_____

Skew Arcsine Distribution

Nuri Celik[1*]

[1]*Department of Mathematics, Gebze Technical University, 41400, Kocaeli, Turkey.*

Author's contribution

The sole author designed, analysed, interpreted and prepared the manuscript.

Original Research Article

_____

Abstract

The arcsine distribution is very important tool in statistics literature especially in Brownian motion studies. However, modelling real data sets, even when the potential underlying distribution is pre-defined, is very complicated and difficult in statistical modelling. For this reason, we desire some flexibility on the underlying distribution. In this study, we propose a new distribution obtained by arcsine distribution with Azzalini's skewness procedure. The main characteristics of the proposed distribution are determined both with theoretically and simulation study.

1 Introduction

A random variable *X* has an arcsine distribution if it has the following probability density function,

$$f(x) = \frac{1}{\pi\sqrt{(x-a)(b-x)}}, \qquad a \leq x \leq b \tag{1}$$

_____

Corresponding author: Email: nuricelik@gtu.edu.tr;

The arcsine distribution has several application areas in statistical theory. In the studies of [1,2] and [3], it is shown that, when the Lindeberg-Levy conditions are satisfied, the limiting distribution of simple random walk is arcsine distribution. The other characteristics of arcsine distribution are undertaken in the studies of [4,5,6] and [7].

The arcsine distribution has the following cumulative distribution function,

$$F(x) = \frac{2}{\pi} arcsine\left(\sqrt{\frac{x-a}{b-a}}\right) \qquad a \leq x \leq b \tag{2}$$

where $a, b \in \mathcal{R}$. The expected value and the variance of the distribution is given by

$$E(X) = \frac{a+b}{2} \qquad and \qquad V(X) = \frac{1}{8}(b-a)^2$$

respectively. Besides that, it has the following properties:

- It is symmetric about $\frac{a+b}{2}$
- It is decreasing on $[a, \frac{a+b}{2})$ and increasing on $(\frac{a+b}{2}, b]$
- $f(x) \rightarrow \infty$ as $x \downarrow a$ and $x \uparrow b$

The shape of the arcsine distribution has "u" form. Different probability density functions (pdf) and cumulative distribution (cdf) are shown in Figs. 1 and 2 respectively.
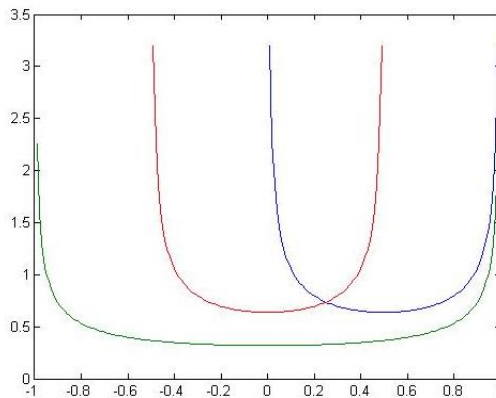

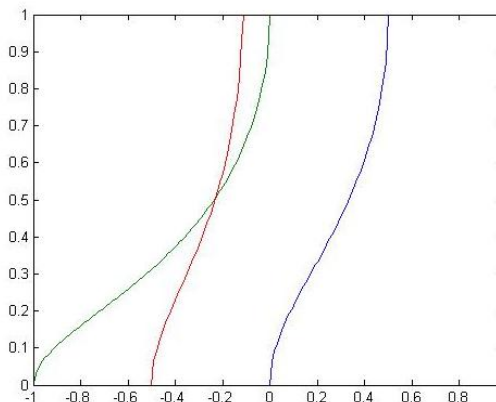**Fig. 1. Probability density functions of arcsine distribution**

Fig. 2. Cumulative distribution functions of arcsine distribution

In recent years, there has been a great interest models for skewed distribution. It allows much better and more flexible fit for data in the presence of the normality (symmetry) departures. [8] realized the importance of skew distributions and their applications in sociology, economics, and in many biological phenomenons, [9]. For this reason, [10,11] has introduced the skew symmetric distribution with normal distribution. Many other symmetric distributions have been studied by using Azzalini's skewness procedure, see [12,13] and [14]. The skewness procedure goes as follows:

Let *g* a probability density function symmetric about zero and *G* scalar distribution function,

$$f(x) = 2g(x)G(\lambda x)$$

is also a probability density function for $\lambda \epsilon \mathcal{R}$. (Azzalini, 1985, 1986). Here λ is the skewness parameter. When $\lambda = 0$, *f(x)* becomes original *g(x)*. [5] developed an alternative process for generating skew distributions. They defined a family of ν-spherical distributions. In this study, we propose a new distribution based on Azzalini's skewness procedure namely skew arcsine distribution. Therefore, let *f(x)* pdf of arcsine distribution ranged in *[-a, a]* and *F(x)* the cdf of it, then $2f(x)F(\lambda x)$ is the new skew arcsine distribution. The aim of proposing of this new distribution is to model stochastic process with skew arcsine distribution.

The rest of the paper organizes as follows: In section 2, we define skew-arcsine distribution, and then we derive the cdf, moments and other certain properties. The parameter estimation of skew arcsine distribution is given in the last part of the section.

2 Skew Arcsine Distribution

Definition: Let *g* an arcsine distribution defined in $[-a, a]$ and *G* distribution function of mentioned distribution,

$$f(x) = \frac{4}{\pi^2} \frac{1}{\sqrt{(a+x)(a-x)}} arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right) \qquad -a < x < a \tag{3}$$

is skew arcsine distribution for $\lambda \in \mathcal{R}$. Then we say that *X* has skew arcsine distribution, we write $X \sim SA(a, \lambda)$. It should be noted that the pdf of $SA(a, \lambda)$ can be found for any skewness parameter, however it can be defined in real number for only $|\lambda| \leq 1$. Fig. 3 illustrates the shape of the $SA(a, \lambda)$ for different λ and *(a=1)*.
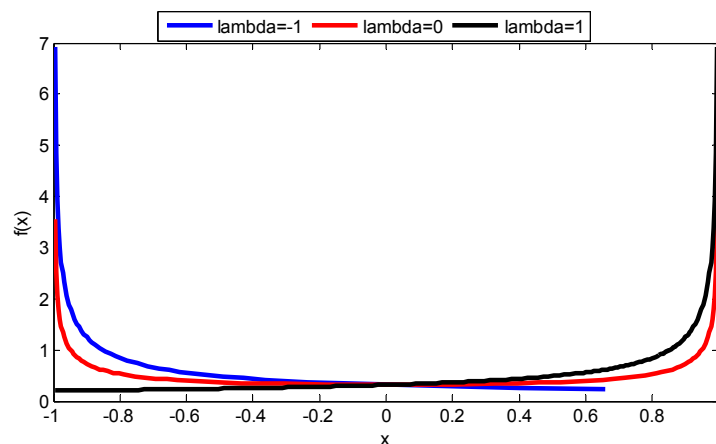

Fig. 3. The probability density functions of skew arcsine distribution

As can be noticed in Fig. 3, the pdf is becoming skew when the skewness parameter is increasing. It is an expected result because of the skew symmetric distributions characteristics.

2.1 Cumulative distribution function

In this section, we derive the cdf of $SA(a, \lambda)$. $F(x)$ of the distribution can be found by the following formula.

$$F(x) = \int_{-a}^{x} \frac{4}{\pi^2} \frac{1}{\sqrt{(a+x)(a-x)}} arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right) dx \tag{4}$$

The integral can be evaluated by partial integration method taking $u = arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right)$ and $= \frac{dx}{\sqrt{a^2 - x^2}}$. Then the distribution function of skew arcsine distribution is

$$F(x) = \frac{4}{\pi^2}\left[arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right) arcsine\frac{x}{a} + \frac{\pi}{2} arcsine\left(\sqrt{\frac{a(1-\lambda)}{2a}}\right)\right] - I \tag{5}$$

where $I = \frac{\lambda}{2}\int_{-a}^{x} arcsine\left(\frac{x}{a}\right)\frac{1}{\sqrt{a^2 - (\lambda x)^2}} dx.$

The integral I cannot be derived analytically due to the structure of the arcsine distribution. Therefore, we benefit from Gauss-Kronrod Quadrature known as an adaptive Gaussian Quadrature method for numerical integration with Matlab programming. Fig. 4 illustrates various cdf of skew arcsine distributions with different skewness parameters in the range [1,1] for the sake of brevity.
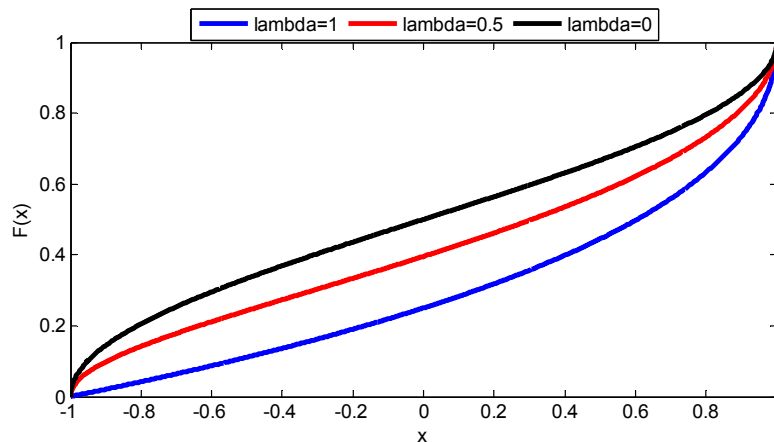

Fig. 4. The Cumulative Distribution Functions of Skew Arcsine Distribution

2.2 Moments

Here, we derive the moments of (3). Because of the structure of skew symmetric distributions, the even moments are equal to the original distributions. Therefore, the even moments of skew arcsine distribution is same as arcsine distributions. The *nth* moment of skew arcsine distribution is;

$$E(X^n) = \int_{-a}^{a} x^n \frac{4}{\pi^2} \frac{1}{\sqrt{(a+x)(a-x)}} arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right) dx \tag{6}$$

The first moment is

$$E(X) = \int_{-a}^{a} x \frac{4}{\pi^2} \frac{1}{\sqrt{(a+x)(a-x)}} arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right) dx \tag{7}$$

The integral can be found by taking $u = arcsine\sqrt{\frac{\lambda x + a}{2a}}$ and $= \frac{x}{\sqrt{a^2-x^2}}dx$. Then, the solution

$$E(X) = \frac{2\lambda}{\pi^2}\int_{-a}^{a}\sqrt{\frac{a^2-x^2}{a^2-(\lambda x)^2}}\,dx \tag{8}$$

When $\lambda = 0$ is taken, then $E(X)=0$ is found. It should be noted that when $\lambda = 0$, $SA(a, \lambda)$ becomes standard arcsine distribution. If λ is equal to 1, then the expected value of $SA(a, 1)$ is $\frac{4a}{\pi^2}$. If λ is equal to -1, the solution is $-\frac{4a}{\pi^2}$. It is also a proof for $SA(-\lambda) = SA(\lambda)$.

The second moment,

$$E(X^2) = \int_{-a}^{a} x^2 \frac{4}{\pi^2}\frac{1}{\sqrt{(a+x)(a-x)}}arcsine\left(\sqrt{\frac{\lambda x + a}{2a}}\right)dx \tag{9}$$

can be solved by taking $u = arcsine\sqrt{\frac{\lambda x + a}{2a}}$ and $dv = \frac{x^2}{\sqrt{a^2-x^2}}dx$. The integral of dv is solved by using

$$\int \frac{P_n(x)}{\sqrt{ax^2+bx+c}}dx = Q_{n-1}(x)\sqrt{ax^2+bx+c} + \delta\int\frac{dx}{\sqrt{ax^2+bx+c}} \tag{10}$$

where P_n and Q_{n-1} are the polynomials of order n and n-1 respectively. Then the second moment of $SA(a, \lambda)$ is

$$E(X^2) = \frac{a^2}{2} \tag{11}$$

Then the variance is found $\frac{1}{2}a^2$ when λ is equal to 0 which is same as arcsine distribution. In a similar way the third and fourth moments are,

$$E(X^3) = \frac{2\lambda}{\pi^2}\left[\frac{1}{3}\int_{-a}^{a} x^2\sqrt{\frac{a^2-x^2}{a^2-(\lambda x)^2}}\,dx + \frac{2}{3}a^2\int_{-a}^{a}\sqrt{\frac{a^2-x^2}{a^2-(\lambda x)^2}}\,dx\right]$$

$$E(X^4) = \frac{3}{8}a^4 \tag{12}$$

respectively. Therefore, the moments of skew arcsine distribution can be generalized using (9) by taking $P_n(x) = x^n$ and $Q_{n-1}(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0$. Then it becomes

$$\int\frac{x^n}{\sqrt{a^2-x^2}}dx = (a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_1x + a_0)\sqrt{a^2-x^2} + \delta\int\frac{dx}{\sqrt{a^2-x^2}}$$

where

$$a_{n-1} = -\frac{1}{n}$$
$$a_{n-2} = 0$$
$$a_{n-3} = -\frac{(n-1)a^2}{n(n-2)}$$
$$a_{n-4} = 0$$
$$a_{n-5} = -\frac{(n-1)(n-3)a^2a^2}{n(n-2)(n-4)}$$
$$a_{n-6} = 0$$
$$a_{n-7} = -\frac{(n-1)(n-3)(n-5)a^2a^2a^2}{n(n-2)(n-4)(n-6)}$$
$$\vdots$$
$$\delta = -a_1a^2$$

Then,

$$E(X^n) = \begin{cases} \frac{2\lambda}{\pi^2} \int_{-a}^{a} Q_{n-1}(x) \sqrt{\frac{a^2 - x^2}{a^2 - (\lambda x)^2}} \, dx & \text{if } n \text{ is odd} \\ \\ \delta & \text{if } n \text{ is even} \end{cases} \tag{13}$$

In Fig. 5, we illustrate the behaviour of E(X), Var(X), Skewness(X), and Kurtosis(X) for $|\lambda| \leq 1$. As it was expected, both E(X) and Skewness(X) are increasing functions of λ, while Var(X) and Kurtosis(X) are even functions of λ.

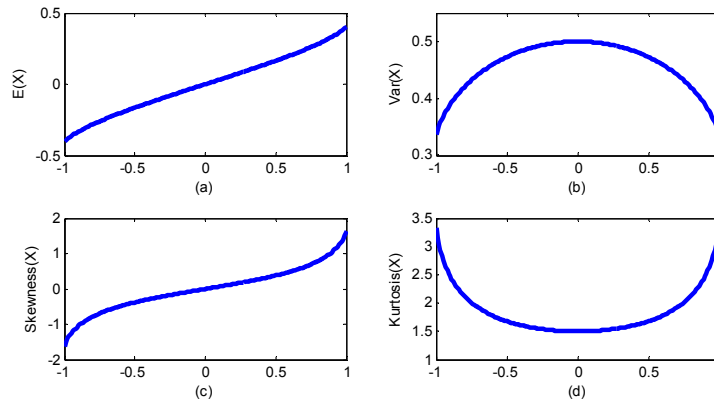**Fig. 5. The behaviour of four measures: (a) E(X), (b) Var(X), (c) Skewness(X), (d) Kurtosis(X) for $|\lambda| \leq 1$**

2.3 Simulation studies for other properties

In this section, we mention the other properties of skew arcsine distribution. The entropy measures the uncertainty in a random variable. As the statistical concept of entropy, Shannon entropy is defined by E[-log(f(X))]. In Fig. 6, we illustrate the behaviour of E[-log(f(X))] for $|\lambda| \leq 1$ with simulation method mentioned in the previous section.
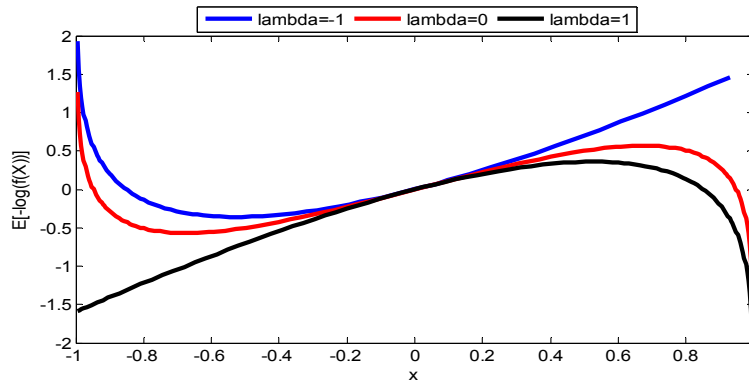

Fig. 6. The Entropy Measures of the Skew Arcsine Distributions

Suppose that (X1,X2,…,Xn) is a random sample from the skew arcsine distribution with pdf in (3). Let X1:n≤X2:n≤…≤Xn:n be the order statistics obtained by arranging the random sample in increasing order of magnitude. The smallest and the largest order statistics are defined by X1:n=min{X1,X2,…,Xn} and Xn:n=max{X1,X2,…,Xn}, respectively. Median of a random sample can be easily defined in terms of order statistics as follows:

Here, we investigate first, second (Median), and third quantiles for the skew arcsine distribution on the basis of a simulation algorithm. Initially, we generate 100000 random numbers from the skew arcsine distribution with acceptance-rejection method as follows:

Acceptance-Rejection Method: Let X a random variable with cdf $F(x)$. If it cannot be found the inverse form of $F(x)$, the alternative methods for random generating are used. The acceptance rejection method based on finding another distribution function with pdf $g(x)$. Then the method proceeds as follows:

- Generate a random variable Y distributed with $G(x)$
- Generate another random variable U distributed uniformly independent from Y
- If $\leq \frac{f(y)}{cg(y)}$, then X=Y, otherwise go back the first step.

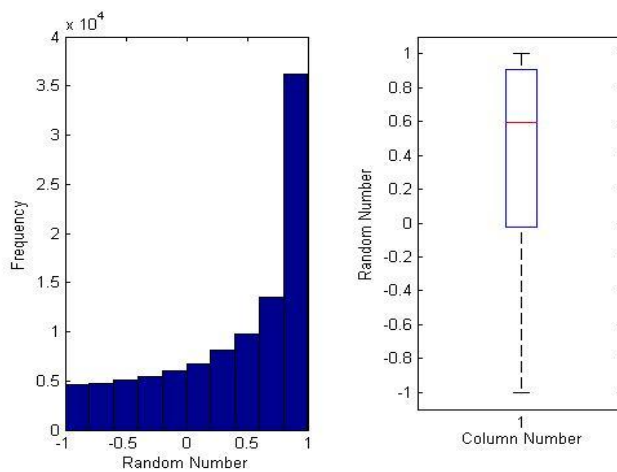We take *g(x)* as the standard arcsine distribution and $c = \frac{\pi}{2}$.

Fig. 7. Histogram and boxplot for generated random numbers

Fig. 7 shows histogram and boxplot for 100000 random numbers generated by this algorithm. This algorithm is repeated 10000 times and the sampling distributions of first, second, and third quartiles are illustrated in Fig. 8.
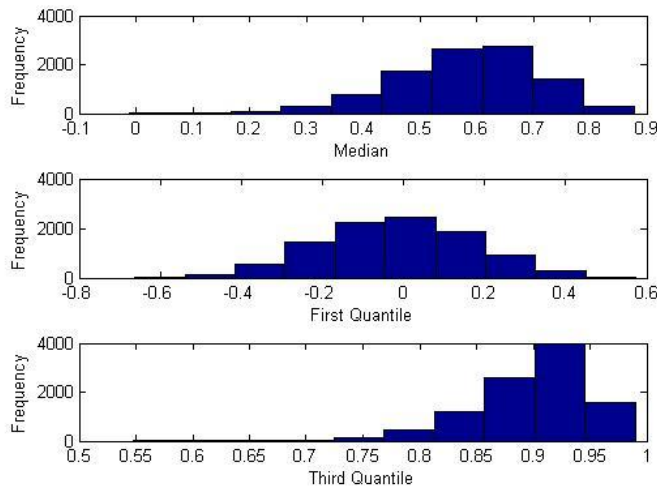

Fig. 8. Histograms of first, second, and third quartiles

2.4 Estimation

In this section we find the estimators of the unknown parameters of the skew arcsine distribution with maximum likelihood (ML) methodology. Consider the pdf of the skew arcsine distribution defined in (3), then the likelihood function can be written as,

$$L(\lambda, a; x) = \prod_{i=1}^{n} f(x_i, a, \lambda) = \prod_{i=1}^{n} \frac{4}{\pi^2} \frac{1}{\sqrt{a^2 - x_i^2}} arcsine\left(\sqrt{\frac{\lambda x_i + a}{2a}}\right) \tag{14}$$

Then the log-likelihood function can be obtained like,

$$lnL = nln4 - 2n\, ln\pi - \frac{1}{2}\sum_{i=1}^{n} ln[a^2 - x_i^2] + \sum_{i=1}^{n} ln\left[arcsine\left(\sqrt{\frac{\lambda x_i + a}{2a}}\right)\right] \tag{15}$$

To obtain ML estimators, the log likelihood function is maximized with respect to unknown parameters, and then we reach the following likelihood equations.

$$\frac{\partial lnL}{\partial a} = \sum_{i=1}^{n}\left[\frac{-\lambda x_i(a^2 - x_i^2)^{-1/2}}{2a. arcsine\sqrt{\frac{\lambda x_i + a}{2a}}} - \frac{a}{(a^2 - x_i^2)}\right] = 0$$

$$\frac{\partial lnL}{\partial \lambda} = \sum_{i=1}^{n}\left[\frac{x_i(a^2 - x_i^2)^{-1/2}}{2. arcsin\sqrt{\frac{\lambda x_i + a}{2a}}}\right] = 0 \tag{16}$$

As can be seen from the equations, ML equations have no explicit solution, therefore iterative methods such as iteratively reweighting algorithms (IRA) or asymptotic methods such as modified maximum likelihood (MML) methods can be used to solve the equations.

The maximum likelihood estimators $(\hat{a}, \hat{\lambda})$ are consistent and $\sqrt{n}(a - \hat{a}, \lambda - \hat{\lambda})$ is asymptotically normal with mean 0 and variance covariance matrix Σ^{-1}, where

$$\Sigma = \begin{bmatrix} E\left(\frac{\partial^2 L}{\partial a^2}\right) & E\left(\frac{\partial^2 L}{\partial a\, \lambda}\right) \\ E\left(\frac{\partial^2 L}{\partial \lambda\, a}\right) & E\left(\frac{\partial^2 L}{\partial \lambda^2}\right) \end{bmatrix} \tag{17}$$

3 Conclusion

In this paper, we propose a new distribution namely skew arcsine distribution using Azzalini's skewness procedure. We derive the pdf of the distribution. The cdf and the other characteristics cannot be solved analytically; therefore, we use Gauss-Kronrod Quadrature method to obtain the certain properties of skew arcsine distribution. The estimation of unknown parameters is obtained at the end of the study. In future works, we are aim to study skew arcsine distribution and its structure in random walk theory and applications where traditional arcsine distribution has been used.

Competing Interests

Author has declared that no competing interests exist.

References

[1]   Norton RM. On properties of the arc sine law. Sankhya Ser. A. 1975;37(2):306-308.

[2]   Billingsley P. Convergence of probability measures. John Wiley & Sons Inc., New York; 1968.

[3]   Erdös P, Frued G. On orthogonal polynomials with regularly distributed zeros. Proc. London Math. Soc; 1974

[4]   Norton RM. Moment properties and the arc sine law. Sankhy¹a Ser. A. 1978;40(2):192-198.

[5]   Fernández C, Osiewalski J, Steel MFJ. Modeling and inference with v-spherical distributions. Journal of the American Statistical Association. 1995;90:1331–1340.

[6]   Arnold BC, Groeneveld RA. Some properties of the arcsine distribution. J. Amer. Statist. Assoc. 1980;75(169):173-175.

[7]   Kemperman JHB, Skibinsky M. On the characterization of an interesting property of the arcsine distribution. Pacific Journal of Mathematics. 1982;103(2):457-465.

[8]   Simon HA. On a class of skew distribution functions. Biometrika. 1955;42:425–440.

[9]   Sastry DVS, Bhati D. A new skew logistic distribution: Properties and applications. Brazilian Journal of Statistics. 2016;30(2):248-271.

[10]   Azzalini A. A class of distributions which includes the normal ones. Scand. J. Stat. 1985;12:171–178.

[11]   Azzalini A. Further results on a class of distributions which includes the normal ones. Statistica. 1986; 46:199–208.

[12]   Gupta AK, Chang FC, Huang WJ. Some skew-symmetric models. Random Oper. Stoch. Equ. 2002;10: 133–140.

[13]   Nadarajah S. The skew logistic distribution. AStA Adv. Stat Anal. 2009;93(2):187-203.

[14]   Dvs S, Deepesh B. A new skew logistic distribution: Properties and applications, Brazilian Journal of Probability and Statistics; 2015, Special Volume.