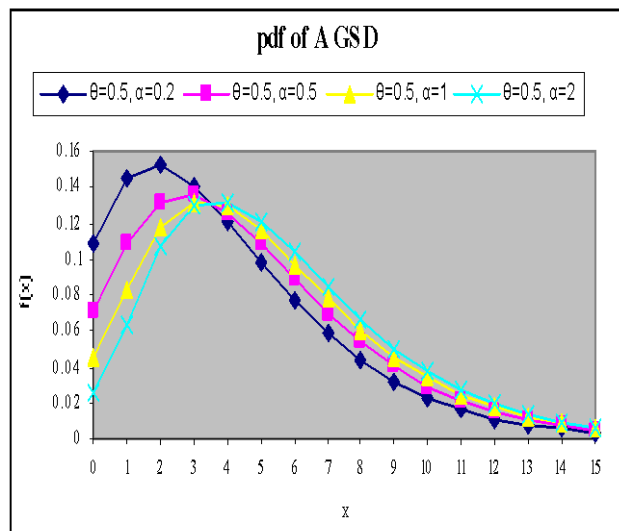
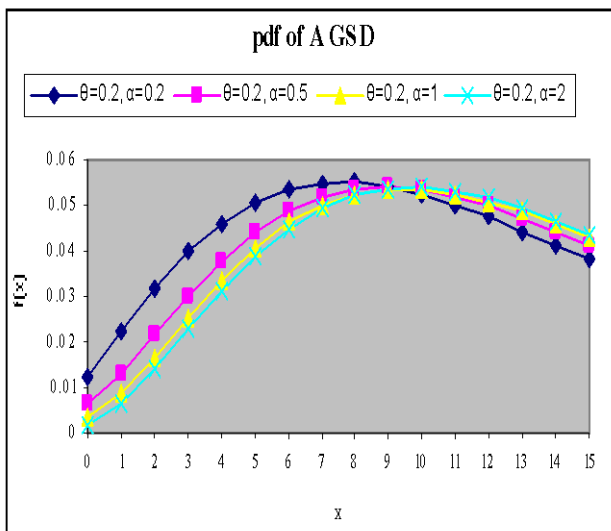
$\alpha = 1$. It can be easily shown that AGSD is a three-component mixture of exponential(θ), gamma ($2, \theta$) and gamma ($3, \theta$) distributions. We have

Further, for $\alpha \rightarrow \infty$, the PDF of AGSD (2.1) reduces to the PDF of gamma distribution with shape parameter 3 and scale parameter θ .

The corresponding cumulative distribution function of (2.1) can be obtained as:

$$F_4(x; \theta, \alpha) = 1 - \left[1 + \frac{\theta x (\alpha \theta x + \theta + 2\alpha)}{\theta^2 + \theta + 2\alpha} \right] e^{-\theta x}; x > 0, \theta > 0, \alpha > 0 \quad (2.2)$$

Graphs of the PDF and the CDF of AGSD are shown in figures 1 (a) and 1 (b) for varying values of parameters θ and α .



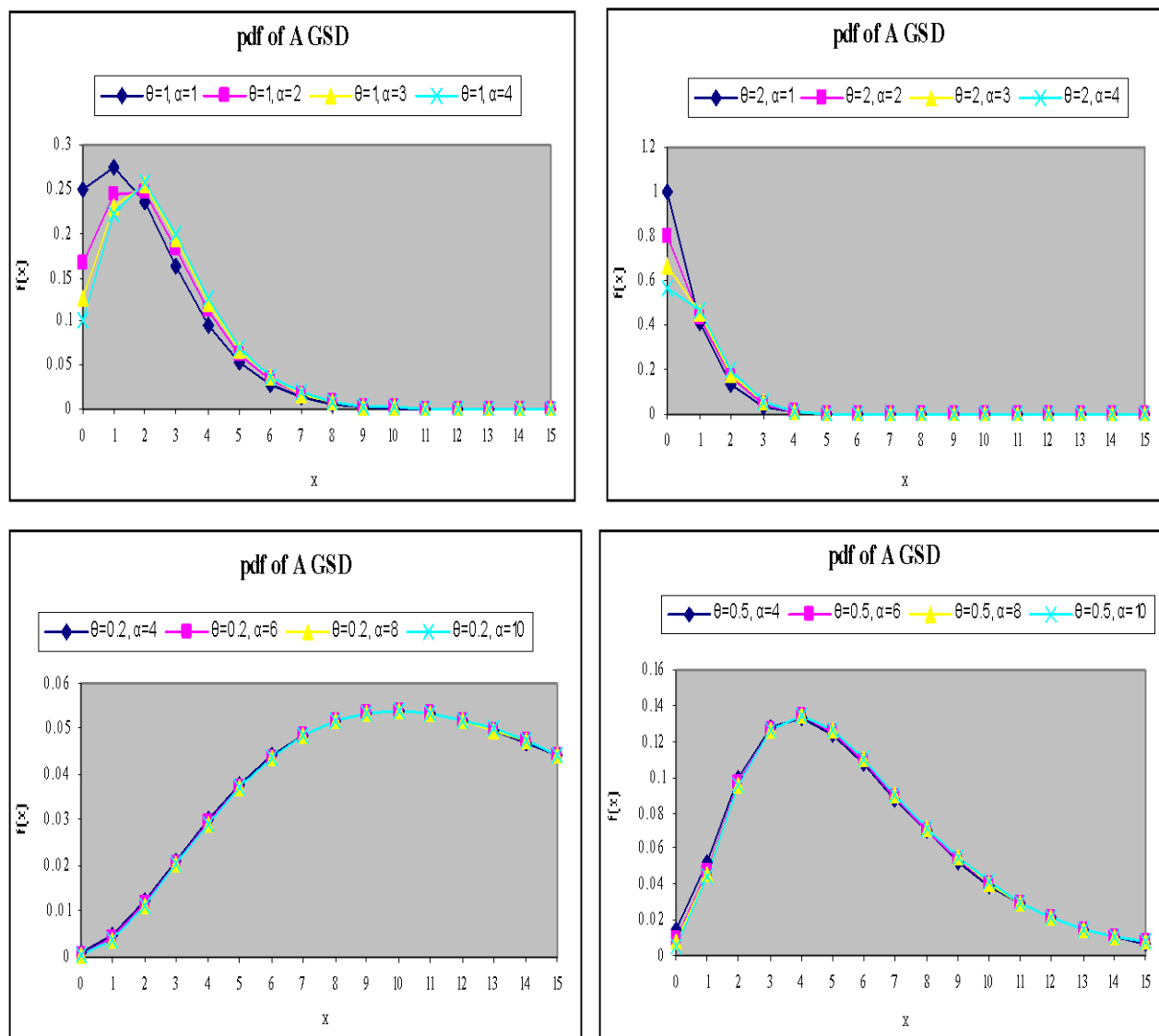
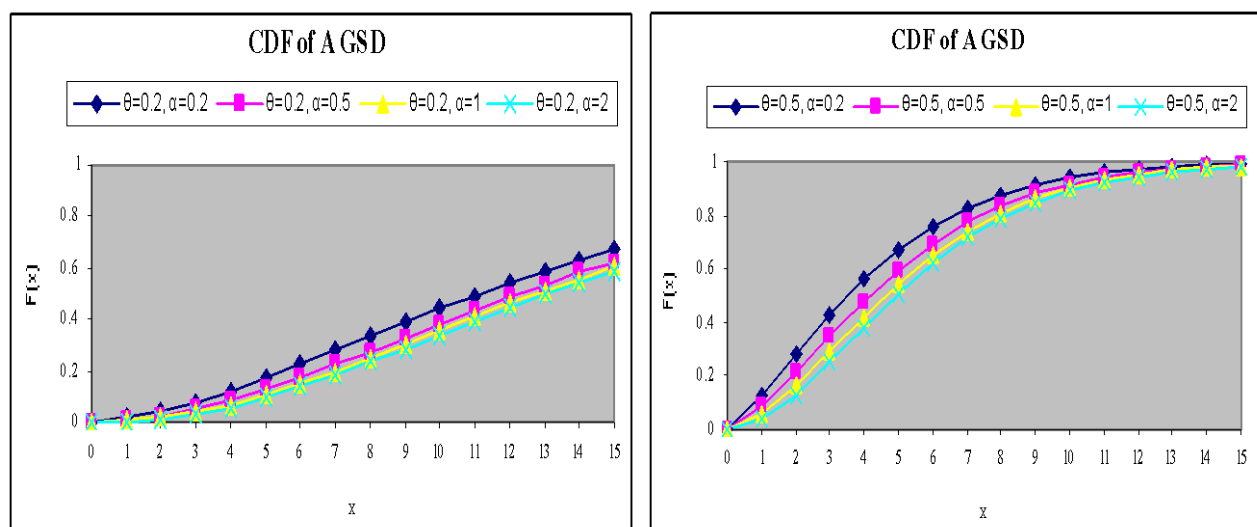


Fig. 1 (a) Graphs of the PDF of AGSD for varying values of parameters θ and α .



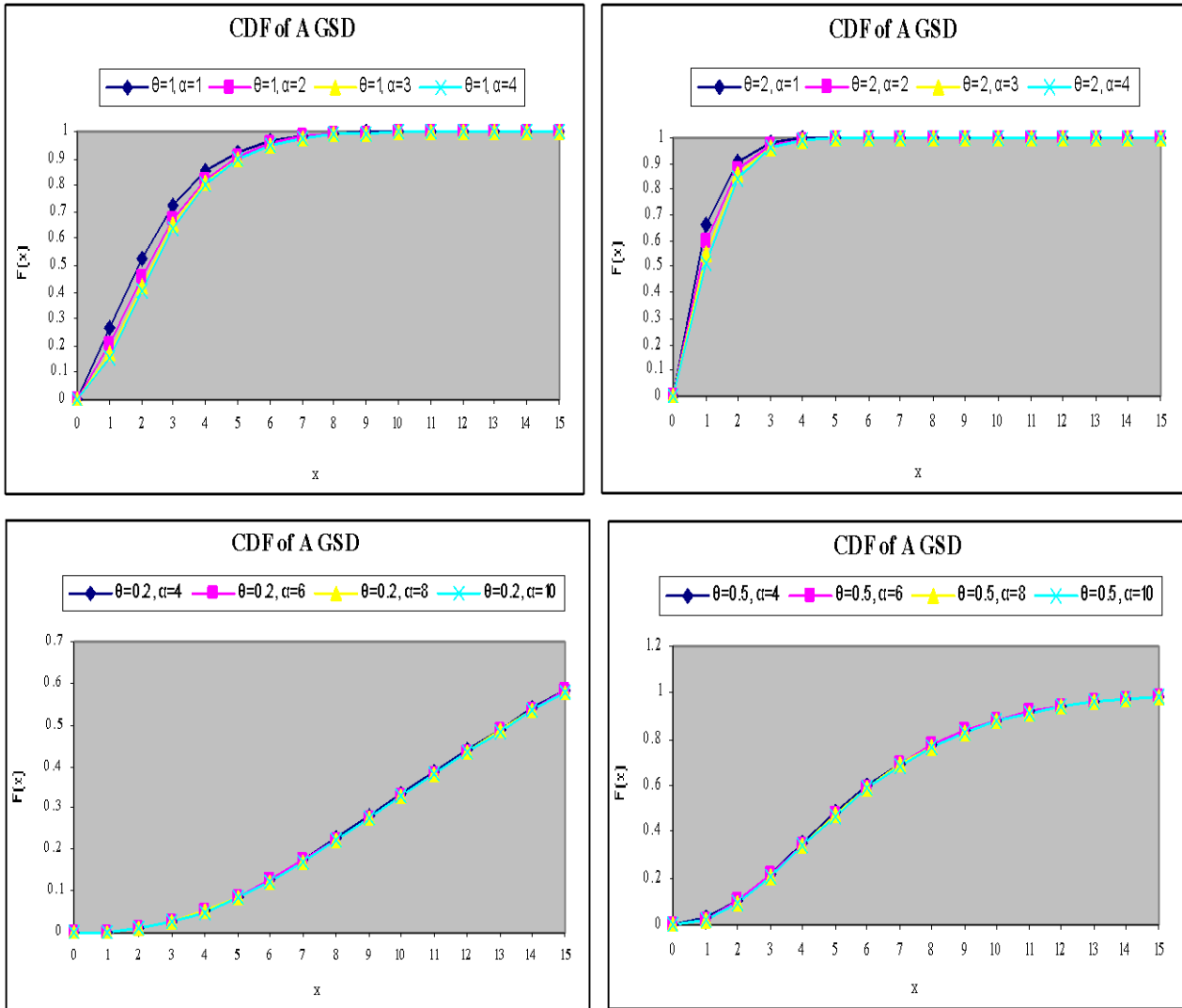


Fig. 1 (b) Graphs of the CDF of AGSD for varying values of parameters θ and α .

MOMENTS

The r^{th} moment about origin, μ_r' of AGSD (2.1) can be obtained as:

$$\mu_r' = \frac{r! [\theta^2 + (r+1)\theta + (r+1)(r+2)\alpha]}{\theta^r (\theta^2 + \theta + 2\alpha)} ; r = 1, 2, 3, 4, \dots \tag{3.1}$$

The first four moments about origin of AGSD (2.1) are thus obtained as:

$$\begin{aligned} \mu_1' &= \frac{\theta^2 + 2\theta + 6\alpha}{\theta(\theta^2 + \theta + 2\alpha)}, & \mu_2' &= \frac{2(\theta^2 + 3\theta + 12\alpha)}{\theta^2(\theta^2 + \theta + 2\alpha)}, \\ \mu_3' &= \frac{6(\theta^2 + 4\theta + 20\alpha)}{\theta^3(\theta^2 + \theta + 2\alpha)}, & \mu_4' &= \frac{24(\theta^2 + 5\theta + 30\alpha)}{\theta^4(\theta^2 + \theta + 2\alpha)} \end{aligned}$$

Using the relationship between moments about mean and the moments about origin, the moments about mean of AGSD (2.1) are obtained as

$$\mu_2 = \frac{\theta^4 + 4\theta^3 + 2(8\alpha + 1)\theta^2 + 12\theta\alpha + 12\alpha^2}{\theta^2(\theta^2 + \theta + 2\alpha)^2}$$

$$\mu_3 = \frac{2 \left[\theta^6 + 6\theta^5 + 6(5\alpha + 1)\theta^4 + 2(21\alpha + 1)\theta^3 + 18(2\alpha + 1)\theta^2\alpha + 36\theta\alpha^2 + 24\alpha^3 \right]}{\theta^3(\theta^2 + \theta + 2\alpha)^3}$$

$$\mu_4 = \frac{3 \left[3\theta^8 + 24\theta^7 + 4(32\alpha + 11)\theta^6 + 8(43\alpha + 4)\theta^5 + 8(51\alpha^2 + 40\alpha + 1)\theta^4 + 96(8\alpha + 1)\theta^3\alpha + 48(9\alpha + 7)\theta^2\alpha^2 + 480\theta\alpha^3 + 240\alpha^4 \right]}{\theta^4(\theta^2 + \theta + 2\alpha)^4}$$

The coefficient of variation (C.V), coefficient of skewness ($\sqrt{\beta_1}$), coefficient of kurtosis (β_2) and index of dispersion (γ) of AGSD (2.1) are thus obtained as

$$C.V = \frac{\sigma}{\mu_1'} = \frac{\sqrt{\theta^4 + 4\theta^3 + 2(8\alpha + 1)\theta^2 + 12\theta\alpha + 12\alpha^2}}{\theta^2 + 2\theta + 6\alpha}$$

$$\sqrt{\beta_1} = \frac{\mu_3}{\mu_2^{3/2}} = \frac{2 \left[\theta^6 + 6\theta^5 + 6(5\alpha + 1)\theta^4 + 2(21\alpha + 1)\theta^3 + 18(2\alpha + 1)\theta^2\alpha + 36\theta\alpha^2 + 24\alpha^3 \right]}{\left[\theta^4 + 4\theta^3 + 2(8\alpha + 1)\theta^2 + 12\theta\alpha + 12\alpha^2 \right]^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{3 \left[3\theta^8 + 24\theta^7 + 4(32\alpha + 11)\theta^6 + 8(43\alpha + 4)\theta^5 + 8(51\alpha^2 + 40\alpha + 1)\theta^4 + 96(8\alpha + 1)\theta^3\alpha + 48(9\alpha + 7)\theta^2\alpha^2 + 480\theta\alpha^3 + 240\alpha^4 \right]}{\left[\theta^4 + 4\theta^3 + 2(8\alpha + 1)\theta^2 + 12\theta\alpha + 12\alpha^2 \right]^2}$$

$$\gamma = \frac{\sigma^2}{\mu_1'} = \frac{\theta^4 + 4\theta^3 + 2(8\alpha + 1)\theta^2 + 12\theta\alpha + 12\alpha^2}{\theta(\theta^2 + \theta + 2\alpha)(\theta^2 + 2\theta + 6\alpha)}$$

To study the nature 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 1,2,3 and 4.

Table 1. C.V of AGSD for varying values of parameters θ and α .

$\alpha \backslash \theta$	0.2	0.5	1	2	3	4	5
0.2	0.668399	0.768226	0.861102	0.935541	0.9632	0.976287	0.983472
0.5	0.62575	0.708329	0.816497	0.922627	0.96225	0.979208	0.987456
1	0.604466	0.662392	0.761739	0.892143	0.95119	0.977525	0.989835
2	0.591771	0.627199	0.702377	0.83666	0.918489	0.96225	0.984948
3	0.587172	0.612596	0.671749	0.795698	0.886072	0.941606	0.973665
4	0.584797	0.604609	0.653155	0.765466	0.85773	0.920447	0.959837
5	0.583347	0.599573	0.640678	0.74246	0.833556	0.900389	0.945247

For a given value of α , C.V increases as the value of θ increases. Again for a given value of $\theta (\leq 3)$, C.V decreases as the value of α increases. But for $4 \leq \theta \leq 5$, C.V increases as the value of α increases between $0.2 \leq \alpha \leq 0.5$ and decreases for $\alpha \geq 1$.

Table 2. $\sqrt{\beta_1}$ of AGSD for varying values of parameters θ and α .

$\alpha \backslash \theta$	0.2	0.5	1	2	3	4	5
0.2	1.193449	1.327366	1.525066	1.739109	1.837044	1.888906	1.91951
0.5	1.150887	1.201582	1.377838	1.662624	1.809416	1.88468	1.92581
1	1.145006	1.145839	1.247611	1.535588	1.733747	1.848046	1.912879
2	1.146979	1.129318	1.154381	1.365976	1.586598	1.745907	1.850452
3	1.148785	1.130201	1.126373	1.27031	1.475249	1.649093	1.776904
4	1.149954	1.132836	1.116957	1.21277	1.3934	1.567307	1.706821
5	1.150749	1.13533	1.11421	1.176244	1.332296	1.499635	1.643745

Since $\sqrt{\beta_1} > 0$, AGSD is always positively skewed.

Table 3. β_2 of AGSD for varying values of parameters θ and α .

$\alpha \backslash \theta$	0.2	0.5	1	2	3	4	5
0.2	4.998199	5.432504	6.255787	7.329721	7.897921	8.223101	8.424354
0.5	4.882734	4.918002	5.53125	6.845226	7.6754	8.154083	8.435548
1	4.898384	4.741512	4.974649	6.143984	7.169484	7.859059	8.293729
2	4.933706	4.741178	4.643535	5.333878	6.323212	7.176	7.813096
3	4.951698	4.781497	4.579203	4.933269	5.759668	6.607315	7.323125
4	4.962123	4.815044	4.580071	4.715733	5.380366	6.167697	6.896751
5	4.968875	4.840695	4.599658	4.59102	5.115984	5.827849	6.538529

Since $\beta_2 > 3$, AGSD is always leptokurtic, which means that AGSD is more peaked than the normal curve.

Table 4. γ of AGSD for varying values of parameters θ and α

$\theta \backslash \alpha$	0.2	0.5	1	2	3	4	5
0.2	5.724085	2.514641	1.297619	0.629076	0.404022	0.294351	0.23035
0.5	5.431358	2.436975	1.333333	0.668831	0.42735	0.308201	0.239049
1	5.252329	2.31348	1.305556	0.696429	0.452381	0.325758	0.251067
2	5.137263	2.194638	1.233333	0.7	0.474537	0.347222	0.26821
3	5.094207	2.140452	1.184524	0.685897	0.479798	0.358059	0.27914
4	5.071703	2.10976	1.151852	0.669643	0.478205	0.363095	0.286084
5	5.057875	2.090047	1.128788	0.654605	0.473737	0.364815	0.290385

As long as $0 < \theta \leq 1$ and $0 \leq \alpha \leq 5$, the nature of AGSD is over-dispersed ($\sigma^2 > \mu_1'$) and for $\theta \geq 2$ and $\alpha > 0$, the nature of AGSD is under-dispersed ($\sigma^2 < \mu_1'$)

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 the failure rate function), $h(x)$ and the mean residual life function, $m(x)$ of X are respectively defined as

$$h(x) = \lim_{\Delta x \rightarrow 0} \frac{P(X < x + \Delta x | X > x)}{\Delta x} = \frac{f(x)}{1 - F(x)} \tag{4.1}$$

and

$$m(x) = E[X - x | X > x] = \frac{1}{1 - F(x)} \int_x^\infty [1 - F(t)] dt \tag{4.2}$$

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

$$h(x) = \frac{\theta^3 (1 + x + \alpha x^2)}{\theta x (\theta \alpha x + \theta + 2\alpha) + (\theta^2 + \theta + 2\alpha)} \tag{4.3}$$

and

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

$$= \frac{\alpha (\theta^2 x^2 + 2\theta x + 2) + \theta (\theta + 2\alpha) (\theta x + 1) + (\theta^2 + \theta + 2\alpha)}{\theta [\theta x (\theta \alpha x + \theta + 2\alpha) + (\theta^2 + \theta + 2\alpha)]} \tag{4.4}$$

It can be easily verified that $h(0) = \frac{\theta^3}{\theta^2 + \theta + 2\alpha} = f(0)$ and $m(0) = \frac{\theta^2 + 2\theta + 6\alpha}{\theta(\theta^2 + \theta + 2\alpha)} = \mu_1'$. The graphs of $h(x)$ and $m(x)$ of AGSD (2.1) for different values of its parameters are shown in figures 2 (a) and 2 (b).

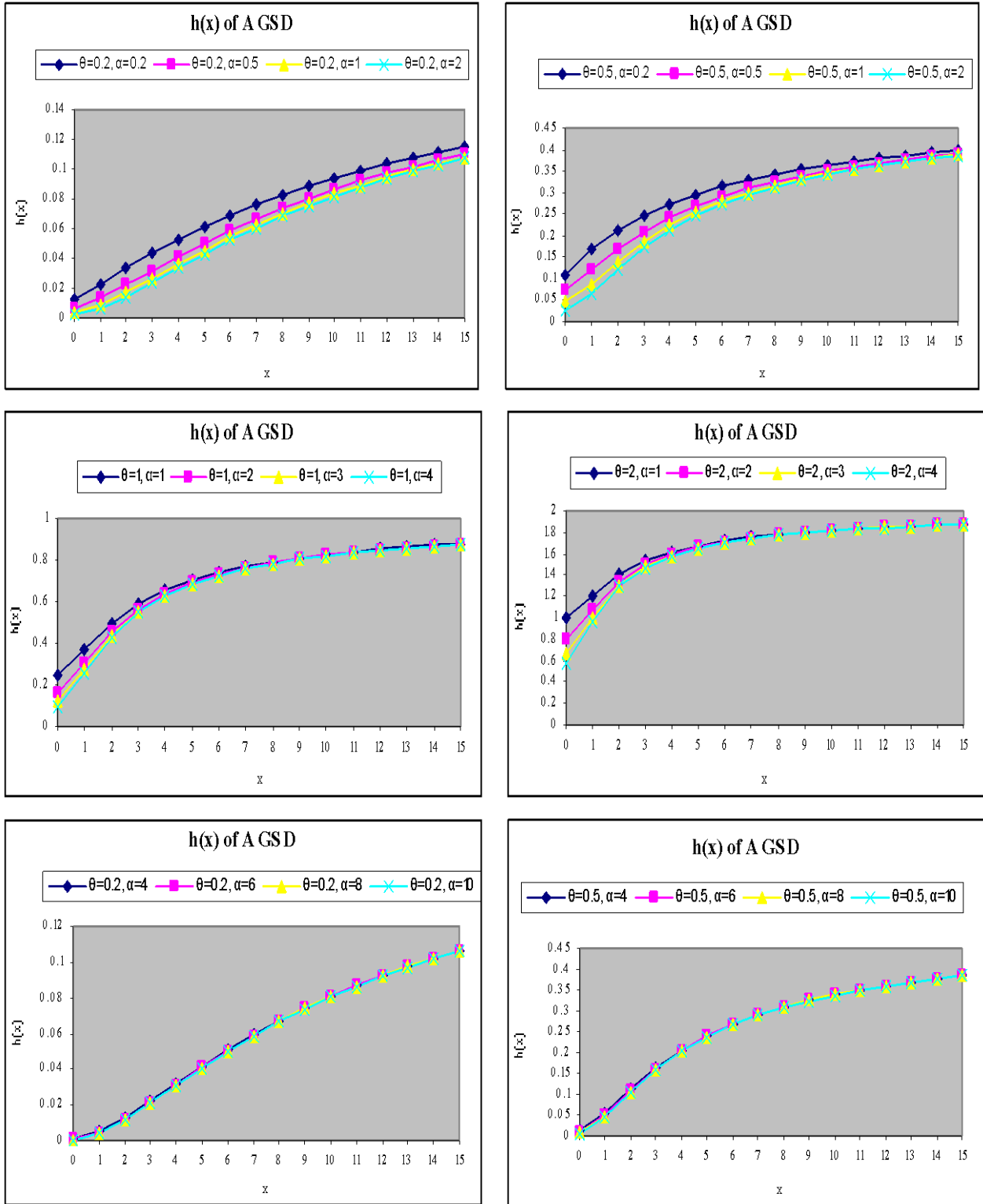


Fig. 2 (a) Graphs of $h(x)$ of AGSD for selected values of parameters θ and α .

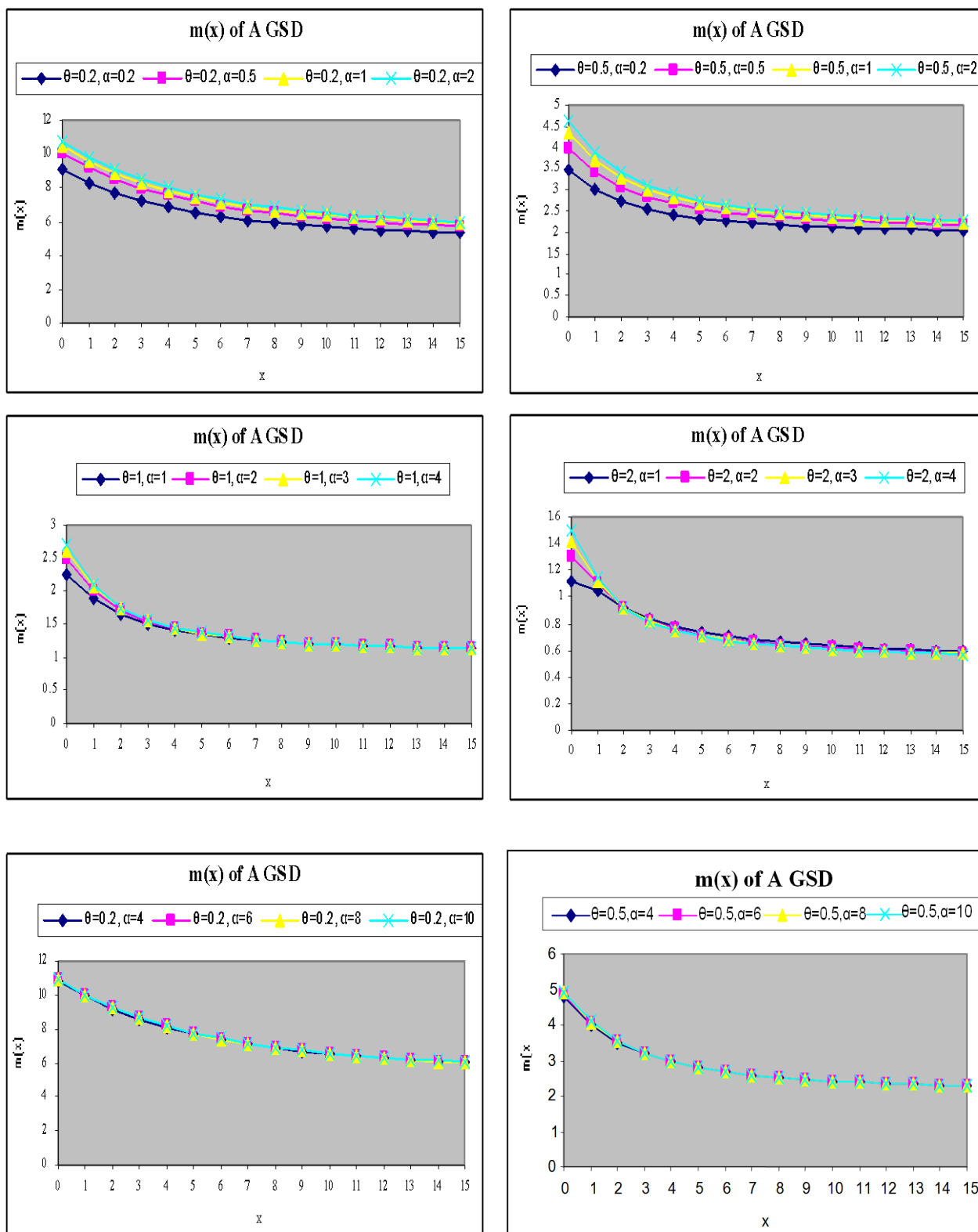


Fig. 2 (b) Graphs of $m(x)$ of AGSD for selected values of parameters θ and α .

It is also obvious from the graphs of $h(x)$ and $m(x)$ that $h(x)$ is monotonically increasing function of x , θ and α whereas $m(x)$ is monotonically decreasing function of x , θ and α .

STOCHASTIC ORDERINGS

Stochastic ordering of positive continuous random variables is an important tool for judging the comparative behaviour of continuous distributions. 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 results 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 AGSD (2.1) is ordered with respect to the strongest ‘likelihood ratio’ ordering as shown in the following theorem:

Theorem: Let $X \sim \text{AGSD}(\theta_1, \alpha_1)$ and $Y \sim \text{AGSD}(\theta_2, \alpha_2)$. If $\theta_1 > \theta_2$ and $\alpha_1 = \alpha_2$ (or $\theta_1 = \theta_2$ and $\alpha_1 < \alpha_2$) then $X \leq_{lr} Y$ and hence $X \leq_{hr} Y$, $X \leq_{mrl} Y$ and $X \leq_{st} Y$.

Proof: We have

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

Now

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

This gives

$$\frac{d}{dx} \ln \frac{f_X(x; \theta_1, \alpha_1)}{f_Y(x; \theta_2, \alpha_2)} = \frac{(\alpha_1 - \alpha_2)x^2 + 2(\alpha_1 - \alpha_2)x}{(1+x+\alpha_1 x^2)(1+x+\alpha_2 x^2)} - (\theta_1 - \theta_2).$$

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

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:

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

where, $\mu = E(X)$ and $M = \text{Median}(X)$.

The measures $\delta_1(X)$ and $\delta_2(X)$ can be calculated using the following relationships

$$\begin{aligned}
 \delta_1(X) &= \int_0^{\mu} (\mu - x)f(x)dx + \int_{\mu}^{\infty} (x - \mu)f(x)dx \\
 &= \mu F(\mu) - \int_0^{\mu} x f(x)dx - \mu [1 - F(\mu)] + \int_{\mu}^{\infty} x f(x)dx \\
 &= 2\mu F(\mu) - 2\mu + 2 \int_{\mu}^{\infty} x f(x)dx \\
 &= 2\mu F(\mu) - 2 \int_0^{\mu} x f(x)dx
 \end{aligned} \tag{6.1}$$

and

$$\begin{aligned}
 \delta_2(X) &= \int_0^M (M - x)f(x)dx + \int_M^{\infty} (x - M)f(x)dx \\
 &= M F(M) - \int_0^M x f(x)dx - M [1 - F(M)] + \int_M^{\infty} x f(x)dx \\
 &= -\mu + 2 \int_M^{\infty} x f(x)dx \\
 &= \mu - 2 \int_0^M x f(x)dx
 \end{aligned} \tag{6.2}$$

Using PDF (2.1) and expression for the mean of AGSD (2.1), we get

$$\begin{aligned}
 \int_0^{\mu} x f_4(x; \theta, \alpha) dx &= \mu - \frac{\left\{ \theta^3 (\alpha \mu^3 + \mu^2 + \mu) + \theta^2 (3\alpha \mu^2 + 2\mu + 1) + 2\theta (3\alpha \mu + 1) + 6\alpha \right\} e^{-\theta \mu}}{\theta (\theta^2 + \theta + 2\alpha)} \\
 \int_0^M x f_4(x; \theta, \alpha) dx &= \mu - \frac{\left\{ \theta^3 (\alpha M^3 + M^2 + M) + \theta^2 (3\alpha M^2 + 2M + 1) + 2\theta (3\alpha M + 1) + 6\alpha \right\} e^{-\theta M}}{\theta (\theta^2 + \theta + 2\alpha)}
 \end{aligned}$$

Using expressions from (6.1), (6.2), (6.3), and (6.4) and after some mathematical simplifications, the mean deviation about the mean, $\delta_1(X)$ and the mean deviation about the median, $\delta_2(X)$ of AGSD (2.1) are obtained as:

$$\delta_1(X) = \frac{2 \left[\theta^2 (\alpha \mu^2 + \mu + 1) + 2\theta (2\alpha \mu + 1) + 6\alpha \right] e^{-\theta \mu}}{\theta (\theta^2 + \theta + 2\alpha)} \tag{6.5}$$

and

$$\delta_2(X) = \frac{2 \left[\theta^3 (\alpha M^3 + M^2 + M) + \theta^2 (3\alpha M^2 + 2M + 1) + 2\theta (3\alpha M + 1) + 6\alpha \right] e^{-\theta M}}{\theta (\theta^2 + \theta + 2\alpha)} - \mu \tag{6.6}$$

BONFERRONI AND LORENZ CURVES AND INDICES

The Bonferroni and Lorenz curves (Bonferroni, 1930) and Bonferroni and Gini indices have applications not only in economics to study income and poverty, but also in other fields like reliability, demography, insurance and medical science. The Bonferroni and Lorenz curves are defined as:

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x) dx = \frac{1}{p\mu} \left[\int_0^\infty x f(x) dx - \int_q^\infty x f(x) dx \right] = \frac{1}{p\mu} \left[\mu - \int_q^\infty x f(x) dx \right] \quad (7.1)$$

and

$$L(p) = \frac{1}{\mu} \int_0^q x f(x) dx = \frac{1}{\mu} \left[\int_0^\infty x f(x) dx - \int_q^\infty x f(x) dx \right] = \frac{1}{\mu} \left[\mu - \int_q^\infty x f(x) dx \right] \quad (7.2)$$

respectively or equivalently

$$B(p) = \frac{1}{p\mu} \int_0^p F^{-1}(x) dx \quad (7.3)$$

and

$$L(p) = \frac{1}{\mu} \int_0^p F^{-1}(x) dx \quad (7.4)$$

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

The Bonferroni and Gini indices are thus defined as:

$$B = 1 - \int_0^1 B(p) dp \quad (7.5)$$

and

$$G = 1 - 2 \int_0^1 L(p) dp \quad (7.6)$$

respectively.

Using PDF of AGSD (2.1), we get

$$\int_q^\infty x f_4(x; \theta, \alpha) dx = \frac{\left\{ \theta^3 (\alpha q^3 + q^2 + q) + \theta^2 (3\alpha q^2 + 2q + 1) + 2\theta (3\alpha q + 1) + 6\alpha \right\} e^{-\theta q}}{\theta (\theta^2 + \theta + 2\alpha)} \quad (7.7)$$

Now using equation (7.7) in (7.1) and (7.2), we get

$$B(p) = \frac{1}{p} \left[1 - \frac{\left\{ \theta^3 (\alpha q^3 + q^2 + q) + \theta^2 (3\alpha q^2 + 2q + 1) + 2\theta (3\alpha q + 1) + 6\alpha \right\} e^{-\theta q}}{\theta^2 + 2\theta + 6\alpha} \right] \quad (7.8)$$

and

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

Now using equations (7.8) and (7.9) in (7.5) and (7.6), the Bonferroni and Gini indices of AGSD (2.1) are thus obtained as:

$$B = 1 - \frac{\left\{ \theta^3 (\alpha q^3 + q^2 + q) + \theta^2 (3\alpha q^2 + 2q + 1) + 2\theta (3\alpha q + 1) + 6\alpha \right\} e^{-\theta q}}{\theta^2 + 2\theta + 6\alpha} \quad (7.10)$$

$$G = \frac{2\{\theta^3(\alpha q^3 + q^2 + q) + \theta^2(3\alpha q^2 + 2q + 1) + 2\theta(3\alpha q + 1) + 6\alpha\}e^{-\theta q}}{\theta^2 + 2\theta + 6\alpha} - 1 \tag{7.11}$$

STRESS-STRENGTH RELIABILITY

The stress- strength reliability of a component illustrates the life of the component which has random strength X that is subjected to a random stress Y . When the stress (Y) of the component applied to it exceeds the strength (X) of the component, 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 applications in almost all areas of knowledge especially in engineering such as structures, deterioration of rocket motors, static fatigue of ceramic components, aging of concrete pressure vessels etc.

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

$$\begin{aligned} R = P(Y < X) &= \int_0^\infty P(Y < X | X = x) f_X(x) dx \\ &= \int_0^\infty f_4(x; \theta_1, \alpha_1) F_4(x; \theta_2, \alpha_2) dx \\ &= 1 - \frac{\left[\begin{aligned} &\theta_2^6 + (4\theta_1 + 3)\theta_2^5 + (6\theta_1^2 + 10\theta_1 + 2\alpha_1 + 6\alpha_2 + 3)\theta_2^4 \\ &+ (4\theta_1^3 + 12\theta_1^2 + 4\alpha_1\theta_1 + 18\alpha_2\theta_1 + 7\theta_1 + 8\alpha_1 + 12\alpha_2)\theta_2^3 \\ &\theta_1^3 + (\theta_1^4 + 6\theta_1^3 + 5\theta_1^2 + 20\alpha_2\theta_1^2 + 2\alpha_1\theta_1^2 + 20\alpha_2\theta_1 + 10\alpha_1\theta_1 + 40\alpha_1\alpha_2)\theta_2^2 \\ &+ (\theta_1^3 + \theta_1^2 + 10\alpha_2\theta_1^2 + 10\alpha_2\theta_1 + 2\alpha_1\theta_1 + 20\alpha_1\alpha_2)\theta_1\theta_2 \\ &+ 2(\alpha_2\theta_1^2 + \alpha_2\theta_1 + 2\alpha_1\alpha_2)\theta_1^2 \end{aligned} \right]}{(\theta_1^2 + \theta_1 + 2\alpha_1)(\theta_2^2 + \theta_2 + 2\alpha_2)(\theta_1 + \theta_2)^5} \end{aligned}$$

It can be easily verified that the above expression reduces to the corresponding expression of Sujatha distribution introduced by Shanker (2016 b) at $\alpha_1 = \alpha_2 = 1$.

MAXIMUM LIKELIHOOD ESTIMATION OF PARAMETERS

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

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

The natural log likelihood function is thus obtained as

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

where \bar{x} is the sample mean.

The maximum likelihood estimates (MLEs) $\hat{\theta}$ and $\hat{\alpha}$ of θ and α are then the solutions of the following non-linear equations

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

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

These two natural log likelihood equations do not seem to be solved directly. However, the Fisher’s scoring method can be applied to solve these equations. We have

$$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{3n}{\theta^2} + \frac{n(2\theta^2 + 2\theta - 4\alpha + 1)}{(\theta^2 + \theta + 2\alpha)^2}$$

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

$$\frac{\partial^2 \ln L}{\partial \theta \partial \alpha} = \frac{2n(2\theta+1)}{(\theta^2 + \theta + 2\alpha)^2} = \frac{\partial^2 \ln L}{\partial \alpha \partial \theta}$$

The following equations can be solved for MLEs $\hat{\theta}$ and $\hat{\alpha}$ of θ and α of AGSD (2.1)

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

where θ_0 and α_0 are the initial values of θ and α , respectively. These equations are solved iteratively till sufficiently close values of $\hat{\theta}$ and $\hat{\alpha}$ are obtained.

APPLICATIONS AND GOODNESS OF FIT

The goodness of fit of a generalization Sujatha distribution (AGSD) using maximum likelihood estimation has been discussed with four real lifetime data sets and the fit has been compared with one parameter Sujatha, Aradhana, Lindley and exponential distributions. The following four real lifetime data sets have been considered for the goodness of fit of considered distributions

Data Set 1: The data set represents the strength of 1.5cm glass fibers measured at the National Physical Laboratory, England. Unfortunately, the units of measurements are not given in the paper, and they are taken from Smith and Naylor (1987)

0.55	0.93	1.25	1.36	1.49	1.52	1.58	1.61	1.64	1.68	1.73	1.81
2.00	0.74	1.04	1.27	1.39	1.49	1.53	1.59	1.61	1.66	1.68	1.76
1.82	2.01	0.77	1.11	1.28	1.42	1.50	1.54	1.60	1.62	1.66	1.69
1.76	1.84	2.24	0.81	1.13	1.29	1.48	1.50	1.55	1.61	1.62	1.66
1.70	1.77	1.84	0.84	1.24	1.30	1.48	1.51	1.55	1.61	1.63	1.67
1.70	1.78	1.89									

Data set 2: This data set represents the lifetime’s data relating to relief times (in minutes) of 20 patients receiving an analgesic and reported by Gross and Clark (1975, P. 105).

1.1 1.4 1.3 1.7 1.9 1.8 1.6 2.2 1.7 2.7 4.1 1.8
 1.5 1.2 1.4 3 1.7 2.3 1.6 2

Data Set 3: This data set is the strength data of glass of the aircraft window reported by Fuller *et al.* (1994):

18.83 20.8 21.657 23.03 23.23 24.05 24.321 25.5 25.52 25.8 26.69 26.77
 26.78 27.05 27.67 29.9 31.11 33.2 33.73 33.76 33.89 34.76 35.75 35.91
 36.98 37.08 37.09 39.58 44.045 45.29 45.381

Data Set 4: 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 the goodness of fit of AGSD, Sujatha, Aradhana, Lindley and exponential distributions, $-2\ln L$, AIC (Akaike Information Criterion), AICC (Akaike Information Criterion Corrected) and BIC (Bayesian Information Criterion) of distributions for four real lifetime data sets have been computed and presented in table 5. The formulae for computing AIC, AICC and BIC are as follows:

$$AIC = -2\ln L + 2k, \quad AICC = AIC + \frac{2k(k+1)}{(n-k-1)}, \quad BIC = -2\ln L + k \ln n, \quad \text{where } k = \text{the number of parameters, } n = \text{the sample size.}$$

Table 5 . MLE’s $-2\ln L$, AIC, AICC and BIC of the fitted distributions of data sets 1, 2, 3 and 4.

	Model	MLE($\hat{\theta}$)	$-2\ln L$	AIC	AICC	BIC
Data 1	AGSD	$\hat{\theta} = 1.9581$ $\hat{\alpha} = 119.826$	110.75	114.75	114.95	119.03
	Sujatha	$\hat{\theta} = 1.3500$	154.80	156.80	156.86	158.94
	Aradhana	$\hat{\theta} = 1.3464$	149.87	151.87	151.93	154.01
	Lindley	$\hat{\theta} = 0.9961$	162.55	164.55	164.62	166.69
	Exponential	$\hat{\theta} = 0.6637$	177.66	179.66	179.73	181.80
Data 2	AGSD	$\hat{\theta} = 1.5712$ $\hat{\alpha} = 222.235$	45.97	49.96	50.67	51.96
	Sujatha	$\hat{\theta} = 1.1367$	57.49	59.49	59.71	60.49
	Aradhana	$\hat{\theta} = 1.1232$	56.37	58.37	58.59	59.36
	Lindley	$\hat{\theta} = 0.8161$	60.49	62.49	62.71	63.49
	Exponential	$\hat{\theta} = 0.5263$	65.67	67.67	67.89	68.67
Data 3	AGSD	$\hat{\theta} = 0.0972$ $\hat{\alpha} = 14.473$	240.54	244.54	244.68	243.97

	Sujatha	$\hat{\theta} = 0.0956$	241.50	243.50	243.63	244.93
	Aradhana	$\hat{\theta} = 0.0943$	242.22	244.22	244.36	245.65
	Lindley	$\hat{\theta} = 0.0630$	253.98	255.98	256.11	257.41
	Exponential	$\hat{\theta} = 0.0324$	274.52	276.52	276.66	277.95
Data 4	AGSD	$\hat{\theta} = 1.2129$ $\hat{\alpha} = 53.567$	186.35	190.35	190.53	194.82
	Sujatha	$\hat{\theta} = 0.9361$	221.60	223.60	223.66	225.83
	Aradhana	$\hat{\theta} = 0.9170$	219.90	221.90	221.96	224.13
	Lindley	$\hat{\theta} = 0.6545$	238.38	240.38	240.44	242.61
	Exponential	$\hat{\theta} = 0.4079$	261.73	263.73	263.79	265.96

The best fit of the distribution is the distribution which corresponds to the lower values of $-2\ln L$, AIC, AICC and BIC. It is obvious from the goodness of fit of distributions for four data sets in the table 5 that AGSD provides better fit than Sujatha, Aradhana, Lindley and exponential distributions for modeling lifetime data.

CONCLUSION

A generalization of Sujatha distribution (AGSD) has been introduced which includes Sujatha distribution, proposed by Shanker (2016c) and Lindley distribution, proposed by Lindley (1958) 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 AGSD 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 maximum likelihood estimation has been discussed for estimating parameters. Four examples of real lifetime data sets have been presented to show the applications and goodness of fit of AGSD over Sujatha, Aradhana, Lindley and exponential distributions and it has been observed that AGSD gives much better fit.

ACKNOWLEDGEMENT

Authors are thankful to the learned referee for the valuable comments to improve the quality of paper.

REFERENCES

- Bader, M. G. and Priest, A. M. (1982). Statistical aspects of fiber and bundle strength in hybrid composites. *In: Hayashi, T., Kawata, K. Umekawa, S. (eds), Progress in Science in Engineering Composites, ICCM-IV, Tokyo, 1129-1136.*
- Bakouch, H. S.; Al-Zaharani, B.; Al-Shomrani, A., Marchi, V. and Louzad, F. (2012). An extended Lindley distribution. *Journal of the Korean Statistical Society, 41: 75-85*
- Bonferroni, C. E. (1930). *Elementi di Statistica Generale*, Seeber, Firenze.
- Deniz, E. and Ojeda, E. (2011). The discrete Lindley distribution-Properties and Applications. *Journal of Statistical Computation and Simulation, 81: 1405-1416.*
- Fuller, E. J.; Frieman, 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, M. E.; Atieh, B. and Nadarajah, S. (2008). Lindley distribution and its Application. *Mathematics Computing and Simulation, 78: 493- 506.*
- Ghitany, M.; Al-Mutairi, D.; Balakrishnan, N. and Al-Enezi, I. (2013). Power Lindley distribution and associated inference.

- Computational Statistics and Data Analysis* **64**: 20-33.
- Gross, A. J. and Clark, V. A. (1975). *Survival Distributions: Reliability Applications in the Biometrical Sciences*, John Wiley, New York.
- Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society, Series B*, **20**: 102- 107.
- Nadarajah, S.; Bakouch, H. S. and Tahmasbi, R. (2011). A generalized Lindley distribution. *Sankhya Series B*, **73**: 331- 359.
- Shaked, M. and Shanthikumar, J. G. (1994). *Stochastic Orders and their Applications*, Academic Press, New York.
- Shanker, R. (2016 a). On Generalized Lindley distribution and its applications to model lifetime data from biomedical science and engineering. *Insights in Biomedicine*, **1**(2): 1-6.
- Shanker, R. (2016 b). Aradhana distribution and its Applications. *International Journal of Statistics and Applications*, **6** (1): 23-34.
- Shanker, R. (2016 c). Sujatha distribution and its Applications. *Statistics in Transition-New Series*, **17** (3): 1-20.
- Shanker, R. (2016 d). The discrete Poisson-Sujatha distribution. *International Journal of Probability and Statistics*, **5** (1): 1-9.
- Shanker, R. (2016 e). A quasi Sujatha distribution. *International Journal of Statistics and Applications*, **6** (6): 376–385.
- Shanker, R. and Hagos, F. (2016). Size-biased Poisson-Sujatha distribution with Applications. *American Journal of Mathematics and Statistics*, **6** (4): 145–154.
- Shanker, R.; Hagos, F., and Sujatha, S. (2015 a). On modeling of Lifetimes data using exponential and Lindley distributions. *Biometrics & Biostatistics International Journal*, **2** (5): 1-9.
- Shanker, R. and Hagos, F. (2015). Zero-truncated Poisson-Sujatha distribution with applications, *Journal of Ethiopian Statistical Association*, **24**: 55–63.
- Shanker, R.; Hagos, F. and Sharma, S. (2016 a). On two parameter Lindley distribution and Its Applications to model Lifetime data. *Biometrics & Biostatistics International Journal*, **3** (2): 1-8.
- Shanker, R.; Hagos, F. and Sharma, S. (2016 b). On Quasi Lindley distribution and its applications to model lifetime data. *International Journal of Statistical distributions and Applications*, **2** (1): 1-7.
- Shanker, R.; Shukla, K. K. and Hagos, F. (2016 c). On Weighted Lindley distribution and applications to model Lifetime data. *Jacobs Journal of Biostatistics*, **1** (1): 1-9.
- Shanker, R. and Mishra, A. (2013 a). A two-parameter Lindley distribution. *Statistics in Transition-new series*, **14** (1): 45-56.
- Shanker, R. and Mishra, A. (2013 b). A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research*, **6** (4): 64-71.
- Shanker, R. and Amanuel, A. G. (2013). A new quasi Lindley distribution. *International Journal of Statistics and Systems*, **8** (2): 143-156.
- Shanker, R.; Sharma, S. and Shanker, R. (2013). A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics*, **4**: 363-368.
- Shanker, R.; Shukla, K. K.; Shanker, R. and Tekie, A. L. (2017). A Three-parameter Lindley Distribution. *American Journal of Mathematics and Statistics*, **7** (1): 15-26.
- Smith, R. L and Naylor, J. C. (1987). A comparison of Maximum likelihood and Bayesian estimators for the three parameter Weibull distribution. *Applied Statistics*, **36**: 358-369.
- Zakerzadeh, H. and Dolati, A. (2009). Generalized Lindley distribution. *Journal of Mathematical extension*, **3**(2): 13-25.