2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Comparison between Modified Weighted Pareto Distribution and Many other Distributions

Mahmood A. Shamran*, ***Asmaa A. Mohammed*** and ***Iden H. Alkanani***

Department of Mathematics, College of Science for Women, University of Baghdad, Baghdad, Iraq.
*Corresponding Author.

Received 26/11/2022, Revised 04/02/2023, Accepted 06/02/2023, Published 20/06/2023

Abstract

In 2020 one of the researchers in this paper, in his first research, tried to find out the Modified Weighted Pareto Distribution of Type I by using the Azzalini method for weighted distributions, which contain three parameters, two of them for scale while the third for shape.This research compared the distribution with two other distributions from the same family; the Standard Pareto Distribution of Type I and the Generalized Pareto Distribution by using the Maximum likelihood estimator which was derived by the researchers for Modified Weighted Pareto Distribution of Type I, then the Mont Carlo method was used–that is one of the simulation manners for generating random samples data in different sizes (n= 10,30,50), and in different initial values for each Pareto distribution family being used in the research. A comparison was done by using Akaike Information Criteria, Corrected Akaike Information Criteria, and Bayesian Information Criteria.

Keywords: Akaike Information Criterion, Bayesian Information Criterion, Corrected Akaike Information Criterion, Generalized Pareto Distribution, Maximum likelihood Estimator, Modified Weighted Pareto Distribution Type I, Standard Pareto Distribution Type I.

Introduction

The aim of this research boils down to three goals, the first goal is to find out and derive what was carried out by the researchers about the Maximum likelihood estimator (MLE) to estimate the three parameters for Modified Weighted Pareto Distribution of Type I (MWPDTI). The second goal is to compare the new distribution with two other distributions of the same family by using Akaike Information Criteria (AIC),(AICc), and (BIC). While the third goal is to generate data by using the Mont Carlo method in order to generate random samples of different sizes; with different initial values to compare the same- Pareto family distribution.

Azzalini[1], found a method for computing the probability density function for weighted distributions according to Eq.1.

$$f_w(x) = \frac{1}{P_r(x_2<\theta x_1)} f(x_1)F(x_1) \qquad 1$$

Where $F(x_1)$ is the cumulative distribution function, this technique has opened up new challenges in the field of weighted probability distributions.

Para and Jan[2], they submitted a generalization of the Pareto distribution of the second type using the idea of weighted distributions, and the statistical properties of this generalization were derived.

Alkanani et al[3], used two methods to estimate the parameters of the Maxwell-Boltzmann distribution, one of which is the method of Maximum Likelihood estimator method used in this paper.. Nagatsuka et al [4] 2021, they presented a new development of the inference for the generalized Pareto distribution operating on all values of the k-

2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

shape parameter, in which they extracted new properties that lead to obtaining new confidence intervals. Omekam et al[5], they used the method of introducing new parameters on the distribution of Pareto of the first type, resulting in a new distribution, and a more flexible distribution with new types of data. Pho KH et al[6], in studying the problem of traffic accidents, using the comparison between Akaike's criterion and Bayes' criterion to get the best estimate of the parameters of the problem model. Portet S.[7], used the Akaike information criterion to study some biological models to obtain the best model in terms of estimating highly accurate parameters.

Pham MH. et al [8], estimated the parameters of the Generalized Pareto Distribution (GPD) and they tested the goodness of fit. Cavanaugh JE.et al [9], they reviewed the Akaike criterion, its properties, its derivation, and its use in predicting the estimation of the parameters of the statistical model [9].

Sahmran MA.[10], used the Azzalini method to find a modified and weighted distribution of a Pareto type I (MWPDTI) according to Eq.2.

$$f_w^*(x;\alpha,k,\theta)=\frac{2\alpha k^{\alpha}x^{-\alpha-1}\left(1-\frac{k^{\alpha}}{\theta^{\alpha}x^{\alpha}}\right)}{\theta^{\alpha}} \quad ,\alpha>0, x\geq\frac{k}{\theta}>0,\theta\geq 1,k>0 \qquad 2$$

Al Sarraf NM et al [11], they studied some methods of parameter estimation for the standard Pareto distribution of type I.

Preliminary:

The probability density function (pdf) and the cumulative distribution function (cdf) for the MWPDTI are equations respectively:

$$f_w^*(x;\alpha,k,\theta)=\frac{2\alpha k^{\alpha}x^{-\alpha-1}\left(1-\frac{k^{\alpha}}{\theta^{\alpha}x^{\alpha}}\right)}{\theta^{\alpha}} \quad ,\ \alpha>0, x\geq\frac{k}{\theta}>0,\theta\geq 1,k>0 \qquad 3$$

$$F_w^*(x)=\frac{(\theta^{\alpha}x^{\alpha}-k^{\alpha})^2}{\theta^{2\alpha}x^{2\alpha}}, x\geq\frac{k}{\theta}>0\ ,\alpha>0\ ,\theta\geq 1 \qquad 4$$

The mean and the variance of the MWPDTI are equations respectively:

$$E_w^*(x)=\frac{2\alpha^2 k}{\theta(1-2\alpha)(1-\alpha)} \quad ,\alpha>1\ ,\theta\geq 1,k>0 \qquad 5$$

$$Var_w^*(x)=\frac{(5\alpha-1)\alpha^2k^2}{(\alpha-2)(\alpha-1)^2(2\alpha-1)\theta^2} \quad ,\alpha>2\ ,\theta\geq 1,k>0 \qquad 6$$

The survival function and the hazard functions are equations respectively

$$\overline{F_w^*}(x)=\left[1-\frac{(\theta^{\alpha}x^{\alpha}-k^{\alpha})^2}{\theta^{2\alpha}x^{2\alpha}}\right]\ ,\theta>1,\alpha>0,x\geq\frac{k}{\theta}>0 \qquad 7$$

$$h_w^*(x)=\frac{2\alpha\theta^{\alpha}x^{\alpha-1}}{2\theta^{\alpha}x^{\alpha}-k^{\alpha}}-\frac{2\alpha k^{\alpha}}{x(2\theta^{\alpha}x^{\alpha}-k^{\alpha})}\ ,\ x\geq\frac{k}{\theta}>0,\alpha>0,\theta\geq 1 \qquad 8$$

And the other properties of the MWPDTI are mentioned in [10].

Materials and Methods

Maximum Likelihood Estimation Method (MLE):

The maximum likelihood estimator is used to find the estimated values for the parameters of MWPDTI.

From Eq.1, obtaining to

$$f_w^*(x;\alpha,\theta,k)=2\alpha k^{\alpha}\theta^{-\alpha}(x^{-\alpha-1}-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}) \qquad 9$$

Then $L(x;\alpha,\theta,k)=2^n\alpha^n k^{n\alpha}\theta^{-n\alpha}\prod_{i=1}^{n}(x^{-\alpha-1}-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})$ 10

Taken the log of Eq.10, getting

$$\ln L(x;\alpha,\theta,k)=n\ln 2+n\ln\alpha+n\alpha\ln k-n\alpha\ln\theta+\sum_{i=1}^{n}\ln(x^{-\alpha-1}-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}) \qquad 11$$

$$\text{Let } f(\hat{\alpha})=\frac{\partial\ln L}{\partial\alpha}=\frac{n}{\alpha}+n\ln k-n\ln\theta+\sum_{i=1}^{n}\frac{-x^{-\alpha-1}\ln x-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}[\ln k-\ln\theta-2\ln x]}{(x^{-\alpha-1}-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})} \qquad 12$$

$$\text{Let } g(\hat{k})=\frac{\partial\ln L}{\partial k}=\frac{n\,\alpha}{k}+\sum_{i=1}^{n}\frac{-x^{-2\alpha-1}k^{\alpha-1}\alpha\theta^{-\alpha}}{(x^{-\alpha-1}-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})} \qquad 13$$

$$\text{Let } h(\hat{\theta})=\frac{\partial\ln L}{\partial\theta}=\frac{-n\,\alpha}{\theta}+\sum_{i=1}^{n}\frac{\alpha\,x^{-2\alpha-1}k^{\alpha}\theta^{-\alpha-1}}{(x^{-\alpha-1}-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})} \qquad 14$$

2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

When $\frac{\partial \ln L}{\partial \alpha} = \frac{\partial \ln L}{\partial k} = \frac{\partial \ln L}{\partial \theta} = 0$, in this case, there is no closed solution for Eqs.12, 13, and 14, which must use numerical methods such as Newton – Raphson method.

$$\begin{bmatrix} \hat{\alpha}_{i+1} \\ \hat{k}_{i+1} \\ \hat{\theta}_{i+1} \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_{i} \\ \hat{k}_{i} \\ \hat{\theta}_{i} \end{bmatrix} - J^{-1} \begin{bmatrix} f(\hat{\alpha}_i) \\ g(\hat{k}_i) \\ h(\hat{\theta}_i) \end{bmatrix} \quad 15$$

Where $$J = \begin{bmatrix} \frac{\partial f(\hat{\alpha})}{\partial \alpha} & \frac{\partial f(\hat{\alpha})}{\partial k} & \frac{\partial f(\hat{\alpha})}{\partial \theta} \\ \frac{\partial g(\hat{k})}{\partial \alpha} & \frac{\partial g(\hat{k})}{\partial k} & \frac{\partial g(\hat{k})}{\partial \theta} \\ \frac{\partial h(\hat{\theta})}{\partial \alpha} & \frac{\partial h(\hat{\theta})}{\partial k} & \frac{\partial h(\hat{\theta})}{\partial \theta} \end{bmatrix} \quad 16$$

$$\frac{\partial f(\hat{\alpha})}{\partial \alpha} = \frac{-n}{\alpha^2} + \sum_{i=1}^{n} \frac{[(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})(\, x^{-\alpha-1}(\ln x)^2 - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}(\ln k - \ln\theta - 2\ln x)^2]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2}$$
$$\frac{[-x^{-\alpha-1}(\ln x)^2 - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}(\ln k - \ln\theta - 2\ln x)]^2}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 17$$

$$\frac{\partial f(\hat{\alpha})}{\partial \theta} = \frac{-n}{\theta} + \sum_{i=1}^{n} [\frac{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})[(-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha-1})(-\alpha(\ln k - \ln\theta - 2\ln x) - 1)]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2}$$
$$\frac{-(\alpha k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha-1})(-x^{-\alpha-1}\ln x - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}(\ln k - \ln\theta - 2\ln x))}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 18$$

$$\frac{\partial f(\hat{\alpha})}{\partial k} = \frac{n}{k} + \sum_{i=1}^{n} \left[\frac{(-x^{-2\alpha-1}k^{\alpha-1}\theta^{-\alpha})[(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})(\alpha(\ln k - \ln\theta - 2\ln x) + 1]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2}\right]$$
$$\frac{+\alpha[(-x^{-\alpha-1}\ln x - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}(\ln k - \ln\theta - 2\ln x))]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 19$$

$$\frac{\partial g(\hat{k})}{\partial \alpha} = \frac{n}{k} + \sum_{i=1}^{n} \frac{(-x^{-2\alpha-1}k^{\alpha-1}\alpha\theta^{-\alpha})[(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})(\alpha(\ln k - \ln\theta - 2\ln x) + 1]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2}$$
$$\frac{+\alpha[(-x^{-\alpha-1}\ln x - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}(\ln k - \ln\theta - 2\ln x))]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 20$$

$$\frac{\partial g(\hat{k})}{\partial \theta} = \sum_{i=1}^{n} \frac{(\alpha^2 x^{-3\alpha-2}k^{\alpha-1}\theta^{-\alpha-1})}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 21$$

$$\frac{\partial g(\hat{k})}{\partial k} = \frac{-n\alpha}{k^2} + \sum_{i=1}^{n} \frac{(-\alpha x^{-2\alpha-1}k^{\alpha-2}\theta^{-\alpha})[(\alpha-1)(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}) + k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 22$$

$$\frac{\partial h(\hat{\theta})}{\partial \alpha} = \frac{-n}{\theta} + \sum_{i=1}^{n} [\frac{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})[(-k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha-1})(-\alpha(\ln k - \ln\theta - 2\ln x) - 1)]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2}$$
$$\frac{-(\alpha k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha-1})(-x^{-\alpha-1}\ln x - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha}(\ln k - \ln\theta - 2\ln x))}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 23$$

$$\frac{\partial h(\hat{\theta})}{\partial \theta} = \frac{n\alpha}{\theta^2} + \sum_{i=1}^{n} \frac{(\alpha k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha-1})[\left(\left(x^{-\alpha-1}(-\alpha-1)\theta^{-1} - (k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha-1})(-\alpha-1)\right)\right) - (\alpha k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})]}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 24$$

$$\frac{\partial h(\hat{\theta})}{\partial k} = \sum_{i=1}^{n} \frac{(\alpha^2 x^{-3\alpha-2}k^{\alpha-1}\theta^{-\alpha-1})}{(x^{-\alpha-1} - k^{\alpha}x^{-2\alpha-1}\theta^{-\alpha})^2} \quad 25$$

2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Because the matrix J is a non-singular symmetric matrix, then $\frac{\partial f(\hat{\alpha})}{\partial \theta} = \frac{\partial h(\hat{\theta})}{\partial \alpha}, \frac{\partial f(\hat{\alpha})}{\partial k} = \frac{\partial g(\hat{k})}{\partial \alpha}$ and $\frac{\partial g(\hat{k})}{\partial \theta} = \frac{\partial h(\hat{\theta})}{\partial k}$.

The maximum likelihood function for SPDTI distribution as in form [11] :

$$L(\alpha|x_n) = \left(\alpha^{nk} k^n\right) \prod_{i=1}^{n} (x_i)^{-(\alpha+1)}$$ 26

And the GPD distribution Maximum likelihood function is in form [8]:

$$L(k, \alpha; X) = \begin{cases} -n \ln \alpha + \frac{1}{\alpha} \sum_{i=1}^{n} x_i & k = 0 \\ -n \ln\alpha + \left(\frac{1}{k} - 1\right) \sum_{i=1}^{n} \ln\left(1 - \frac{k x_i}{\alpha}\right) & k \neq 0 \end{cases}$$ 27

The Three Criteria (AIC, AIC$_c$ and BIC):

The following statistical criteria are used (AIC, AICc and BIC) to compare the MWPDTI with two Pareto distributions (SPDTI, GPD).

$$AIC = 2m - \hat{L}$$ 28

$$AIC_c = AIC + \frac{2m(m+1)}{n-m-1}$$ 29

and $$BIC = m \ln(n) - 2 \ln(\hat{L})$$ 30

Where m is the number of parameters in the distribution, $\hat{L}$ is the maximum log-likelihood function value for the distribution and n represents the sample size.

The Simulation Technique:

The Monte Carlo method is the most popular simulation technique used to generate the observations (samples) for any distribution. The simulation method is flexible for the test of experiments by replications many times.

1. Applying the Monte Carlo method for Modified Weighted Pareto distribution Type I (MWPDTI), using the cumulative distribution function as follows:

$$F(x) = \frac{(\theta^\alpha x^\alpha - k^\alpha)^2}{\theta^{2\alpha} x^{2\alpha}}, x \geq \frac{k}{\theta} > 0, \alpha > 0, \theta \geq 1$$ 31

By substituting $F(x)$ as u, with a random number the $u = F(x)$, that is $x = F^{-1}(u)$

$$\frac{(\theta^\alpha x^\alpha - k^\alpha)^2}{\theta^{2\alpha} x^{2\alpha}} = u$$ 32

$$\frac{(\theta^\alpha x^\alpha - k^\alpha)^2}{(\theta^\alpha)^2 (x^\alpha)^2} = u$$ 33

$$\frac{\theta^\alpha x^\alpha - k^\alpha}{\theta^\alpha x^\alpha} = \sqrt{u}$$ 34

$$\frac{\theta^\alpha x^\alpha}{\theta^\alpha x^\alpha} - \frac{k^\alpha}{\theta^\alpha x^\alpha} = \sqrt{u}$$ 35

$$1 - \frac{k^\alpha}{\theta^\alpha x^\alpha} = \sqrt{u}$$ 36

$$[1 - \sqrt{u}] = \frac{k^\alpha}{\theta^\alpha x^\alpha}$$ 37

$$x^\alpha = \frac{k^\alpha}{\theta^\alpha [1 - \sqrt{u}}$$ 38

$$x = [\frac{k^\alpha}{\theta^\alpha [1 - \sqrt{u}}]^{\frac{1}{\alpha}}$$ 39

2. Generating random samples which are $n = 10, 30, 50$.
3. The initial values of the (SPDTI),(GPD) and (MWPDTI) are:

Table No.	α	k	θ
1	1	0.1	**1**
2	1	0.002	**0.1**
3	1.5	0.008	**2**

4. The replicate is $R = 1000$.

Then compare three distributions by (AIc, AICc, and BIC) criteria to find the best distribution than another's.

The result Estimated Value of Parameters with corresponding Criteria AIC, AICc, and BIC as in Table 1-3.

2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Table 1. The result when $\alpha = 1$, $\theta = 1$, $k = 0.1$ with sample size 10, 30 and 50

Initial values of parameters	distribution name	Sample Size	estimated parameters	AIC	AIC_C	BIC
α=1 **k=0.1**	SPDTI (2parameters)	10	α̂=1.004799 k̂=0.099993	44.871	46.575	45.931
		30	α̂=1.001398 k̂=0.0999820	83.5466	82.8741	82.8453
		50	α̂=1.000898 k̂=0.0997536	115.5643	115.8243	116.06
α=1 **k=0.1**	GPD (2parameters)	10	α̂=0.94008147 k̂=0.99987687	44.5949	48.5949	45.4182
		30	α̂=1.000490 k̂=0.1000265	83.6325	84.5555	84.4183
		50	α̂=0.9993400 k̂=0.099792	114.6952	115.2169	115.2830
α=1 **k=0.1** **θ=1**	MWPDTI (3 parameters)	10	α̂=1.000493 k̂=0.099489 θ̂=1.0070863	42.8642	46.8642	44.772
		30	α̂=1.0002314 k̂=0.1000213 θ̂=1.0004647	81.2756	82.0086	82.356 3
		50	α̂=1.00004996 k̂=0.10090915 θ̂=0.9973012	114.1271	114.6488	115.7854

Table. 1 showing the value of the three criteria increases with the sample size increasing.

1. When $n = 10$ the MWPDTI distribution has minimum values of AIC, and BIC, but the SPDTI distribution has less value in AICc.
2. When $n = 30$the MWPDTI distribution has less values AIC, AICc, and BIC.
3. When $n = 50$ the MWPDTI distribution has less values AIC, and AICc, but the GPD distribution has less BIC.
4. Although the results for AIC, AICc, and BIC are converged from each other.

Table 2. The result when $\alpha = 1$, $\theta = 0.1$, $k = 0.002$ with sample size 10, 30 and 50

Initial values of parameters	distribution name	Sample Size	estimated parameters	AIC	AIC_C	BIC
α=1 k=0.002	SPDTI (2parameters)	10	α̂=1.004799 k̂=.00200043	32.2321	33.9461	34.21735
		30	α̂=1.0007543 k̂=0.0019949	53.5466	54.0081	54.8432
		50	α̂=1.0000542 k̂=0.0019730	87.3421	87.6021	88.1796
α=1 k=0.002	GPD (2parameters)	10	α̂=0.99400814 k̂=0.00201905	34.9765	38.9765	39.6276
		30	α̂=1.000490 k̂=0.0019023	55.2381	56.1611	56.796
		50	α̂=1.0097859 k̂=0.0019853	88.1346	88.6563	89. 4798
α=1 k=0.002 θ=0.1	MWPDTI (3 parameters)	10	α̂=1.0000493 k̂=0.00194891 θ̂=0.0989708	30.3651	34.3651	34.7876
		30	α̂=1.0083949 k̂=0.0021436 θ̂=0.0998735	52.6579	53.5809	54.1352
		50	α̂=0.997499 k̂=0.00998646 θ̂=0.10009418	86.1436	86.6653	87.0631

2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

Showing in Table 2 that

1. When $n = 10$ the SPDTI distribution has less values in AICc, and BIC, but the MWPDTI distribution has less value in AIC.
2. When $n = 30$ the MWPDTI distribution has less values in AIC, AICc, and BIC.
3. When $n = 50$ the MWPDTI distribution has less values in AIC, AICc, and BIC.
4. Although the results for AIC, AICc, and BIC are converged from each other.

Table 3. The result when $\alpha = 1.5$, $\theta = 2$, $k = 0.008$ with sample sizes 10, 30 and 50

Initial values of parameters	distribution name	Sample Size	estimated parameters	AIC	AIC$_C$	BIC
α=1.5 k=0.008	SPDTI (2parameters)	10	$\hat{\alpha}$=1.499950 $\hat{k}$=0.00099768	42.7364	44.4504	45.2376
		30	$\hat{\alpha}$=1.5007543 $\hat{k}$=0.0079983	73.546	74.0075	74.8326
		50	$\hat{\alpha}$=1.50000542 $\hat{k}$=0.00095360	118.396	118.656	119.879
α=1.5 k=0.008	GPD (2parameters)	10	$\hat{\alpha}$=1.499905632 $\hat{k}$=0.00092743	43.247	47.247	48.6965
		30	$\hat{\alpha}$=1.500490 $\hat{k}$=0.007956	71.212	72.135	72.9342
		50	$\hat{\alpha}$=1.5000635 $\hat{k}$=0.00906854	118.164	118.6857	119.2365
α=1.5 k=0.008 θ=2	MWPDTI (3 parameters)	10	$\hat{\alpha}$=1.4999873 $\hat{k}$=0.00800674 $\hat{\theta}$=1.99936724	41.641	43.641	44.3421
		30	$\hat{\alpha}$=1.500039 $\hat{k}$=0.00814231 $\hat{\theta}$=2.00001999	70.584	71.507	72.1231
		50	$\hat{\alpha}$=1.500004998 $\hat{k}$=0.00800443 $\hat{\theta}$=1.99991681	117.465	117.9867	118.3843

Table 3 shows that

1. When $n = 10$ the MWPDTI distribution has less values in AIC, AICc, and BIC.
2. When $n = 30$ the MWPDTI distribution has less values in AIC, AICc, and BIC.
3. When $n = 50$ the MWPDTI distribution has less values in AIC, AICc, and BIC.
4. Although the results for AIC, AICc, and BIC are converged from each other.

Conclusion

The researchers found out and derived the (MLE) to estimate the three parameters for (MWPDTI), then they compared the three distributions used in this research by (AIC),(AICc) and (BIC) to represent the efficiency of the statistical distribution, the three criteria used in this research must be in minimum value i.e.(AIC, AICc, BIC). By using the simulation manner, it was found that the new (MWPDTI) contains a minimum value for all sizes of the generated random samples, except for a few samples. It was obvious that the new (MWPDTPI) is more efficient and better 23 times, (GPD) is more efficient one time, and (SPDTI) is three times more effective.

In the future, it can be confirmed the efficiency of (MWPDTI) by using other estimation methods to estimate the distribution parameters.

Author’s Declaration

- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are ours. Furthermore,

any Figures and images that are not ours have been included with the necessary permission for re-publication, which is attached to the manuscript.

- Ethical Clearance: The project was approved by the local ethical committee at the University of Baghdad.

Author's Contribution Statement

M.A. and I. H. presented the idea and discussed it theoretically and checked the equations. Authors M. A. and A. A. made the derivations for the equations. Authors M. A. and I. H. discuss and validate the results. M. A. prepared the computer programs for the simulation method. I. H. Conceptualization (supporting the idea and goal, methodology, and reviewing).

References

1. Azzalini A. A Class of Distributions Which Includes the Normal Ones. Scand J Stat. 1985 Jan 1; 12(2): 171-178. http://www.jstor.org/stable/4615982.
2. Para BA, Jan TR. On Three Parameter Weighted Pareto Type II Distribution: Properties and Applications in Medical Sciences. Amisl. 2018; 6(1): 13-26. http://dx.doi.org/10.18576/amisl/06010.
3. Alkanani IH, Salman SG. Bayes and Non-Bayes Estimation Methods for the Parameter of Maxwell-Boltzmann Distribution. Baghdad Sci J. 2017; 14(4). http://dx.doi.org/10.21123/bsj.2017.14.4.0808
4. Nagatsuka H, Balakrishnan N. Efficient Likelihood-Based Inference for the Generalized Pareto Distribution. Ann Inst Stat Math. 2021 Dec; 73(6): 1153-85.
5. Omekam IV, Popoola J, Gatta NF, Adejumo AO. Some Extended Pareto Type I Distributions. Ife J Sci. 2022 Oct 13; 24(2): 265-76. https://dx.doi.org/10.4314/ijs.v24i2.8.
6. Pho KH, Ly S, Ly S, Lukusa TM. Comparison Among Akaike Information Criterion, Bayesian Information Criterion and Vuong's Test in Model Selection: A Case Study of Violated Speed Regulation in Taiwan. J Adv Eng Comput. 2019 Mar 31;3(1): 293-303. http://dx.doi.org/10.25073/jaec.201931.220.
7. Portet S. A Primer on Model Selection Using the Akaike Information Criterion. Infect Dis Model. 2020 Jan 1; 5: 111-28. https://doi.org/10.1016/j.idm.2019.12.010.
8. Pham MH, Tsokos C, Choi BJ. Maximum likelihood estimation for the generalized pareto distribution and goodness-of-fit test with censored data. J Mod Appl Stat Methods. 2019;17(2):11.. https://dx.doi.org/10.22237/jmasm/1553261471.
9. Cavanaugh JE, Neath AA. The Akaike Information Criterion: Background, Derivation, Properties, Application, Interpretation, and Refinements. WIREs Comput Stat. 2019: e1460. https://doi.org/10.1002/wics.1460.
10. Sahmran MA. Modified Weighted Pareto Distribution Type I (MWPDTI). Baghdad Sci J. 2020 Sep 1; 17(3): 869-877. https://doi.org/10.21123/bsj.2020.17.3.0869.
11. Al Sarraf NM, Mohana FA, Kamal GI. About Estimating Pareto Distribution Parameters. J *Rafidain* Uni Coll Sci. 2020; 2020(46): 431-40.

2023, 20(3 Suppl.): 1108-1115
https://dx.doi.org/10.21123/bsj.2023.8169
P-ISSN: 2078-8665 - E-ISSN: 2411-7986
Baghdad Science Journal

مقارنة بين توزيع باريتو المعدل الموزون والعديد من التوزيعات الأخرى

محمود عريبي شمران، اسماء عبدالحسين محمد و ايدن حسن حسين

قسم الرياضيات، كلية العلوم للبنات، جامعة بغداد، بغداد، العراق.

الخلاصة

قام احد الباحثين في هذه الورقة في سنة 2020 في بحثه الأول إيجاد توزيع باريتو الموزون المعدل من النوع الأول من خلال الاعتماد على طريقة Azzalini للتوزيعات الموزونة , وبعدها قام باشتقاق كافة الخواص الرياضية والاحصائية للتوزيع الجديد والذي يحتوي على ثلاثة معلمات اثنان منها معلمات القياس والثالثة معلمة الشكل.هذا البحث تمت مقارنة توزيع باريتو الموزون المعدل من النوع الأول مع توزيعين من نفس عائلة توزيع باريتو وهما توزيع باريتو القياسي من النوع الاول وتوزيع باريتو المعمم من خلال استخدام طريقة الإمكان الأعظم التي قام الباحثين باشتقاقها ومن ثم استخدمت طريقة مونت كارلو وهي احد أساليب المحاكاة لتوليد بيانات عينات عشوائية باحجام مختلفة n=10,30,50 وبقيم ابتدائية مختلفة لكل التوزيعات من عائلة توزيع باريتو المستخدمة في البحث وتمت المقارنة باستخدام معيار معلومات أكاكي، معيار معلومات اكاكي المصحح ومعيار المعلومات البيزية .

الكلمات المفتاحية: معيار معلومات أكاكي، معيار معلومات البيزية، معيار معلومات أكاكي المعدل، توزيع باريتو المعمم، دالة الإمكان الاعظم، توزيع باريتو الموزون المعدل من النوع I، توزيع باريتو القياسي من النوع I.