



Science

TRANSMUTED ARCSINE DISTRIBUTION PROPERTIES AND APPLICATION

Salma Omar Bleed ^{*1}, Arwa Elsunousi Ali Abdelali ²

^{*1} Department of Statistics, Collage of Science, Alasmarya University Zliten-Libya

² Libyan Academic of Postgraduate Studies, Misurata – Libya

Abstract

The distribution of ArcSine will be developed to another new distribution using the Quadratic Rank Transmutation (QRT) method proposed by Shaw and Buckley (2007). The new distribution will be called the Transmuted ArcSine distribution, some of its mathematical characteristics such as variance, expectation, residual function, risk function, moments, moment generating function and characteristic function will be presented. The model parameters will be estimated by the maximum likelihood method. Finally, two real data sets are analyzed to illustrates the usefulness of the TAS distribution.

Keywords: ArcSine Distribution; Quadratic Rank Transmutation; Residual Function; Risk Function; Estimated of Parameters.

Cite This Article: Salma Omar Bleed, and Arwa Elsunousi Ali Abdelali. (2018). “TRANSMUTED ARCSINE DISTRIBUTION PROPERTIES AND APPLICATION.” *International Journal of Research - Granthaalayah*, 6(10), 38-47. <https://doi.org/10.5281/zenodo.1473819>.

1. Introduction

ArcSine distribution is an extensive model of data analysis for multiple situations, which is of great importance in probability applications. It is a symmetric over the period $(-1,1)$. This distribution can be applied to several fields, such as the random variables behavior model, which is limited to periods of specific length in a wide range of fields, in genetics of the statistical description of allele frequencies, in theory of statistical communication as a model of amplitude periodic signal in thermal noise as a specific spectrum, in economic modeling to describe the function density of functions such as allocation of time in project management and control systems, in modeling ratios in mathematics (e.g rates, fractions and ratios), such as the rate of unemployment, the rate of poverty, etc [Arnold and Groeneveld (1980)]. Also, it is used in the genetic section, such as the knowledge of the amount of cholesterol for a part of the total, the percentage of mammography intensity (MD height is the breast cancer sign) and in the environment for knowing a percentage of sand in the soil in a part of the surface covered with plants, [Gauss *et. al.* (2016)].

The ArcSine distribution is considered a continuous probability distribution and is defined by the following probability function (PDF)

$$f(x) = \frac{1}{\pi\sqrt{1-x^2}}, \quad -1 < x < 1$$

with cumulative function (CDF)

$$F(x) = \frac{1}{2\pi}(\pi + 2 \arcsin x), \quad -1 < x < 1$$

and arithmetic mean = median = skewness coefficient = zero, variance = 0.5, kurtosis coefficient = -1.5 and moment generating function $MGF = \sum_{n=0}^{\infty} \frac{t^{2n}(2n-1)!!}{(2n)!!}$

Where $(2n - 1)!! = 1.3 \dots (2n - 1)$ and $(2n)!! = 2.4 \dots (2n)$, [Ahsanullah (2015)].

In the past few years, new families have been attracted by distributions that generate many statisticians to develop new models as well as provide great flexibility in modeling real data. These families are obtained by inserting one or more of the additional form parameters into the basic distribution. For example, In 2013, Aryal presented the transmuted Log-Logistic distribution, Gauss and others (2016) proposed a new continuous model with two parameters called the extended arcsine distribution using the method of exponentiated generalized, Yousof and others (2017) proposed the beta Weibull-G family of distributions, and Salma (2018) presented the distribution of the quadratic formula (L) with the study of distribution characteristics.

In this paper, the QRT method suggested by Shaw and Buckley (2007) was used to identify a new model, which is a generalization of the ArcSine distribution called the transmuted ArcSine distribution and will be denoted by TAS distribution. According to the QRT method, the cumulative distribution function (CDF) was obtained using the following relationship:

$$F(x) = (1 + \lambda)F_1(x) - \lambda F_1^2(x) \rightarrow (1)$$

where $F_1(x)$ is the cumulative distribution function (CDF) of the basic distribution, and the probability function of the new distribution can be obtained as follows:

$$f(x) = (1 + \lambda)f_1(x) - 2\lambda f_1(x)F_1(x) \rightarrow (2)$$

In the rest of this paper, the mathematical structures of the new distribution and some of its properties are presented with an explanation of two applied examples to support the distribution properties. For more information about the QRTM, see [Shaw and Buckley (2007)].

2. Transmuted Arcsine Distribution

The TAS distribution in the interval $(-1,1)$ by using Eq.1 and Eq.2 with the distribution function $F(X)$ is defined as

$$F(x) = \left[\frac{\pi + 2\arcsin(x)}{2\pi} \right] \left[1 + \lambda \left(\frac{\pi - 2\arcsin(x)}{2\pi} \right) \right] \rightarrow (3)$$

where $\lambda > 0$ is the shape transformer parameter. The probability density function $f(x)$ of the new distribution will be as follows:

$$f(x) = \frac{1}{\pi\sqrt{1-x^2}} \left[1 - \frac{2\lambda}{\pi} \arcsin(x) \right] \rightarrow (4)$$

Lemma: $f(x)$ given in Eq.4 is a well-defined probability density function.

Proof: Note that $f(x)$ given in Eq.4 is a non-negative and the integration over the support of the random variable is equal one as follows:

$$\int_{-1}^1 f(x) dx = 1 = \int_{-1}^1 \frac{1}{\pi\sqrt{1-x^2}} \left[1 - \frac{2\lambda}{\pi} \arcsin(x) \right] dx$$

by assuming $u = 2\lambda \arcsin(x)$, $dx = \frac{\sqrt{1-x^2}}{2\lambda} du$ the integration over the support of the random variable is equal one. Hence $f(x)$ given in Eq.4 is a well-defined probability density function.

3. Some Statistical Properties

In this section, some statistical properties of the new distribution are summarized, [Christian (2007)].

Residual Function: The residual or the reliability function (RF) is defined by the relationship $RF = 1 - F(x)$ and for the TAS distribution the residual function (RF_{TAS}) is defined as

$$: RF_{TAS}(x) = \left[\frac{\pi - 2\arcsin(x)}{2\pi} \right] \left[1 - \lambda \left(\frac{\pi + 2\arcsin(x)}{2\pi} \right) \right].$$

Risk Function: The risk function (RiF) is defined by the relationship $RiF_{TAS}(x) = \frac{f(x)}{RF(x)}$ and for the TAS distribution the risk function (RiF_{TAS}) is defined as:

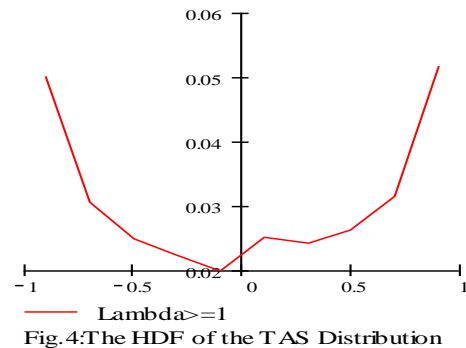
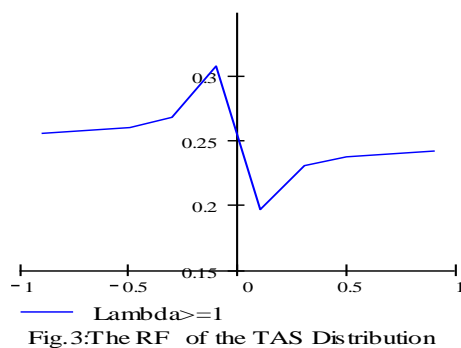
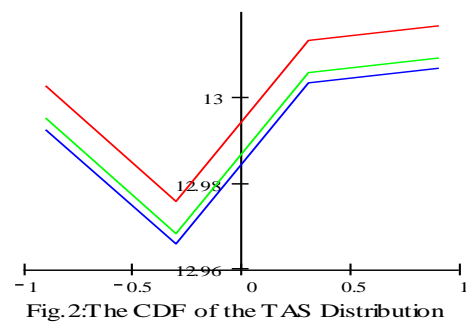
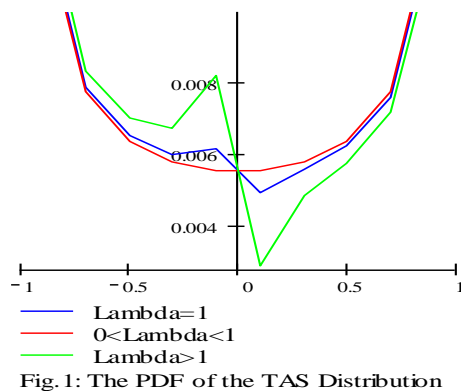
$$RiF_{TAS}(x) = \left[\frac{1 - \frac{2\lambda}{\pi} \arcsin(x)}{\pi\sqrt{1-x^2}} \right] \left\{ \left[\frac{\pi - 2\arcsin(x)}{2\pi} \right] \left[1 - \lambda \left(\frac{\pi + 2\arcsin(x)}{2\pi} \right) \right] \right\}^{-1}.$$

Cumulative Risk Function: The cumulative risk function (CRiF) is defined by the relationship $CRiF_{TAS}(x) = -\ln F(x)$ and for the TAS distribution the cumulative hazard function ($CRiF_{TAS}$) is defined as:

$$CRiF_{TAD}(x) = -\ln \left\{ \left[\frac{\pi + 2\arcsin(x)}{2\pi} \right] \left[1 + \lambda \left(\frac{\pi - 2\arcsin(x)}{2\pi} \right) \right] \right\}$$

Fig. 1 to 4 illustrate plot the PDF, CDF, RF, and HDF of the TAS distribution under different values of the shape parameter (λ). It is noted that the PDF of the new distribution is U-shaped, which is a symmetric function around zero when ($0 < \lambda < 1$).

Also, the plots of the CDF, RF and HDF indicates increasing cumulative distribution function, decreasing instantaneous residual function and then increasing and the HDF takes the form of the letter U, which is an incremental function on the right and the left side.



Quantile Function and Estimate Parameter: By inverting Eq.3, the quantile function of the random variable X is defined as follows:

$$x(U) = \sin \left(\frac{\pi}{4} [2(2U - 1) - \lambda] + \frac{\lambda}{\pi} \arcsin^2(x) \right) \rightarrow (5)$$

where U is a random variable followed the uniform distribution (0.1). Eq.5 does not have a closed form and to generate random variables from the TAS distribution, Matacad 15 program can be used. By substituting $U = 0.5$ in Eq.5, the median of the TAS distribution is obtained.

4. Useful Integration Expansions

In this section, useful integration expansions for the raw moments are provided, [Gradshteyn and Ryzhik (2007)]. The generalized infinite integration expansion is given by

$$\int \sin^m x \, dx = \begin{cases} \sum_{k=0}^{\frac{m-1}{2}} (-1)^{\frac{1}{2}} M_k [\cos((m-2k)x)] & \text{for } m = 1, 3, \dots \\ \pi_m x + \sum_{k=0}^{\frac{m-1}{2}} M_k [\sin((m-2k)x)] & \text{for } m = 2, 4, \dots \end{cases} \rightarrow (6)$$

And

$$\int x \sin^m x \, dx = \begin{cases} \sum_{k=0}^{\frac{m-1}{2}} (-1)^{\frac{1}{2}} M_k \left[\frac{\sin((m-2k)x)}{(m-2k)} - x \cos((m-2k)x) \right] & \text{for } m = 1, 3, \dots \\ \frac{\pi_m x^2}{2} + \sum_{k=0}^{\frac{m-1}{2}} M_k \left[\frac{\cos((m-2k)x)}{(m-2k)} + x \sin((m-2k)x) \right] & \text{for } m = 2, 4, \dots \end{cases} \rightarrow (7)$$

The useful generalized finite integration expansion in the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$ by using Eq.6 and Eq.7 is given by

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^m x \, dx = \begin{cases} \pi_m \pi + 2 \sum_{k=0}^{\frac{m-1}{2}} M_k \left[\sin \left((m-2k) \left(\frac{\pi}{2} \right) \right) \right] & \text{for } m = 2, 4, \dots \\ 0 & \text{for } m = 1, 3, \dots \end{cases} \rightarrow (8)$$

and

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} x \sin^m x \, dx = \begin{cases} \sum_{k=0}^{\frac{m-1}{2}} M_k \left[\frac{2 \sin \left((m-2k) \left(\frac{\pi}{2} \right) \right)}{(m-2k)} - \pi \cos \left((m-2k) \left(\frac{\pi}{2} \right) \right) \right] & \text{for } m = 1, 3, \dots \\ 0 & \text{for } m = 2, 4, \dots \end{cases} \rightarrow (9)$$

5. The Raw Moments

Theorem 1: If X is a random variable followed the TAS distribution with PDF.(2), then the m^{th} raw moments is defined as follows

$$\mu'_m = \begin{cases} \frac{2\lambda}{\pi^2} \sum_{k=0}^{\frac{m-1}{2}} (-1)^{\frac{1}{2}} M_k \left[\pi \cos(m-2k) \left(\frac{\pi}{2} \right) - \frac{2}{(m-2k)} \sin(m-2k) \left(\frac{\pi}{2} \right) \right] & \text{for } m = 1, 3, \dots \\ \pi_m + \frac{2}{\pi} \sum_{k=0}^{\frac{m-1}{2}} M_k \left[\sin(m-2k) \left(\frac{\pi}{2} \right) \right] & \text{for } m = 2, 4, \dots \end{cases}$$

Such that

$$\pi_m = \frac{\binom{m}{m/2}}{2^m}, \quad M_k = \frac{(-1)^{\frac{m}{2}+k} \binom{m}{k}}{2^{m-1} (m-2k)}$$

Proof: Since $\mu'_m = \int_{-1}^1 x^m f(x) dx = \int_{-1}^1 \frac{x^m}{\pi\sqrt{1-x^2}} \left[1 - \frac{2\lambda(\arcsin x)}{\pi}\right] dx$

Put $y = \arcsin x \rightarrow x = \sin y \rightarrow dx = \cos y dy$
 Then, the integration can be rewritten as follows:

$$= \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{(\sin y)^m}{\sqrt{1 - (\sin y)^2}} \cos y dy - \frac{2\lambda}{\pi^2} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \frac{y (\sin y)^m}{\sqrt{1 - (\sin y)^2}} \cos y dy \quad \rightarrow (10)$$

Substituting (8) and (9) in (10), the integration becomes

$$\mu'_m = \begin{cases} \frac{2\lambda}{\pi^2} \sum_{k=0}^{\frac{m-1}{2}} (-1)^{\frac{1}{2}} M_k \left[\pi \cos(m-2k) \left(\frac{\pi}{2}\right) - \frac{2}{(m-2k)} \sin(m-2k) \left(\frac{\pi}{2}\right) \right] & \text{for } m = 1, 3, \dots \\ \pi_m + \frac{2}{\pi} \sum_{k=0}^{\frac{m-1}{2}} M_k \left[\sin(m-2k) \left(\frac{\pi}{2}\right) \right] & \text{for } m = 2, 4, \dots \end{cases} \quad \rightarrow (11)$$

Where $\pi_m = \frac{\binom{m}{m/2}}{2^m}, \quad M_k = \frac{(-1)^{\frac{m}{2}+k} \binom{m}{k}}{2^{m-1} (m-2k)}$.

Therefore, $\mu'_1 = \frac{4\lambda}{\pi^2}, \mu'_2 = \frac{1}{2} - \frac{1}{\pi}, \mu'_3 = \frac{5\lambda}{3\pi^2}, \mu'_4 = \frac{3}{8} - \frac{1}{2\pi}$.

Note that, when $\lambda = 0$ then

$$\mu'_m = \begin{cases} \pi_m + \frac{2}{\pi} \sum_{k=0}^{\frac{m-1}{2}} M_k \left[\sin(m-2k) \left(\frac{\pi}{2}\right) \right] & \text{for } m = 2, 4, \dots \\ 0 & \text{for } m = 1, 3, \dots \end{cases} \quad \rightarrow (12)$$

Thus, μ'_m does not exist for every m of the basic distribution, where the odd moments are zeros and the even moments are defined by Eq.12. But for the TAS distribution, μ'_m exist for every m, where the odd moments and the even moments are defined by Eq.11, and this is one of the benefits of the TAS distribution.

6. The Moment Generating Function

Theorem 2: If X is a r.v followed the TAS distribution with PDF.(2), then the moment generating function (MGF) is defined as follows: $MGF_x(t) = \sum_{m=0}^{\infty} \frac{t^m}{m!} \mu'_m$

proof: By definition $MGF_x(t) = E(e^{tx}) = \int \frac{e^{tx}}{\pi\sqrt{(1-x^2)}} dx - \frac{2\lambda}{\pi^2} \int \frac{e^{tx} \arcsin x}{\sqrt{(1-x^2)}} dx$

so that

$$MGF_x(t) = \sum_{m=0}^{\infty} \frac{t^m}{m!} \left\{ \int \frac{x^m}{\pi\sqrt{(1-x^2)}} dx - \frac{2\lambda}{\pi^2} \int \frac{x^m \arcsin x}{\sqrt{(1-x^2)}} dx \right\} = \sum_{m=0}^{\infty} \frac{t^m}{m!} \mu'_m \rightarrow (13)$$

By substituting $t = it$ in Eq.13, the characteristic function of the TAS distribution is obtained, i.e., $QF_x(t) = MGF_x(it) = \sum_{m=0}^{\infty} \frac{(it)^m}{m!} \mu'_m$.

7. Estimation of Parameter

The maximum likelihood method (MLM) is used to estimate the TAS unknown parameter (λ). The likelihood function for the TAS distribution is given by:

$$L = \frac{1}{\pi^n} \prod_{i=1}^n \left\{ (1+x_i^2)^{-\frac{1}{2}} \left(1 - \frac{2\lambda}{\pi} \arcsin(x_i) \right) \right\}.$$

The log-likelihood function is obtained as

$$\ell = \log l = -n \log \pi - \frac{1}{2} \sum_{i=1}^n \log(1+x_i^2) + \sum_{i=1}^n \log \left(1 - \frac{2\lambda}{\pi} \arcsin(x_i) \right)$$

Hence, the log-likelihood equation for estimating (λ) is given by

$$\frac{d\ell}{d\lambda} = 0 \rightarrow \frac{-2}{\pi} \sum_{i=1}^n \frac{\arcsin(x_i)}{1 - \frac{2\hat{\lambda}}{\pi} \arcsin(x_i)} = 0 \rightarrow (14)$$

Eq.14 is an implicit equation in (λ), hence it can't be solved analytically for (λ), then the Matabad 15 program will be used to find ($\hat{\lambda}$). The expected information matrix is given by

$$I(\lambda) = \frac{4}{\pi^2} \sum_{i=1}^n \frac{\arcsin(x_i)^2}{\left(1 - \frac{2\hat{\lambda}}{\pi} \arcsin(x_i) \right)^2} \rightarrow (15)$$

Solving the inverse for the observed information matrix (15), yields the asymptotic variance of the ML estimator ($\hat{\lambda}$). By using Eq.15 approximate $100(1 - \alpha)\%$ asymptotic confidence intervals for (λ) is: $\hat{\lambda} \mp Z_{\frac{\alpha}{2}} \sqrt{v(\hat{\lambda})}$ where $Z_{\frac{\alpha}{2}}$ is the upper α^{th} percentile of the standard normal distribution.

8. Application

This section illustrates the usefulness of the TAS distribution with real data.

Data I: The 1st data presented by Linhart and Zucchini (1986) is used. The data set (0.46, 0.46, 0.46, 0.50, 0.58, 0.58, 0.58, 0.67, 0.67, 0.83, 0.87) is the failure times (Days) of the air conditioning system of an airplane.

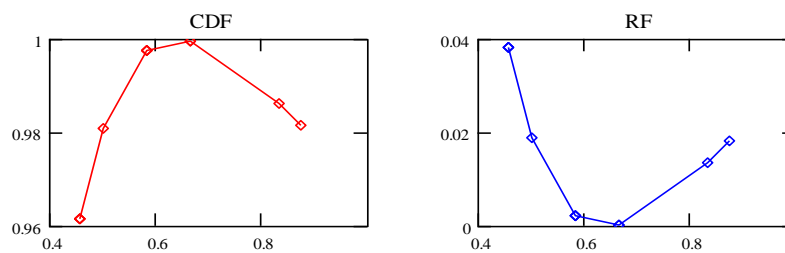


Figure 5: CDF and RF of TAS-distribution

Table 1: Estimated Parameter, Raw moments and Some Properties of TAS-distribution

λ_0	$\hat{\lambda}$	MSE	Variance	LLC.I	ULC.I	K-S Test
1	1.002	4.09510E-6	4.94310E-3	0.886	1.118	0.467
1.25	1.254	1.28710E-5	4.99910E-3	1.137	1.37	C($\alpha=0.01$) = 0.46770
λ_0	Mean	Median	Mode	Variance	Kurtosis	Skewness
1	5.09910E-3	0.06	6.61110E-3	0.502	1.497	-8.27310E-3
1.25	6.37910E-3	0.095	7.64610E-3	0.502	1.497	-0.01
λ_0	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
1	5.09910E-3	0.502	4.73810E-3	0.378	4.43110E-3	0.315
1.25	6.37910E-3	0.502	5.92810E-3	0.378	5.54310E-3	0.3151

Data II: The 2nd data (0.05, 0.10, 0.34, 0.60, 0.79, 0.06, 0.12, 0.35, 0.60, 0.79) was representing the stress-rupture life of Kevlar 49/epoxy strands (unit: hours), [Insuk et. al. (2015)].

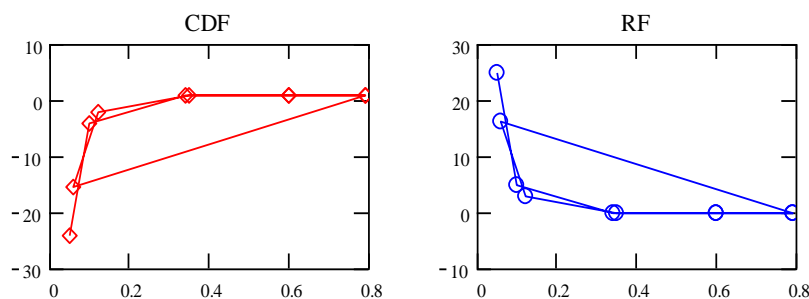


Figure 6: CDF and RF of TAS-distribution

Table 2: Estimated Parameter, Raw moments and Some Properties of TAS-distribution

λ_0	$\hat{\lambda}$	MSE	Variance	LLC.I	ULC.I	K-S Test
0.36	0.369	7.75210E-5	0.124	-0.211	0.949	0.487
0.761	0.764	6.86510E-6	0.147	0.134	1.393	$C(\alpha=0.01)= 0.48893$
λ_0	Mean	Median	Mode	Variance	Kurtosis	Skewness
0.36	1.87710E-3	0.631	0.035	0.502	1.497	-3.04510E-3
0.761	3.88610E-3	0.041	0.069	0.502	1.497	-6.30510E-3
λ_0	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6
0.36	1.87710E-3	0.502	1.74410E-3	0.378	1.63110E-3	0.315
0.761	3.88610E-3	0.502	3.61110E-3	0.378	3.37710E-3	0.315

Statistical analysis: The plots of cumulative distribution function (CDF) and residual function (RF) of TAS-distribution having pdf (4) for both data (data I, data II) indicates increasing cumulative distribution function and then decreasing residual function, As seen from Figure 6, 7. The estimated parameters for the TAS distribution with their corresponding 95% C.I, the values of the mean square error (MSE) and Kolmogorov-Smirnov (K-S) test of the TAS distribution are displayed in Table 1, 2. The results indicates that the data provides a better fit to the TAS distribution and estimate the true parameter λ well with relatively small MSE with short confidence intervals. In general, the results indicate that the proposed model and the asymptotic approximation work well under the situation, therefore the TAS distribution provides a better fit, and then can be chosen as a suitable model for lifetime data, as seen from these numerical results in Table 1, 2.

9. Conclusion

A new distribution named the TAS distribution, which is an extension of the AS distribution is proposed. In this model the new parameter λ provides more flexibility in modeling reliability data and provides better results than the AS distribution. The variance, expectation, residual function, risk function, moments, moment generating function and the characteristic function are derived and all the raw moments exist. The model parameter is estimated by maximum likelihood method with two real data sets are analyzed to illustrates the usefulness of the TAS distribution. Moreover, the asymptotic variance-covariance matrix and the asymptotic confidence intervals are determined.

References

- [1] Ahsanullah, M. (2015). "Some characterizations of Arcsine distribution". Moroccan J. Pure and Appl. Anal. (MJPA). Volume 1(2), 2015, Pages 70–75.
- [2] Arnold B. C., and Groeneveld R. A. (1980). "Some Properties of the arcsine Distribution". J. Amer. Statist. Assoc., 75 (369): 173-175.
- [3] Aryal, G. R. (2013). Transmuted Log-Logistic Distribution. J. Stat. Appl. Pro. 2, No. 1, pp. 11-20.
- [4] Christian Walck. (2007). "Hand-book on Statistical Distributions for Experimentalists". Particle Physics Group Fysikum University of Stockholm.
- [5] Gauss M., Artur J., and Anak C. (2016). "Extended Arcsine distribution to Properties and Applications". Studia Scientiarum Mathematicarum Hungarica. 53 (4), 440-466.
- [6] Gradshteyn, I. S., and Ryzhik, I. M. (2007). "Table of Integrals, Series, and Products". Seventh Edition, Translated from Russian by Scripta Technica, Inc.
- [7] Insuk, T., Bodhisuwan, W., and Jaroengeratikun, U. (2015). "A new mixed beta distribution and structural properties with applications". Songklanakarin J. Sci. Technol. 37 (1), 97-108.

- [8] *Linhart, H. and Zucchini, W. (1986). "Model selection". New York: Wiley.*
- [9] *Salma Omar Bleed. (2018). "L-Quadratic Distribution". International Journal of Computer Applications, V. 179, No.13, Pp:6-11.*
- [10] *Shaw, W. T., and Buckley, I. R. C. (2007). "The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtosis normal distribution from a rank transmutation map". arXiv preprint arXiv: 0901.0434.*
- [11] *Yousof, H. M., Rasekhi, M., Afify, A. Z., Ghosh, I., Alizadeh, M., and Hamedani, G. G. (2017). "The beta Weibull-G family of distributions: Theory, characterizations and applications". Pakistan Journal of Statistics, 33(2), 95-116.*

*Corresponding author.

E-mail address: SalmaBleed@ yahoo.com