SIA-Lognormal Power Distribution

Safoora Samuel
Student of MPhil Statistics
Kinnaird College for Women
Lahore, Pakistan
e-mail: safoorasamuel@gmail.com

Saleha Naghmi Habibullah
Professor of Statistics
Kinnaird College for Women
Lahore, Pakistan
e-mail: salehahabibullah@gmail.com

***Abstract*— During the past few years, some researchers have worked on distributions that are invariant under the reciprocal transformation. Such distributions are now referred to as being Self-Inverse at Unity. Only very recently, a generalized version of this class of distributions has been introduced --- distributions "Self-Inverse at A" where A, an arbitrary positive number, represents the median of the distribution. The self-inversion property permits development of estimators of distribution parameters that are more efficient than their well-known counterparts. The lognormal distribution with scale parameter zero belongs to the class of SIU distributions. In this paper, we obtain the "SIA-Lognormal-Power distribution" and derive some of its fundamental properties such as the first four moments, the quantile function and the hazard function. The shape of the density provides optimism that this newly derived probability model will turn out to be a suitable candidate for modeling a variety of real-life data-sets and the fact that it belongs to the class of SIA distributions will enable efficient estimation of the shape parameter of this distribution.**

Keywords-Self-Inverse distributions; Lognormal distribution; SIA-Lognormal-Power distribution; moments; hazard function.

I. INTRODUCTION

A number of authors have focused on distributions that are invariant under the reciprocal transformation. (See [1], [2] and [3].) The nomenclature "Self-Inverse at Unity (SIU)" has been adopted for this class of distributions in [4]. A generalized version of SIU distributions has been given in [5] which have been called "Self-Inverse at A (SIA)" where A is an arbitrary positive number.

The remarkable property of SIA distributions is that, due to self-inversion, it is possible to modify the formulae of well-known estimators in order to obtain estimators of distribution parameters that are more efficient than the well-known ones.

II. DISTRIBUTIONS SELF-INVERSE AT UNITY

The self-inversion property can be defined as that property by which reciprocal of a non negative continuous random variable possesses exactly the same distribution as the one possessed by the original random variable. One of the fundamental properties of this class of distributions is that the $(1-p)^{th}$ quantile is the reciprocal of the pth quantile and the median is equal to unity. Some simple examples are the half Cauchy distribution, the F distribution having $\mu_1 = \mu_2$ and lognormal distribution where $\mu = 0$. It is well-known that each of these distributions finds applications in a number of areas.

III. DISTRIBUTIONS SELF-INVERSE AT 'A'

The distribution of a non-negative continuous random variable X will be regarded as being self-inverse at A if the distribution of X/A is identical to the distribution of A/X where A is an arbitrary positive real number. The median of this distribution will be equal to A.

The property that the median of every SIU distribution is unity is, in fact, a limitation of this class of distributions. 'A' being an arbitrary positive number, it is obvious that the class of SIA distributions is much wider than the class of SIU distributions.

IV. LOGNORMAL DISTRIBUTION

The lognormal distribution is one of the most well-known distributions all over the world and finds applications in a wide variety of disciplines including economics, finance, biology, medicine, engineering and human behaviors as well.

The probability density function of lognormal distribution is given by

$$g_y(y;\mu,\sigma) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y-\mu)^2}{2\sigma^2}}, \ y>0 \qquad (1)$$

where μ can be regarded as the scale parameter and σ the shape parameter.

V. LOGNORMAL POWER DISTRIBUTION

It has been shown in [1] that application of the power transformation to an SIU distribution results in another SIU distribution. In this section, we apply the power transformation to the lognormal distribution with scale parameter equal to zero.

Applying the transformation $Z = Y^r$ to the lognormal distribution given in eq. (1), we obtain the following probability density function:

$$w(z) = \frac{1}{zr\sigma\sqrt{2\pi}}.e^{-\frac{1}{2r^2\sigma^2}(\ln z)^2}, \ 0<z<\infty, \ r>0 \qquad (2)$$

We call it the Lognormal Power distribution. It is easy to verify that this distribution is self-inverse at unity.

VI. SIA-LOGNORMAL POWER DISTRIBUTION

Applying the transformation $X = \mathrm{A}Z$ to the Lognormal Power distribution given in eq. (2), we obtain the probability density function

$$f(x) = \frac{1}{xr\sigma\sqrt{2\pi}} e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2}, \quad 0 < x < \infty \qquad (3)$$

where

$$\sigma > 0, r > 0, \mathrm{A} > 0.$$

for which we adopt the nomenclature "SIA-Lognormal Power distribution". The graph of the density is given in Figure 1.

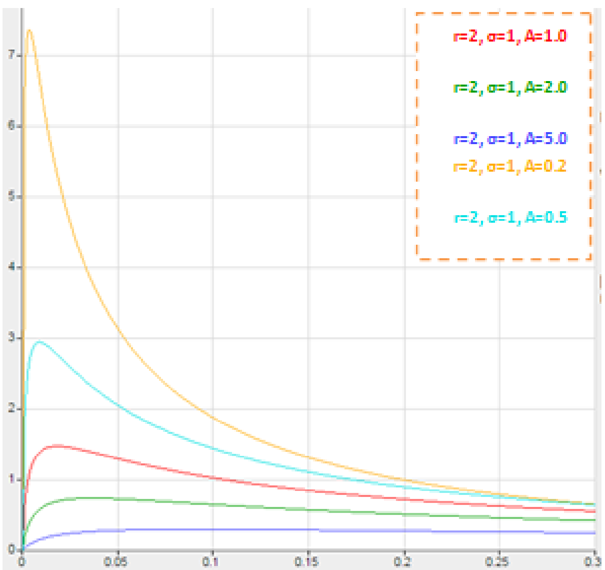

Fig. 1: Graph of the density function of the SIA- Lognormal-Power distribution

Clearly, the shape of the density function is unimodal and positively skewed.

The cumulative distribution function is given by

$$\frac{1}{2}\left[erf\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma} \right) + 1 \right] \qquad (4)$$

The error function $erf(x)$ is defined as follows:

$$erf(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

where

$$\int_0^x e^{-t^2} dt = x - \frac{x^3}{3} + \frac{x^5}{5.2!} - \frac{x^7}{7.3!} + \ldots\ldots$$

VII. FUNDAMENTAL PROPERTIES

In this section, we present some of the fundamental properties of the SIA-Lognormal Power distribution. We begin with the well-known measures of central tendency.

A. *Arithmetic Mean*

The mean of the distribution is given by

$$E(X) = \int_0^\infty \frac{x}{xr\sigma\sqrt{2\pi}} e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2} dx = \mathrm{A}e^{\frac{r^2\sigma^2}{2}} \qquad (5)$$

B. *Geometric Mean*

The logarithm of the geometric mean G_X of a distribution with random variable X is the expected value of ln(X). As such, we have

$$E[\ln X] = \int_0^\infty \ln x . \frac{1}{xr\sigma\sqrt{2\pi}} e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2} dx = \ln(\mathrm{A})$$

Therefore, we have

$$G_X = \mathrm{A} \qquad (6)$$

It is interesting to note that the geometric mean of the distribution is equal to the median.

C. *Harmonic Mean*

The harmonic mean (H_X) of a distribution of the random variable X is the reciprocal of the expected value of 1/X. Therefore, we have

$$H_X = \left[\int_0^\infty \frac{1}{x^2 r\sigma\sqrt{2\pi}} e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{A}\right)\right)^2} dx \right]^{-1} = \mathrm{A}e^{-\frac{r^2\sigma^2}{2}} \qquad (7)$$

It is interesting to note that the arithmetic and harmonic means are related by the equation:

$$\frac{\mathrm{AM}}{\mathrm{A}} = \frac{\mathrm{A}}{\mathrm{HM}}$$

VIII. QUANTILE FUNCTION

The quantile function is one way of describing a probability distribution, and it is an alternative to the probability density function (pdf), the cumulative distribution function (CDF) and the characteristic function.

By definition, the qth quantile is obtained by solving for X_q the equation

$$\int_0^{X_q} f(x)dx = q$$

As such, we obtain

$$X_q = \mathrm{A}\exp\left(\left[erfinv(2q-1)\right]\sqrt{2}r\sigma\right) \tag{8}$$

The first quartile and the third quartile of the distribution come out to be

$$Q_1 = \mathrm{A}e^{-0.67449r\sigma} \tag{9}$$

and

$$Q_3 = \mathrm{A}e^{0.67449r\sigma} \tag{10}$$

It is noteworthy that the first and third quartiles are related by the equation:

$$\frac{Q_1}{\mathrm{A}} = \frac{\mathrm{A}}{Q_3}$$

IX. MEASURES OF DISPERSION

The variance and the standard deviation are regarded as some of the most important measures of dispersion. The variance of SIA-Lognormal Power distribution is derived below:

$$E(X^2) = \int_0^\infty x^{\not 2}.\frac{1}{\not x r\sigma\sqrt{2\pi}} e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2} dx = \mathrm{A}^2 e^{2r^2\sigma^2}$$

Therefore

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \left[\mathrm{E}(X)\right]^2 = \mathrm{A}^2 e^{r^2\sigma^2}\left(e^{r^2\sigma^2}-1\right)$$

As such, the standard deviation of the SIA-Lognormal-Power distribution is given by

$$\mathrm{S.D.} = \sqrt{\mathrm{A}^2 e^{r^2\sigma^2}\left(e^{r^2\sigma^2}-1\right)} = \mathrm{A}e^{r\sigma}\sqrt{e^{r^2\sigma^2}-1} \tag{11}$$

The coefficient of variation is given by

$$\mathrm{C.V.} = \frac{\mathrm{S.D.}}{\mathrm{Mean}} = \frac{\mathrm{A}e^{r\sigma}\sqrt{e^{r^2\sigma^2}-1}}{\mathrm{A}e^{\frac{r^2\sigma^2}{2}}}$$

As such, we have

$$\mathrm{C.V.} = e^{r\sigma\left(1-\frac{r\sigma}{2}\right)}\sqrt{e^{r^2\sigma^2}-1} \tag{12}$$

It is interesting to note that the coefficient of variation of the SIA-Lognormal Power distribution is independent of A.

X. MODE

By definition, the mode is obtained by equating the first derivative of the density function to zero. Here, we have

$$f'(x) = -\left[\frac{\sqrt{2\pi}e^{\frac{\left(\ln\left(\frac{x}{A}\right)\right)^2}{2r^2\sigma^2}}\left\{\ln\left(\frac{x}{\mathrm{A}}\right)+r^2\sigma^2\right\}}{r\sigma\left\{xr\sigma\sqrt{2\pi}e^{\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2}\right\}^2}\right] = 0$$

Hence, the mode of the distribution is given by

$$\hat{X} = Ae^{-r^2\sigma^2} \tag{13}$$

XI. HIGHER MOMENTS AND MOMENT-RATIOS

In this section, we obtain the third and fourth moments of the SIA-Lognormal Power distribution.

The third and fourth moments about the origin come out to be

$$E(X^3) = A^3 e^{\frac{9r^2\sigma^2}{2}}$$

and

$$E(X^4) = A^4 e^{8r^2\sigma^2}$$

The third and fourth moments about the mean are given by

$$\mu_3 = \mu_3' - 3\mu_2'\mu_1' + 2\left(\mu_1'\right)^3$$

and

$$\mu_4 = \mu_4' - 4\mu_3'\mu_1' + 6\mu_2'\left(\mu_1'\right)^2 - 3\left(\mu_1'\right)^4$$

As such, we have

$$\mu_3 = \mathrm{A}^3\left(e^{\frac{9r^2\sigma^2}{2}} - 3e^{\frac{5r^2\sigma^2}{2}} + 2e^{\frac{3r^2\sigma^2}{2}}\right) \tag{14}$$

and

$$\mu_4 = \mathrm{A}^4\left(e^{8r^2\sigma^2} - 4e^{5r^2\sigma^2} + 6e^{3r^2\sigma^2} - 3e^{2r^2\sigma^2}\right) \quad (15)$$

The two moment ratios are given by $\beta_1 = (\mu_3)^2/(\mu_2)^3$ and $\beta_2 = \mu_4/(\mu_2)^2$. For the SIA-Lognormal Power distribution, these come out to be

$$\beta_1 = \frac{\left(e^{\frac{9r^2\sigma^2}{2}} - 3e^{\frac{5r^2\sigma^2}{2}} + 2e^{\frac{3r^2\sigma^2}{2}}\right)^2}{\left(e^{r^2\sigma^2}\left(e^{r^2\sigma^2} - 1\right)\right)^3} \quad (16)$$

and

$$\beta_2 = \frac{e^{8r^2\sigma^2} - 4e^{5r^2\sigma^2} + 6e^{3r^2\sigma^2} - 3e^{2r^2\sigma^2}}{\left(e^{r^2\sigma^2}\left(e^{r^2\sigma^2} - 1\right)\right)^2} \quad (17)$$

The moment-ratios of the SIA-Lognormal Power distribution are independent of A.

XII. QUANTILES-BASED MEASURES OF CENTRE, SPREAD, SKEWNESS AND KURTOSIS

In this section, we obtain measures of central tendency, dispersion, skewness and kurtosis based on the quantiles of the SIA-Lognormal Power distribution.

The Mid-Quartile Range is given by

$$\frac{Q_1 + Q_3}{2} = \frac{\mathrm{A}\left(e^{-0.67449r\sigma} + e^{0.67449r\sigma}\right)}{2}$$

The Inter-Quartile Range is

$$Q_3 - Q_1 = \mathrm{A}\left(e^{0.67449r\sigma} - e^{-0.67449r\sigma}\right)$$

The Bowley's Coefficient of Skewness is given by

$$Sk = \frac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1} = \frac{e^{-0.67449r\sigma} + e^{0.67449r\sigma} - 2}{e^{0.67449r\sigma} - e^{-0.67449r\sigma}} \quad (18)$$

The Percentile Coefficient of Kurtosis is

$$\frac{Q_3 - Q_1}{2\left(D_9 - D_1\right)} = \frac{e^{0.67449r\sigma} - e^{-0.67449r\sigma}}{2\left(e^{1.28155r\sigma} - e^{-1.28155r\sigma}\right)} \quad (19)$$

The formulae of the Bowley's Coefficient of Skewness and the Percentile Coefficient of Kurtosis of the SIA-Lognormal Power distribution do not involve A.

XIII. SURVIVAL AND HAZARD FUNCTIONS

The survival function is defined as

$$S(x) = 1 - F(x)$$

Here, we have

$$S(x) = \frac{1}{2}\left[1 - erf\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma}\right)\right]$$

or in other words

$$S(x) = \frac{1}{2}erfc\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma}\right) \quad (20)$$

where the complementary error function $erfc(x)$ is defined as

$$erfc(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty e^{-t^2}dt$$

The hazard function is defined as

$$h(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{S(x)}$$

Here, we have

$$h(x) = \frac{2e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2}}{xr\sigma\sqrt{2\pi}\left\{\left[erfc\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma}\right)\right]\right\}} \quad (21)$$

The graph of the hazard function is given in Figure 2. The graph shows upside down bathtub shaped hazard rate. This is also called unimodal hazard rate.

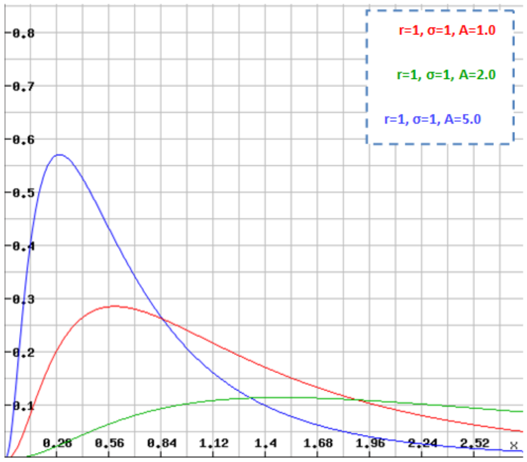

Fig. 2: Graph of the hazard function of the SIA- Lognormal-Power distribution

By definition, the Cumulative Hazard Function is given by

$$\Lambda(x) = -\log S(x)$$

As such, the CHF of the SIA-Lognormal Power distribution is

$$\Lambda(x) = -\left[\log(1) - \log(2) + \log\left\{1 - erf\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma}\right)\right\}\right]$$

or

$$\Lambda(x) = \log(2) - \log\left\{erfc\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma}\right)\right\} \qquad (22)$$

The Reverse Hazard Rate is defined as

$$r(x) = \frac{f(x)}{F(x)}$$

For the SIA-Lognormal Power distribution, we have

$$r(x) = \frac{2e^{-\frac{1}{2r^2\sigma^2}\left(\ln\left(\frac{x}{\mathrm{A}}\right)\right)^2}}{xr\sigma\sqrt{2\pi}\left\{\left[erf\left(\frac{\ln\left(\frac{x}{\mathrm{A}}\right)}{\sqrt{2}r\sigma}\right)+1\right]\right\}} \qquad (23)$$

XIV. CONCLUDING REMARKS

In this paper, we have developed the SIA-Lognormal Power distribution which can be regarded as a generalization of the lognormal distribution having scale parameter μ equal to zero. Some of the fundamental properties of the distribution such as moments and moment-ratios, quartiles and deciles, survival function and hazard function have been obtained. The shape of the density being unimodal and positively skewed, it can be expected that this newly derived probability distribution will turn out to be a pertinent model for real-life data-sets exhibiting an upside down bathtub-shaped hazard rate. More importantly, the self-inversion property provides the capability of developing an SIA-estimator of σ that will be more efficient than the estimator obtained by the ordinary method of moments. This work is under way.

REFERENCES

[1] Seshadri, V. (1965). On Random Variables which have the Same Distributions as their Reciprocals. *Can. Math. Bull*., 8(6), 819-824.

[2] Saunders, S.C. (1974). A Family of Random Variables Closed Under Reciprocation . *J. Amer. Statist Assoc*., 69(346),533-539.

[3] Habibullah, S.N., Memon, A.Z. and Ahmad, M. (2010).*On a Class of Distributions Closed Under Inversion*, Lambert Academic Publishing (LAP), ISBN 978-3-8383-4868-1

[4] Habibullah, S.N. and Saunders, S.C. (2011), "A Role for Self-Inversion", *Proceedings of International Conference on Advanced Modeling and Simulation (ICAMS, Nov 28-30, 2011)* published by Department of Mechanical Engineering, College of Electrical and Mechanical Engineering, National University of Science and Technology (NUST), Islamabad, Pakistan, Copyright 2011, ISBN 978-869-8535-11-7.

[5] Habibullah, S.N. and Fatima, S.S. (2015), "On a Newly Developed Estimator for More Accurate Modeling with an Application to Civil Engineering", Proceedings of the 12th International Conference on Applications of Statistics and Probability in Civil Engineering (ICASP12) organized by CERRA (Vancouver, BC, Canada, July 12-15, 2015). *Sponsoring Agency: Higher Education Commission, Pakistan.*