

Transmuted Weibull distribution and its applications

Ivana Pobočiková¹, Zuzana Sedliačková¹, Mária Michalková^{1,*}

¹Department of Applied Mathematics, Faculty of Mechanical Engineering, University of Žilina, Univerzitná 1, 010 26 Žilina, Slovakia

Abstract. In this paper we study new distribution called transmuted Weibull distribution. Some properties of this distribution are described. The usefulness of the distribution for modelling data is illustrated using real data set.

Keywords: transmuted Weibull distribution, properties, maximum likelihood estimation, application; Akaike's information criterion, Bayesian information criterion, coefficient of determination, root mean square error

1 Introduction

The Weibull distribution was originally introduced by Swedish physicist Waloddi Weibull [1]. He applied it on modelling the distribution of the yield strength of materials. The Weibull distribution is popular and widely used, with many applications in areas such as engineering (reliability, failure analysis, lifetime analysis ...), material science, quality control, physics, medicine, meteorology, hydrology and others [2].

However, there are cases when standard Weibull distribution fails to model data suitably enough. This is where it is necessary to apply generalized distribution because of its flexibility and better fit of the data than standard Weibull distribution. The importance of such generalization has been proved in recent years on various problems and many standard distributions have been generalized. Some recent generalizations which are based on the Weibull distribution, are discussed in [3].

An interesting idea of generalization where the distribution is derived using the quadratic rank transmutation map, was studied first by Shaw and Buckley in [4]. Recently, many transmuted distributions have been proposed in literature. See, for example, transmuted Gumbel distribution [5], transmuted Rayleigh distribution [6], transmuted modified Weibull distribution [7], transmuted inverse Weibull distribution [8], transmuted Pareto distribution [9] and others.

Aryal and Tsokos [10] used the quadratic rank transmutation map to develop new generalization of the Weibull distribution, called the transmuted Weibull distribution. They studied mathematical properties of this distribution and the maximum likelihood estimates of the parameters.

* Corresponding author: maria.michalkova@fstroj.uniza.sk

Reviewers: *Vladimír Dekýš, Urszula Siedlecka*

In this paper, we focus on demonstrating the usefulness of the transmuted Weibull distribution for modelling lifetime, illustrated on three real data sets. All three data sets are modelled with the transmuted Weibull distribution, the 2-parameter and 3-parameter Weibull distribution. For comparison of the distributions the following criteria are used: the Akaike's information criterion, the corrected Akaike's information criterion, the Bayesian information criterion, the coefficient of determination and the root mean square error.

The parameters of the probability distributions are estimated using the maximum likelihood method. The calculations are performed using statistical software STATISTICA and software MATLAB.

The rest of the paper is organized as follows: In Section 2 we introduce the transmuted Weibull distribution. In Section 3 we define basic statistical properties including moments, quantiles, reliability and hazard rate functions. The maximum likelihood estimates of the transmuted Weibull parameters are presented in Section 4. Finally, in Section 5 we provide the real data applications of the transmuted Weibull distribution.

2 Transmuted Weibull distribution

A random variable X is said to have the 3-parameter Weibull distribution (WD) $W(a, b, c)$ with parameters $a > 0$, $b > 0$, $c \geq 0$ if its cumulative distribution function (CDF) for $x \geq c$ is given by

$$G(x) = 1 - \exp\left(-\left(\frac{x-c}{b}\right)^a\right) \quad (1)$$

and probability density function (PDF) is given by

$$g(x) = \frac{a}{b^a} (x-c)^{a-1} \exp\left(-\left(\frac{x-c}{b}\right)^a\right) \quad (2)$$

where a is the dimensionless shape parameter, b is the scale parameter and c is the location parameter.

Setting $c=0$, (1) and (2) transform into CDF and PDF, respectively, of the 2-parameter Weibull distribution $W(a, b)$.

The transmuted distribution can be obtained by adding real number λ , $|\lambda| \leq 1$, into CDF that provides more flexibility in the form of new distribution [7]. λ denotes the transmuted parameter.

A random variable X is said to have the transmuted distribution if its CDF $F(x)$ and PDF $f(x)$ are given by

$$F(x) = (1 + \lambda)G(x) - \lambda G(x)^2 \quad (3)$$

$$f(x) = g(x)[(1 + \lambda) - 2\lambda G(x)] \quad (4)$$

where $|\lambda| \leq 1$, $G(x)$ and $g(x)$ are CDF and PDF of the base distribution, respectively. Observe that for $\lambda = 0$ we have the base distribution of the random variable X .

Let us consider the 2-parameter Weibull distribution. Adding (1) into (3) we have CDF of the transmuted Weibull distribution (TWD) $W(a, b, \lambda)$ with parameters $a > 0$, $b > 0$ and $|\lambda| \leq 1$

$$F(x) = \left[1 - \exp\left(-\left(\frac{x}{b}\right)^a\right)\right] \left[1 + \lambda \exp\left(-\left(\frac{x}{b}\right)^a\right)\right] \quad (5)$$

Similarly, combining (2), (3) and (4) we have PDF of the TWD

$$f(x) = \frac{a}{b^a} x^{a-1} \exp\left(-\left(\frac{x}{b}\right)^a\right) \left[1 - \lambda + 2\lambda \exp\left(-\left(\frac{x}{b}\right)^a\right)\right] \quad (6)$$

3 Statistical properties

In this section we define statistical properties including moments, quantiles, reliability and hazard rate functions. The origin and other aspects of this distribution can be found in [10].

3.1 Moments and quantiles

The r -th moment, $r = 1, 2, \dots$, of the TWD is defined as

$$E(X^r) = b^r \Gamma\left(\frac{r}{a} + 1\right) [1 - \lambda + \lambda 2^{-r/a}] \quad (7)$$

Consequently, the mean $E(X)$ and the variance $Var(X)$ are given by

$$E(X) = b \Gamma\left(\frac{1}{a} + 1\right) [1 - \lambda + \lambda 2^{-1/a}] \quad (8)$$

$$Var(X) = b^2 \left\{ \Gamma\left(\frac{2}{a} + 1\right) [1 - \lambda + \lambda 2^{-2/a}] - \Gamma^2\left(\frac{1}{a} + 1\right) [1 - \lambda + \lambda 2^{-1/a}]^2 \right\} \quad (9)$$

respectively, where $\Gamma(p)$ is gamma function defined by $\Gamma(p) = \int_0^\infty x^{p-1} e^{-x} dx$, $p > 0$.

The α -th quantile, $\alpha \in (0, 1)$, of the TWD is given by

$$x_\alpha = b \left[-\ln \left(1 - \left(\frac{1 + \lambda - \sqrt{(1 + \lambda)^2 - 4\lambda\alpha}}{2\lambda} \right) \right) \right]^{1/a} \quad (10)$$

3.2 Reliability analysis

The TWD can be useful in characterization of the lifetime data analysis. The reliability function if the TWD is given by

$$R(x) = \exp\left(-\left(\frac{x}{b}\right)^a\right) \left[1 - \lambda + \lambda \exp\left(-\left(\frac{x}{b}\right)^a\right)\right] \quad (11)$$

This function can be interpreted as probability that an item does not fail prior to some time x .

Another important characteristics is the hazard rate function that can be interpreted as conditional probability of failure when the item has already survived to time x .

The hazard rate function of the TWD is given by

$$H(x) = \frac{a}{b^a} x^{a-1} \left[\frac{1 - \lambda + 2\lambda \exp\left(-\left(\frac{x}{b}\right)^a\right)}{1 - \lambda + \lambda \exp\left(-\left(\frac{x}{b}\right)^a\right)} \right] \quad (12)$$

3.3 Random number generation

The random numbers from the TWD can be generated using the inverse function of CDF (5)

$$X = b \left[-\ln \left(1 - \left(\frac{1 + \lambda - \sqrt{(1 + \lambda)^2 - 4\lambda U}}{2\lambda} \right) \right) \right]^{1/a} \quad (13)$$

where U has uniform distribution over the interval $(0,1)$.

4 Maximum likelihood estimates

As a method for estimation of parameters of the TWD we choose the maximum likelihood method (MLM).

Let X_1, X_2, \dots, X_n be a random sample of size n from the TWD with PDF (6) and let x_1, x_2, \dots, x_n be a realization of the random sample. The likelihood function $L = L(x_1, x_2, \dots, x_n; a, b, \lambda)$ of this random sample is given by

$$L = \left(\frac{a}{b^a} \right)^n \exp\left(-\frac{1}{b^a} \sum_{i=1}^n x_i^a\right) \prod_{i=1}^n x_i^{a-1} \left[1 - \lambda + 2\lambda \exp\left(-\left(\frac{x_i}{b}\right)^a\right) \right] \quad (14)$$

and the log-likelihood function $\ln L = \ln L(x_1, x_2, \dots, x_n; a, b, \lambda)$ of this random sample is given by

$$\ln L = n \ln \frac{a}{b^a} - \frac{1}{b^a} \sum_{i=1}^n x_i^a + (a-1) \sum_{i=1}^n \ln x_i + \sum_{i=1}^n \ln \left[1 - \lambda + 2\lambda \exp\left(-\left(\frac{x_i}{b}\right)^a\right) \right] \quad (15)$$

Differentiating (15) with respect to a , b and λ , respectively, and equating each derivative to zero we obtain the equations

$$\frac{n}{a} + \sum_{i=1}^n \ln\left(\frac{x_i}{b}\right) \left[1 - \left(\frac{x_i}{b}\right)^a \right] - 2\lambda \sum_{i=1}^n \frac{\left(\frac{x_i}{b}\right)^a \ln\left(\frac{x_i}{b}\right) \exp\left(-\left(\frac{x_i}{b}\right)^a\right)}{1 - \lambda + 2\lambda \exp\left(-\left(\frac{x_i}{b}\right)^a\right)} = 0 \quad (16)$$

$$\frac{a}{b} \sum_{i=1}^n \left[\left(\frac{x_i}{b}\right)^a - 1 \right] - 2\lambda \sum_{i=1}^n \frac{\left(\frac{x_i}{b}\right)^a \exp\left(-\left(\frac{x_i}{b}\right)^a\right)}{1 - \lambda + 2\lambda \exp\left(-\left(\frac{x_i}{b}\right)^a\right)} = 0 \quad (17)$$

$$\sum_{i=1}^n \frac{2 \exp\left(-\left(\frac{x_i}{b}\right)^a\right) - 1}{1 - \lambda + 2\lambda \exp\left(-\left(\frac{x_i}{b}\right)^a\right)} = 0 \quad (18)$$

The MLM estimates \hat{a} , \hat{b} , $\hat{\lambda}$ of the parameters a , b , λ are obtained iteratively from the nonlinear system of equations (16) – (18).

5 Application

In this section we present the analysis of the real data sets using the TWD and compare it with the WD. In order to compare the distributions we consider the criteria:

- the Akaike's information criterion (*AIC*) – calculated using [11]

$$AIC = -2 \ln L(\hat{\theta}) + 2m \quad (19)$$

where $L(\hat{\theta}) = L(x_1, x_2, \dots, x_n; \hat{\theta})$ is the maximized value of the likelihood function for the estimated model, $\hat{\theta}$ is the MLM estimate of the parameter θ , m is number of parameters to be estimated, n is number of observed data;

- the corrected Akaike's information criterion (*AICC*) – calculated using [12]

$$AICC = AIC + \frac{2m(m+1)}{n-m+1} \quad (20)$$

- the Bayesian information criterion (*BIC*) – calculated using [13]

$$BIC = -2 \ln L(\hat{\theta}) + m \ln(n); \quad (21)$$

- the coefficient of determination (R^2) – calculated using

$$R^2 = \frac{\sum_{i=1}^n (\hat{F}(x_i) - \bar{F})^2}{\sum_{i=1}^n (\hat{F}(x_i) - \bar{F})^2 + \sum_{i=1}^n (F_n(x_i) - \hat{F}(x_i))^2}, \quad (22)$$

where $\hat{F}(x)$ is estimated cumulative distribution function and

$$\bar{F} = \frac{1}{n} \sum_{i=1}^n \hat{F}(x_i) \quad (23)$$

Function $F_n(x)$ is the empirical distribution function, defined as follows

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n I(x_{(i)} \leq x) \quad (24)$$

where $I(x_{(i)} \leq x) = 1$ if $x_{(i)} \leq x$ and 0 otherwise, $x_{(1)}, x_{(2)}, \dots, x_{(n)}$ are observations in ascending order, i.e. $x_{(1)} \leq x_{(2)} \leq \dots \leq x_{(n)}$;

- the root mean square error ($RMSE$) – calculated using

$$RMSE = \left[\frac{1}{n} \sum_{i=1}^n (F_n(x_i) - \hat{F}(x_i))^2 \right]^{1/2} \quad (25)$$

The distribution better fitting the data corresponds to smaller values of AIC , BIC , $AICC$ and $RMSE$ and larger value of R^2 .

Data set 1

The first data set represents lifetimes of Kevlar 49/epoxy strands subjected to constant sustained pressure at 90% stress level until the strand failure. For previous studies with the data see [14 – 17]. The data are as follows:

0.0251, 0.0886, 0.0891, 0.2501, 0.3113, 0.3451, 0.4763, 0.5650, 0.5671, 0.6566, 0.6748, 0.6751, 0.6753, 0.7696, 0.8375, 0.8391, 0.8425, 0.8645, 0.8851, 0.9113, 0.9120, 0.9836, 1.0483, 1.0596, 1.0773, 1.1733, 1.2570, 1.2766, 1.2985, 1.3211, 1.3503, 1.3551, 1.4595, 1.4880, 1.5728, 1.5733, 1.7083, 1.7263, 1.7460, 1.7630, 1.7746, 1.8275, 1.8375, 1.8503, 1.8808, 1.8878, 1.8881, 1.9316, 1.9558, 2.0048, 2.0408, 2.0903, 2.1093, 2.1330, 2.2100, 2.2460, 2.2878, 2.3203, 2.3470, 2.3513, 2.4951, 2.5260, 2.9911, 3.0256, 3.2678, 3.4045, 3.4846, 3.7433, 3.7455, 3.9143, 4.8073, 5.4005, 5.4435, 5.5295, 6.5541, 9.0960.

Table 1 gives the descriptive statistics for the data set. We can see that the empirical distribution is skewed to the right. Table 2 presents the MLM estimates of the parameters together with the log-likelihood $\ln L$, AIC , $AICC$ and BIC values. Results are rounded to four decimal places.

Table 1. Descriptive statistics for data set 1

Min	Max	Mean	Variance	Lower Quantile	Median	Upper Quantile	Skewness	Kurtosis
0.0251	9.0960	1.9592	2.4774	0.8982	1.7362	2.3041	2.0196	5.6004

Table 2. MLM estimates of the parameters and model selection criteria for data set 1

Distribution	Parameter estimates	Statistical test					
		$\ln L$	AIC	$AICC$	BIC	R^2	$RMSE$
$W(a, b)$	$\hat{a} = 1.3256$ $\hat{b} = 2.1328$	-122.5247	249.0494	249.2094	253.7108	0.9753	0.0423
$W(a, b, c)$	$\hat{a} = 1.3169$ $\hat{b} = 2.1228$ $\hat{c} = 0.0058$	-122.5141	251.0282	251.3525	258.0204	0.9750	0.0425
$W(a, b, \lambda)$	$\hat{a} = 1.0509$ $\hat{b} = 1.4419$ $\hat{\lambda} = -0.7955$	-121.4300	248.8600	249.1843	255.8522	0.9838	0.0351

In Table 2 we see that values of AIC , $AICC$ and $RMSE$ are smallest ones and value of $\ln L$ and R^2 are highest ones for TWD, compared to values for 2-parameter and 3-parameter WDs. Hence, we can conclude that the TWD provides better fit to the data than the other two distributions.

Table 5 gives the descriptive statistics for data set 3. Table 6 presents the MLM estimates of the parameters and the comparison of TWD with the 2-parameter and 3-parameter WDs.

Table 5. Descriptive statistics for data set 3

Min	Max	Mean	Variance	Lower Quantile	Median	Upper Quantile	Skewness	Kurtosis
12	90	43.6516	146.0597	36	42	50	0.6915	2.2738

Table 6. MLM estimates and model selection criteria for data set 3

Distribution	Parameter estimates	Statistical test					
		$\ln L$	AIC	$AICC$	BIC	R^2	$RMSE$
$W(a, b)$	$\hat{a} = 3.6871$ $\hat{b} = 48.1100$	-610.2967	1224.5934	1224.6713	1230.6802	0.9640	0.0491
$W(a, b, c)$	$\hat{a} = 2.9515$ $\hat{b} = 38.3366$ $\hat{c} = 9.2768$	-607.4815	1220.9629	1221.1189	1230.0932	0.9656	0.0489
$W(a, b, \lambda)$	$\hat{a} = 4.0312$ $\hat{b} = 55.0570$ $\hat{\lambda} = 0.8380$	-605.5600	1217.1200	1217.2768	1226.2502	0.9733	0.0438

Values in Table 6 indicate that the TWD provides better fit to data than the 2-parameter and 3-parameter WDs.

Figures 1, 2 and 3 illustrate the histograms and the fitted PDFs of the TWD, 2-parameter WD and 3-parameter WD for data set 1, 2 and 3, respectively.

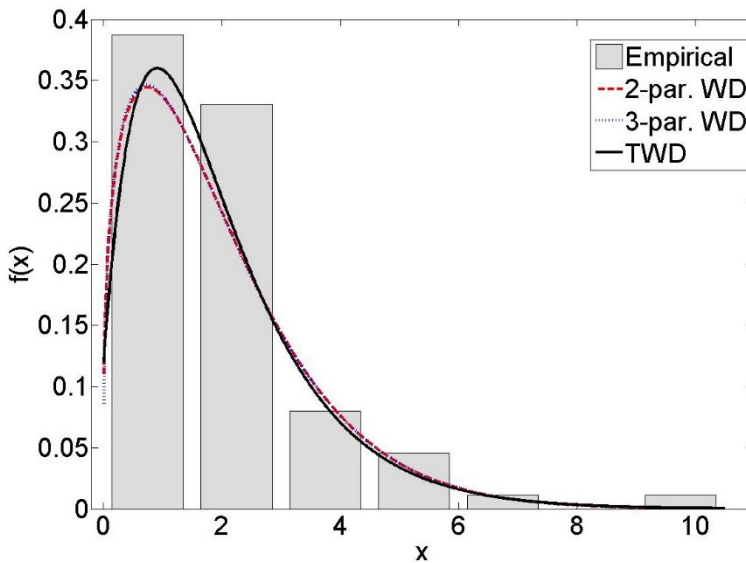


Fig. 1. Histogram of data set 1 and the fitted PDFs

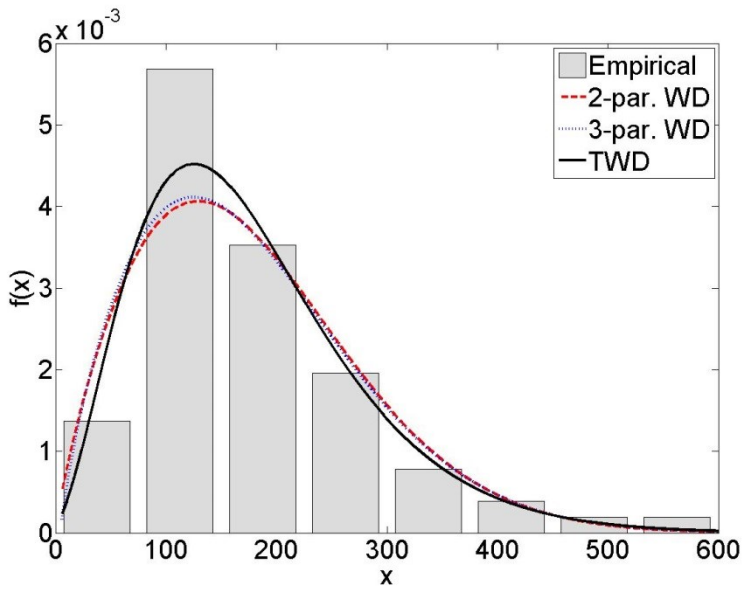


Fig. 2. Histogram of data set 2 and the fitted PDFs

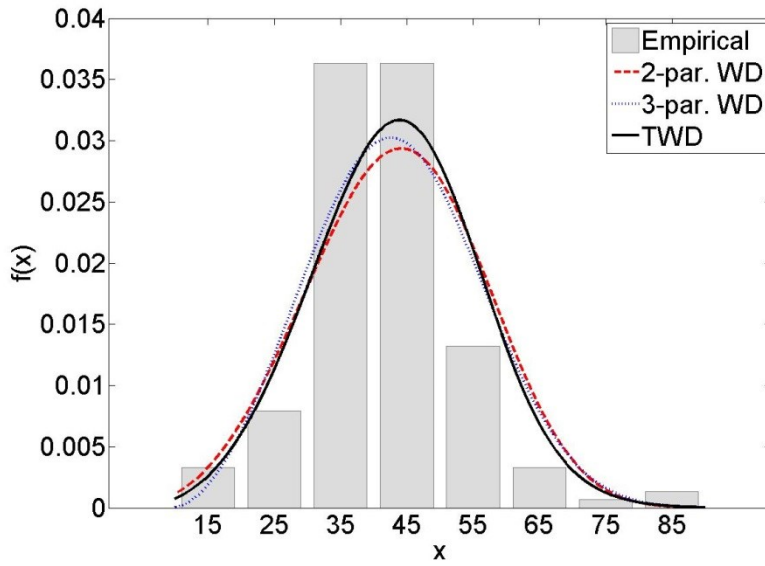


Fig. 3. Histogram of data set 3 and the fitted PDFs

Conclusion

In this paper we studied modification of the Weibull distribution, called the transmuted Weibull distribution. The usefulness of this distribution for modelling lifetime was illustrated using three real data sets. We compared the performance of the transmuted Weibull distribution with performances of the 2-parameter and the 3-parameter Weibull distributions. Comparison of distributions was based on criteria as the Akaike's information criterion, the corrected Akaike's information criterion, the Bayesian information criterion, the coefficient of determination and the root mean square error. We demonstrated that the transmuted Weibull distribution is more flexible and it better models lifetime data of animate (breast tumour patients, infected guinea pigs) and inanimate (Kevlar/epoxy strands) objects than the 2-parameter or the 3-parameter Weibull distribution.

This paper was supported by the Slovak Grant Agency VEGA through the project No. 1/0812/17.

References

1. W. Weibull, *A statistical distribution function of wide applicability*. Journal of Applied Mechanics **18** (3), 293-297 (1951)
2. H. Rinne, *The Weibull distribution: a handbook*. (Chapman & Hall, CRC Press, 2008)
3. H. Pham, C. D. Lai, *On recent generalizations of the Weibull distributions*. IEEE Transactions on Reliability **56** (3), 454-458 (2007)
4. W T. Shaw, I. R. Buckley, *The alchemy of probability distributions: beyond Gram-Charlier expansions and a skew kurtotic-normal distribution from a rank transmutation map*. Research report (2007)
5. G. R. Aryal, Ch. P. Tsokos, *On the transmuted extreme value distribution with application*. Nonlinear Analysis: Theory, Methods and Applications **71**, 1401-1407 (2009)
6. F. Merovci, *Transmuted Rayleigh Distribution*. Austrian Journal of Statistics **42** (1), 21-31 (2013)
7. M. S. Khan, R. King, *Transmuted Modified Weibull Distribution: A Generalization of the Modified Weibull Probability Distribution*. European Journal of Pure and Applied Mathematics **6**, 66-88 (2013)
8. M. S. Khan, R. King, R. Hudson, *Characterization of the transmuted inverse Weibull Distribution*. ANZIAM **55**, 197-217, (2013)
9. F. Merovci, L. Puka, *Transmuted Pareto Distribution*. ProbStat Forum **7**, 1-11 (2014)
10. G. R. Aryal, Ch. P. Tsokos, *Transmuted Weibull Distribution: A Generalization of the Weibull Probability Distribution*. European Journal of Pure and Applied Mathematics **4** (2), 89-102 (2011)
11. H. Akaike, *A new look at the statistical model identification*, Automatic Control, IEEE Transactions **19** (6), 716-723 (1974)
12. C. M. Hurvich, C. L. Tsai, *Regression and time series model selection in small samples*. Biometrika **76** (2), 297-307 (1989)
13. G. Schwarz, *Estimating the dimension of a model*, Annals of Statistics **6** (2), 461-464, (1978)
14. R. E. Barlow, R. H. Toland, T. Freeman, *A Bayesian analysis of stress rupture life of Kevlar 49/epoxy spherical pressure vessels*. Proceedings conference on applications of statistics, Marcel dekker, New York (1984)

15. D. F. Andrews, A. M. Herzberg, *Data: A Collection Of Problem from Many Fields for the Student and Researcher Worker*. (Springer Series in Statistics, New York, 1985)
16. I. B. Abdul-Moniem, *Transmuted Burr Type III Distribution*. Journal of Statistics: Advances in Theory and Applications **14** (1), 37-47 (2015)
17. E. A. Owoloko, P. E. Oguntunde, A. O. Adejumo, *Performance rating of the transmuted exponential distribution: an analytical approach*. Springer Plus **4**, 818 (2015)
18. T. Bjerkedal, *Acquisition of resistance in guinea pigs infected with different doses of virulent tubercile bacilli*. American Journal of Hygiene **72**, 100-148 (1960)
19. E. A. Owoloko, P. E. Oguntunde, A. O. Adejumo, *A Comparative Analysis on the Performance of the Convoluted Exponential Distribution and the Exponential Distribution in terms of Flexibility*. Journal of Mathematics and Statistics **12** (1) (2016)
20. M. M. Mansour, E. M. Abd Elrazik, M. S. Hamed, S. M. Mohamed, *A New Additive Weibull Distribution: Based on a New Method for Adding Parameter to a Family Distributions*. International Journal of Applied Mathematical Sciences **8** (1), 01-54 (2015)