Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/090307

MG Exponentiated Pareto Distribution

Hussein Eledum[1,2,*] *and S. I. Ansari* [1,3]

[1]Department of Statistics, Faculty of Science, University of Tabuk, KSA
[2]Department of Applied Statistics, Shendi University, Sudan
[3]Department of Business Administration, Azad Institute of Engineering and Technology, Lucknow, India

Received: 1 Jul. 2019, Revised: 18 Oct. 2019, Accepted: 19 Oct. 2019
Published online: 1 Nov. 2020

Abstract: In this paper, we have introduced a new generalization of the Exponentiated Pareto distribution named as the MG Exponentiated Pareto (MGEP) distribution based on MG transmutation map introduced by Kumar et al. (2017). Furthermore, we have derived some statistical properties of the MGEP distribution. The parameters for the proposed distribution is estimated using maximum likelihood method and the performance of estimators is studied by using simulation. An application of MGEP distribution to a real data set for the purpose of illustration is conducted.

Keywords: Exponential Pareto distribution, Survival function, Maximum likelihood estimation, Hazard function, Quantile function

1. Introduction

The statistical literature filled with several methods to propose new distribution, by the use of some available distributions, are called baseline distributions. For instance, Gupta et al [1] suggested a family of distributions termed as exponentiated exponential distribution by using cumulative distribution function of a new distribution corresponding to the cumulative distribution function of the baseline distribution function. Shaw and Buckley [2] developed the transmutation maps to solve financial mathematics problems. Based on Shaw and Buckley's quadratic transmuted family, several forms of cubic ranking have been proposed. For example, Granzottoa et al [3] suggested the cubic ranking transmutation map with two transmutation parameters. Rahman et al [4] proposed an extension of the quadratic transmuted distribution and named the resulting family as the cubic transmuted family of distribution. Al- Kadim [5] suggested cubic ranking transmutation map with single parameter. Another idea of getting a new distribution is to transform the baseline distribution. Kumar et al [6] introduced DUS transformation. Kumar et al [7] suggested SS transformation by use of the sine function. Kumar et al [8] proposed MG transmutation map which was used to generalizes the exponential distribution.

In this article MG transmutation map suggested by Kumar et al [8] is used to propose a new model which generalizes the exponentiated Pareto distribution. This new version of the Exponentiated Pareto distribution called MG Exponentiated Pareto (MGEP) distribution. Some statistical properties are studied including survival and hazard function, quantile function, mean, variance and random number generation. The parameters of MGEP distribution is estimated using Maximum Likelihood method and the performance of the estimators is studied using simulation. Moreover, an application of MGEP distribution to a real data set for the purpose of illustration is conducted.

The proposed MGEP distribution consider as one of a lifetime distribution, it is useful to analyze the skewed extremely positive data sets concerning the distribution of wealth.

* Corresponding author e-mail: heledum@yahoo.com

1.1. Exponentiated Pareto distribution

Let X be a random variable with the exponentiated Pareto distribution. The probability density function (pdf) and the cumulative distribution function (cdf) are defined, respectively, as

$$g(x) = ak^a e^{-ax} \tag{1}$$

and

$$G(x) = 1 - k^a e^{-ax}; \qquad x \geq \ln k, \qquad a,k > 0 \tag{2}$$

1.2. MG Transmutation

According to the MG transmutation approach proposed by Kumar et al[8], the cumulative distribution function (cdf) satisfies the following relationship

$$F(x) = \exp\left(1 - \frac{1}{G(x)}\right) \tag{3}$$

and the density function (pdf) is given by

$$f(x) = \frac{\exp(1 - \frac{1}{G(x)})}{[G(x)]^2} g(x) \tag{4}$$

where $G(x)$ and $g(x)$ is the cdf and pdf of the base distribution respectively.

1.3. Organization of the paper

The remainder of this paper is organized as follows. The new proposed distribution MGEP and its hazard and survival function is presented in Section 2. Some statistical properties is derived in Section 3. Section 4 provides parameter estimation of MGEP. Simulation is conducted in section 5. An application of MGEP to a real data set for the purpose of illustration is provided in Section 6. Finally, Section 7 gives some concluding remarks.

2. MG Exponentiated Pareto (MGEP) Distribution

In this Section, the new proposed distribution MG Exponentiated Pareto (MGEP) is demonstrated.

2.1. Cumulative and Density Function

Lemma 1. Let X be a continuous random variable; follows an exponentiated Pareto distribution then the cumulative distribution function (cdf) and probability density function (pdf) of the MGEP are respectively given by

$$F(x) = \exp\left(-\frac{k^a e^{-ax}}{1 - k^a e^{-ax}}\right) \tag{5}$$

and

$$f(x) = \frac{ak^a e^{-ax}}{(1 - k^a e^{-ax})^2} \exp\left(-\frac{k^a e^{-ax}}{1 - k^a e^{-ax}}\right); \qquad x \geq \ln k, \qquad a,k > 0 \tag{6}$$

Proof. Proof is straightforward.
Figure 1 illustrates some of possible shapes of the pdf and cdf of MGEP for selected values of parameters a where $k = 1.5$.

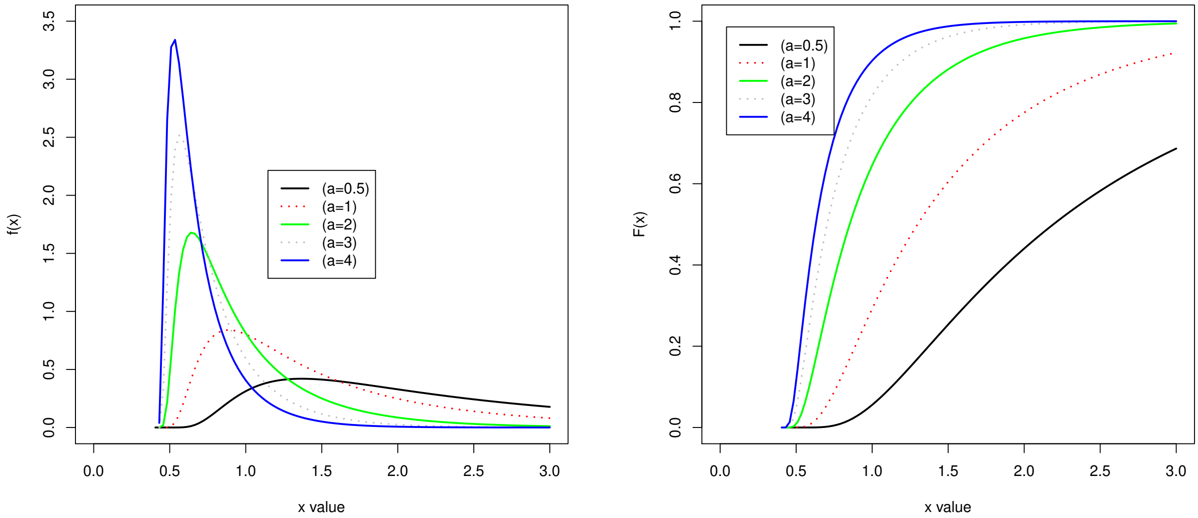

Fig. 1: The pdf and cdf of MGEP for different value of a.

From plot of pdf of Figure 1, we can observe that as the shape parameter a increases, the distribution is skewed extremely positive and the kurtosis is becomes more leptokurtic.

Lemma 2. The limit of MGEP density as $x \to \ln k$ is ∞ and the limit as $x \to \infty$ is 0.

Proof. The proof is straightforward.

Lemma 3. $f(x)$ of Eq.(6) is a probability density function.

To prove $f(x)$ is a pdf, we need to prove $f(x) \geq 0$ and $\int_{\ln k}^{\infty} f(x)dx = 1$.

(1) Proof of $f(x) \geq 0$

From Lemma 2, since $\lim_{x \to \ln k} f(x) = \infty$ and $\lim_{x \to \infty} f(x) = 0$ then $f(x) \geq 0$

(2) Proof of $\int_{\ln k}^{\infty} f(x)dx = 1$

$$\int_{\ln k}^{\infty} f(x)dx = \int_{\ln k}^{\infty} \left[\frac{ak^a e^{-ax}}{(1-k^a e^{-ax})^2} \exp\left(- \frac{k^a e^{-ax}}{1-k^a e^{-ax}} \right) \right] dx$$

let $\frac{k^a e^{-ax}}{(1-k^a e^{-ax})} = u$ then $\frac{ak^a e^{-ax}}{(1-k^a e^{-ax})^2} dx = -du$, for $x = \ln k \to u = \infty, x = \infty \to u = 0$

$$\int_{\ln k}^{\infty} f(x)dx = \int_{\infty}^{0} e^{-u}(-du) = \int_{0}^{\infty} e^{-u}du = \left[e^{-u} \right]_0^{\infty} = 1$$

From 1 and 2 we conclude that $f(x)$ is probability density function.

2.2. *Survival and hazard function*

The survival (reliability) function of MGEP is

$$S(x) = 1 - \exp\left(- \frac{k^a e^{-ax}}{1-k^a e^{-ax}} \right) \tag{7}$$

The hazard function of MGEP is

$$h(x) = \frac{\frac{ak^a e^{-ax}}{(1-k^a e^{-ax})^2} \exp\left(- \frac{k^a e^{-ax}}{1-k^a e^{-ax}} \right)}{1 - \exp\left(- \frac{k^a e^{-ax}}{1-k^a e^{-ax}} \right)} \tag{8}$$

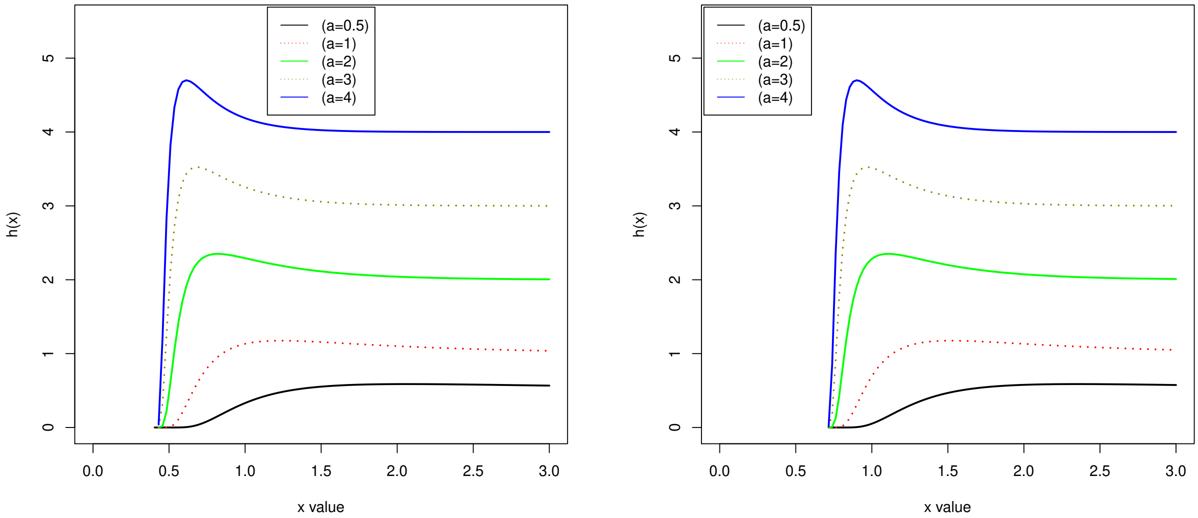

Fig. 2: The hazard function of MGEP for different values of a at $k = 1.5$ and $k = 2$.

Figure 2 illustrates the hazard function of MGEP for different values of a at $k = 1.5$ and $k = 2$.
From Figure 2, it is shown that as a increases the hazard function curve moves upward. Moreover, as k increases the curve shifts to the right.

3. Statistical Properties

In this Section, some statistical properties for MGEP are demonstrated include arithmetic mean, variance and quantile function. This section also provides simulation of the random sample.

3.1. Arithmetic Mean

Theorem 1. If X is a random variable having the MGEP then the arithmetic mean μ_X is

$$\mu_X = E(X) = \ln k + \frac{1.173563}{a} \tag{9}$$

and

$$Var(X) = \frac{0.862770}{a^2}$$

Proof. We know that

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx$$

Substitute $f(x)$ in above equation by its value in Eq.(6) to get

$$E(X) = \int_{\ln k}^{\infty} x \Big[\frac{a k^a e^{-ax}}{(1 - k^a e^{-ax})^2} \exp\Big(- \frac{k^a e^{-ax}}{1 - k^a e^{-ax}} \Big) \Big] dx$$

let $\frac{k^a e^{-ax}}{(1-k^a e^{-ax})} = u$ (see proof (2) for Lemma 3) and it follows that $x = \ln k + \frac{1}{a}\ln(\frac{1+u}{u})$

$$\begin{aligned} E(X) &= \int_0^\infty \left[\ln k + \frac{1}{a}\ln(\frac{1+u}{u})\right] e^{-u} du \\ &= \ln k \int_0^\infty e^{-u} du + \frac{1}{a}\left[\int_0^\infty e^{-u}\ln(1+u)du - \int_0^\infty e^{-u}\ln u\, du\right] \\ &= \ln k + \frac{1}{a}\left[-eE_i(-1) - (-\gamma)\right] \end{aligned}$$

where $E_i(x)$ is the exponential integral E_i , $-eE_i(-1) \approxeq 0.596347$ and $\gamma \approxeq 0.577216$ is the Euter-Mascheroni constant. Therefore,

$$E(X) = \ln k + \frac{1.173563}{a}$$

This completes the proof.
Similarly, we can prove that

$$Var(X) = \frac{0.862770}{a^2}$$

3.2. *Quantile function*

The quantile function for MGEP is derived by finding the value of Q for which $F(x) = p$:

$$Q(p,a) = \ln k + \frac{1}{a}\left[\ln(1-(\ln p)^{-1})\right] \quad \text{for} \quad 0 \le p \le 1 \tag{10}$$

The three quartiles $Q1, Q2$ and $Q3$ can be obtained by using $p = 0.25, 0.50$ and 0.75 in Eq.(10) respectively.
The mean, median and variance of MGEP for various values of a are given in Table 1.

Table 1: Mean, median and variance of MGEP distribution for different values of a.

Measure	$a = 0.5$	$a = 1$	$a = 2$	$a = 3$
Mean	2.7526	1.5790	0.9922	0.7967
Median	2.1917	1.2986	0.8520	0.7032
Variance	3.4511	0.8628	0.2157	0.0959

3.3. *Simulating the Random Sample*

Random numbers from the MGEP can be obtained by equating cdf of the distribution in Eq.(5) with a uniform random number and inverting the expression; that is the random number from MGEP is obtained by solving

$$\exp\left(-\frac{k^a e^{-ax}}{1-k^a e^{-ax}}\right) = u$$

for x. The random sample from MGEP can be further expressed as

$$x = \ln k + \frac{1}{a}\left[\ln(1-(\ln u)^{-1})\right] \tag{11}$$

where u is an arbitrary continuous uniform point over $(0,1)$.

4. Parameters Estimation

Maximum likelihood approach can be used for the estimation of model parameters. The maximum likelihood estimates (MLE) of the parameters that are inherent within the MGEP is given by the following.
Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from MGEP distribution. Then the likelihood function is given by

$$L = \prod_{i=1}^{n} f(x_i; a, k) = \prod_{i=1}^{n} \left[\frac{a k^a e^{-a x_i}}{(1 - k^a e^{-a x_i})^2} \exp\left(- \frac{k^a e^{-a x_i}}{1 - k^a e^{-a x_i}} \right) \right]$$

$$= a^n k^{na} e^{-a \sum_{i=1}^{n} x_i} \prod_{i=1}^{n} \left[\frac{1}{(1 - k^a e^{-a x_i})^2} \right] \exp\left(\sum_{i=1}^{n} - \frac{k^a e^{-a x_i}}{1 - k^a e^{-a x_i}} \right)$$

so, the log likelihood function is

$$\ln L = n \ln a + a n \ln k - a \sum_{i=1}^{n} x_i - 2 \sum_{i=1}^{n} \ln(1 - k^a e^{-a x_i}) - k^a \sum_{i=1}^{n} \frac{e^{-a x_i}}{1 - k^a e^{-a x_i}} \tag{12}$$

Therefore, the maximum likelihood estimates of a and k which maximize Eq.(12), must satisfy the two normal Eq.'s(13) and (14).

$$\frac{\partial \ln L}{\partial a} = k^a \sum_{i=1}^{n} \frac{(x_i - \ln k) e^{-a x_i}}{(1 - k^a e^{-a x_i})^2} + 2 k^a \sum_{i=1}^{n} \frac{(x_i - \ln k) e^{-a x_i}}{1 - k^a e^{-a x_i}} - \sum_{i=1}^{n} x_i + n \ln k + \frac{n}{a} = 0 \tag{13}$$

$$\frac{\partial \ln L}{\partial k} = n + 2 k^{a-1} \sum_{i=1}^{n} \frac{e^{-a x_i}}{1 - k^a e^{-a x_i}} - k^{a-1} \sum_{i=1}^{n} \frac{e^{-2 a x_i}}{(1 - k^a e^{-a x_i})^2} = 0 \tag{14}$$

The maximum likelihood estimates $\hat{\theta} = (\hat{a}, \hat{k})$ of $\theta = (a, k)$ is obtained by solving the above nonlinear system of Eq. (13) and (14). It is more convenient to use nonlinear optimization algorithms such as the quazi-Newton or Newton-Raphson to numerically maximize the log-likelihood function in Eq. (12).

5. Simulation Study

In this Section, the simulation study is conducted to see the performance of the maximum likelihood estimators of MGEP distribution. The simulation study have been performed by drawing random samples of sizes $10, 25, 50$ and 100 from the MGEP distribution for $k = 1.5, a = 0.5, 1, 2$ and 3. For each sample size, the maximum likelihood estimators are obtained and the procedure is repeated for 10000 times. We have computed average for the estimated value and MSE of parameter estimates for these 10000 values and the results are shown in Table 2.

Table 2: The estimated value and *MSE* of parameter estimates of MGEP for different values of a and sample sizes.

n	$a = 0.5$		$a = 1$		$a = 2$		$a = 3$	
	$\hat{a}$	MSE	$\hat{a}$	MSE	$\hat{a}$	MSE	$\hat{a}$	MSE
10	0.5088	0.026588	1.0200	0.109299	2.0441	0.434601	3.0580	0.97919
25	0.5039	0.010316	1.0067	0.040358	2.0144	0.162820	3.0234	0.365272
50	0.5018	0.004981	1.0042	0.019921	2.0084	0.079207	3.0126	0.178266
100	0.5006	0.002435	1.0026	0.009815	2.0042	0.039656	3.0049	0.088807

It can be noticed that from Table 2, that the estimated values of the parameter a are very close to the true values, and all *MSE*s decrease as the sample size increases, while they increase with increasing of the true parameter.

6. Application of MG Exponentiated Pareto distribution

In this section, we provide an application of MGEP distribution. Therefore, in order to test MGEP's goodness of fit, it has been fitted to a real lifetime data set. Moreover, MGEP has been compared with some related distributions involving Pareto, exponential Pareto, exponential, and MG exponential distribution.

The data set reported by Gross and Clark [9] represents the relief times (in minutes) of 20 patients receiving an analgesic. This dataset was analyzed by Shanker et al [10] and Kumar et at [8].
For the purpose of the analysis we set $\ln k = \exp[\min(x)]$. In order to compare MGEP with Pareto, exponential Pareto, exponential, and MG exponential distribution we use some different comparison measures includes $-2\times$ log-likelihood $(-2log(L))$, Akaike's information criterion (AIC), Corrected Akaike's information criterion $(AICC)$, Schwarz's Bayesian information criterion(BIC) and Kolmogorov-Smirnov (ks) test. Table 3 shows the results.

Table 3: Comparison criteria for the relief times data set

Distribution	$-2log(L)$	AIC	$AICC$	BIC	ks
Pareto	59.4755	63.4755	64.1814	65.4669	0.3922
Exponentiated Pareto	46.6206	50.6206	51.3265	52.612	0.2986
Exponential	65.2509	67.2509	67.4731	68.2466	0.4218
MG exponential	52.4501	54.4501	54.6723	55.4458	0.2386
MG exponential Pareto	**45.2972**	**49.2972**	**50.003**	**51.2886**	**0.2178**

From table 3, concerning the relief times of 20 patients receiving an analgesic data set, we observe that the calculated values of the five comparison criteria (the smaller the better) reveal that the MGEP distribution is the most appropriate model.

7. Conclusions

In this article, a new generalization of the exponentiated Pareto distribution named as the MG Exponentiated Pareto (MGEP) distribution is introduced and studied. Furthermore, some properties of the MGEP including survival and hazard functions, quantile function, mean, variance and random number generation are derived. The estimation of the distribution parameters are conducted using maximum Likelihood method and the simulation study is used to evaluate them. In order to test a goodness of fit for MGEP, the distribution is fitted to a real lifetime data set and compared with some related distributions. It is observed that the new distribution works better than these distributions.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article

8. References

[1] R. C. Gupta, R. D. Gupta and P. L. Gupta. Modeling Failure Time Data by Lehman Alternatives. *Commun. Statist.-Theory, Meth.*, **27(4)**, 887-904, 1998.

[2] W. T. Shaw and I. R. C. Buckley. *The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map*, Conference on Computational Finance, IMA, 0901-0434, Research Report, 2009.

[3] D. C. T. Granzotto, F. Louzada and N. Balakrishnan. *Cubic rank transmuted distributions: inferential issues and applications*. Journal of Statistical Computation and Simulation, **87(14)**, 2760-2778 (2017).

[4] Md.M. Rahman, B. Al-Zahrani and MQ. Shahbaz. A general transmuted family of distributions, *Pak J Stat Oper Res*, **14**, 451-469, 2018.

[5] K. A. AL-Kadim. Proposed Generalized Formula for Transmuted Distribution, *Journal of Babylon University, Pure and Applied Sciences*, **26(4)**, 66-74, 2018.

[6] D. Kumar, U. Singh, and S. K. Singh. A Method of Proposing New Distribution and its Application to Bladder Cancer Patients Data, *Journal of Statistics Applications & Probability Letters*, **2 (3)**, 235-245, 2015.

[7] D. Kumar, U. Singh, and S. K. Singh. A New Distribution Using Sine Function- Its Application to Bladder Cancer Patients Data, *Journal of Statistics Applications & Probability*, **4 (3)**, 417- 427, 2015.

[8] D. Kumar, U. Singh, and S. k. Singh. Life Time Distributions: Derived from some Minimum Guarantee Distribution, *Sohag Journal of Mathematics*, **4(1)**, 7-11, 2017.
[9] A.J. Gross. and V. A. Clark. *Survival Distributions: Reliability Applications in the Biometrical Sciences*,John Wiley, New York, USA, 1975.
[10] R. Shanker, H. Fesshaye, and S. Selvaraj. On Modeling of Lifetimes Data Using Exponential and Lindley Distributions, *Biometrics and Biostatistics International Journal*, **2 (5)**, 1-9, 2015.

Hussein Yousif Eledum is Associate Professor of Applied Statistics at department of Statistics at University of Tabuk, Tabuk, KSA. His permanent address is in Sudan, Shendi University, faculty of Science and Technology, department of Applied Statistics. He received his MSc. and Ph.D. in 2000 and 2005 respectively from Al-Mustansiriya University, Iraq, Sudan University of Science and Technology, Sudan. His research areas of interest are linear and nonlinear Regression, Ridge Regression, Statistical Simulation and Applied Probability.

Saiful Islam Ansari received his M.Phil. and Ph.D. in 2008 and 2011 respectively from Aligarh Muslim University, Aligarh, India. Currently he is working as Associate Professor in the Department of Business Administration, Azad Institute of Engineering and Technology, Lucknow, India. His research areas of interest are: Stochastic Programming, Reliability Optimization, Distribution Theory, Statistical computing etc.