The Kumaraswamy-Generalized Exponentiated Pareto Distribution

Dr. Tarek M. Shams
Assistant Professor of Statistics, Department of Mathematical Statistics
Faculty of Commerce, Al-Azhar University

Abstract

Based on the Kumaraswamy distribution Jones[12], we study the so-called Kum-generalized Exponentiated Pareto distribution that is capable of modeling bathtub-shaped hazard rate functions. For the first time the Kum-GEP distribution is introduced and studied. This distribution can have a decreasing and upside-down bathtub failure rate function depending on the value of its parameters; it's including some special sub-model like exponentiated Pareto Distribution and its original form. Some structural properties of the proposed distribution are studied including explicit expressions for the moments. We provide the density function of the order statistics and obtain their moments. The method of maximum likelihood is used for estimating the model parameters and the observed information matrix is derived. The real data is provided to illustrate the theoretical results in the complete data.

Key Words and Phrases: Hazard function, Kumaraswamy distribution, Moment, Maximum likelihood estimation, Exponentiated Pareto distribution.

Introduction

The Pareto distribution is the most popular model for analyzing skewed data. The Pareto distribution was first proposed by Pareto [15] as a model for the distribution of income. It can be used to represent various other forms of distributions (other than income data) that arise in human life. It has played a very important role in the investigation of city population sizes, occurrence of natural resources, insurance-risk, and business failures. Arnald et el [1]gives an extensive historical survey of its use in the context of income distribution.

The cumulative distribution function (cdf) of the two parameter exponentiated Pareto distribution is:

$$F(x;\theta,\lambda) = \left[1-(1+x)^{-\lambda}\right]^{\theta} \quad , \quad x>0 \quad ,(\theta,\lambda)>0, \tag{1}$$

a random variable X is said to follow the Pareto distribution with four parameters, if the probability density function (pdf) of X is as follows:

$$f(x;\theta,\lambda) = \theta\,\lambda\left[1-(1+x)^{-\lambda}\right]^{\theta-1}(1+x)^{-(\lambda+1)} \quad ,x>0 \quad ,(\theta,\lambda)>0 \tag{2}$$

where θ *and* λ are the shape parameters.

The generalized Pareto distribution was introduced by Pickands[16] and has since been applied to a number of areas including socio-economic phenomena, physical and biological processes Saksena et el [18], reliability studies and the analysis of environmental extremes. Davison et el [5] pointed out that the GP distribution might form the basis of a broad modeling approach to high-level exceedances. DuMouchel [8] applied it to estimate the stable index α to measure tail thickness, whereas Davison [6], [7] modeled contamination due to long-range atmospheric transport of radio nuclides. Van Montfort et el [25], [26] and van Montfort et el [24] applied the GP distribution to model the peaks over a threshold (POT) stream flows and rainfall series, and Smith[20,21] and[22] applied it to develop a POT model for flood frequencies and wave heights. Similarly, Joe [10] employed it to estimate quintiles of the maximum of N observations. Wang [27] applied it to develop a POT model for flood peaks with Poisson arrival time, whereas Rosbjerg et al [17] compared the use of the 2-parameters GP and exponential distributions as distribution models for exceedances with the parent distribution being a generalized GP distribution. In an extreme value analysis of the flow of Buebage Brook, Barrett [2] used the GP distribution to model the Pot flood series with Poisson inter-arrival times. Davison et el[7] presented a comprehensive analysis of the extremes of data by use of the GP distribution for modeling the sizes and occurrences of exceedances over high thresholds, Abdul Fattah et el.introduced the new model of generalized Pareto distribution.

In this context, we propose an extension of the exponentiated Pareto distribution based on the family of Kumaraswamy generalized denoted with the prefix "Kw-G" for short distributions introduced by Cordeiro and de Castro [4]. Nadarajah et al. [14] studied some mathematical properties of this family. The Kumaraswamy (Kw) distribution is not very common among statisticians and has been little explored in the literature. Its cdf (for $0 < x < 1$) is $F(x) = 1 - (1 - x^a)^b$, where $a > 0$ and $b > 0$ are shape parameters, and the density function has a simple form$f(x) = ab\, x^{a-1}(1 - x^a)^{b-1}$, which can be unimodal, increasing, decreasing or constant, depending on the parameter values. It does not seem to be very familiar to statisticians and has not been investigated systematically in much detail before, nor has its relative interchangeability with the beta distribution been widely appreciated. However, in a very recent paper, Jones [11] explored the background and genesis of this distribution and, more importantly, made clear some similarities and differences between the beta and Kw distributions.

In this note, we combine the works of Kumaraswamy [12] and Shawky et el [19], to derive some mathematical properties of a new model, called the Kumaraswamy Generalized exponentiated Pareto (Kw-GEP) distribution, which stems from the following general construction: if G denotes the baseline cumulative function of a random variable, then a generalized class of distributions can be defined by:

$$F(x;a,b) = 1-(1-G(x)^{a})^{b} \tag{3}$$

where $a > 0$ and $b > 0$ are two additional shape parameters which govern skewness and tail weights. Because of its tractable distribution function (2), the Kw-G distribution can be used quite effectively even if the data are censored.

Correspondingly, its density function is distributions has a very simple form:

$$f(x; a, b) = ab\, g(x)\, G(x)^{a-1}(1 - G(x)^{a})^{b-1} \qquad (4)$$

The density family (3) has many of the same properties of the class of beta-G distributions see Eugene et al [10] , but has some advantages in terms of tractability, since it does not involve any special function such as the beta function.

Equivalently, as occurs with the beta-G family of distributions, special Kw-G distributions can be generated as follows: the Kw-normal distribution is obtained by taking G(x) in (2) to be the normal cumulative function. Analogously, the Kw-Weibull Cordeiro et al [5] Kw-generalized gamma Pascoa et al [21] , Kw-Birnbaum-Saunders Saulo et al. [23] andKw-Gumbel Cordeiro et al [3] distributions are obtained by taking G(x) to be the cdf ofthe Weibull, generalized gamma, Birnbaum-Saunders and Gumbel distributions, respectively, among several others. Hence, each new Kw-

G distribution can be generated from a specified G distribution.

This paper is outlined as follows. In section 2, we define the KW-GEP distribution and provide expansions for its cumulative and density functions. A range of mathematical properties of this distribution is considered in sections 3 and4. These include quantile function, simulation, skewness and kurtosis. Maximum likelihood estimation is performed and the observed information matrix is determined in section 5. In section 6, we provide simulation study for the generated data. Finally, some conclusions are addressed.

2- The Kumaraswamy-Generalized Pareto Distribution

If $G(x;\underline{\theta})$ is the exponentiated Pareto cumulative distribution with parameter $\underline{\vartheta} = (\theta, \lambda)$ then equation (1) yields the Kw-GEP cumulative distribution for (x$\geq$ 0)

$$F\left(x;\underline{\xi}\right) = 1 - \left\{1 - \left[1 - (1 + x)^{-\lambda}\right]^{\theta a}\right\}^{b} \quad (5)$$

where $\underline{\xi} = (a, b, \theta, \lambda)$, $a, b, \theta, \lambda > 0$ are non-negative shape Parameter. The corresponding pdf and hazard rate function are:

$$f\left(x;\underline{\xi}\right) = ab\theta\,\lambda\left[1 - (1 + x)^{-\lambda}\right]^{\theta a - 1}(1 + x)^{-(\lambda+1)}\left\{1 - \left[1 - (1 + x)^{-\lambda}\right]^{\theta a}\right\}^{b-1}$$

(6)

and

$$S\left(x;\underline{\xi}\right) = 1 - F\left(x;\underline{\xi}\right) = \left\{1 - \left[1 - (1+x)^{-\lambda}\right]^{\theta a}\right\}^{b}$$

$$H\left(x;\underline{\xi}\right) = \frac{f\left(x;\underline{\xi}\right)}{S\left(x;\underline{\xi}\right)}$$

$$= \frac{ab\theta\,\lambda\left[1-(1+x)^{-\lambda}\right]^{\theta-1}(1+x)^{-(\lambda+1)}\left[\left[1-(1+x)^{-\lambda}\right]^{\theta}\right]^{a-1}}{\{1-[1-(1+x)^{-\lambda}]^{\theta a}\}}$$

respectively

2.1- Special Distributions

The following well-known and new distributions are special sub-models of the Kum-GP distribution.

If $b = 1$ in (6) we get the Kum-GEP distribution reduces to

$$f\left(x;\underline{\xi}\right) = ab\theta\,\lambda\left[1-(1+x)^{-\lambda}\right]^{\theta-1}(1+x)^{-(\lambda+1)}\left[\left[1-(1+x)^{-\lambda}\right]^{\theta}\right]^{a-1}$$

which is the exponentiated exponentiated Pareto (EGP) For $a = b = 1$, we obtain the exponentiated Pareto distribution, for $a = b = \theta = 1$, we obtain the Pareto distribution.

2.2- Expansions for the cumulative and density functions

Here, we give simple expansions for the Kw-GEP cumulative distribution. By using the generalized binomial theorem (for $0 < a < 1$)

$$(1+a)^{v} = \sum_{i=0}^{\infty} \binom{v}{i} a^{i} \tag{7}$$

in equation (5), we can write:

$$F\left(x;\underline{\xi}\right) = 1 - \sum_{i=0}^{\infty} (-1)^{i} \binom{b}{i} \left[1-(1+x)^{-\lambda}\right]^{\theta a i} = 1 - \sum_{i=0}^{\infty} \eta_i \tau\left(x;\underline{\zeta}\right)$$

where $\eta_i = (-1)^{i}\binom{b}{i}$ and $\tau\left(x;\underline{\zeta}\right)$ denotes the EP cumulative distribution with parameters $\underline{\zeta} = (\theta, ai, \lambda)$.

Now, using the power series (7) in the last term of (6), we obtain:

$$f\left(x;\underline{\xi}\right) = \frac{b\lambda\theta a(i+1)}{(i+1)} \sum_{i=0}^{\infty} (-1)^{i} \binom{b-1}{i} (1+x)^{-(\lambda+1)} \left[1-(1+x)^{-\lambda}\right]^{\theta a(i+1)-1}$$

we can write

$$f\left(x;\underline{\xi}\right) = \sum_{i=0}^{\infty} k_i g(x;\underline{\vartheta}) \tag{8}$$

where:

$$k_i = \frac{b}{(i+1)} \sum_{i=0}^{\infty} (-1)^i \binom{b-1}{i}$$

and

$g\left(x;\underline{\vartheta}\right)$, denotes the Exponentiated Pareto Distribution with parameters

$\underline{\vartheta} = (\theta, a(i+1), \lambda)$.

Thus, the Kw-GEP density function can be expressed as an infinite linear combination of Pareto densities. Thus, some of its mathematical properties can be obtained directly from those properties of the Exponentiated Pareto distribution. For example, the ordinary, inverse and factorial moments, moment generating function (mgf) and characteristic function of the Kw-GEP distribution follow immediately from those quantities of the Pareto distribution.

3- Quantile function and simulation

We present a method for simulating from the Kw-GEP distribution (6). The quantile function corresponding to (5) is:

$$Q(u) = F^{-1}(u) = \left\{1 - \left(1 - (1-u)^{\frac{1}{b}}\right)^{\frac{1}{a\theta}}\right\}^{\frac{-1}{\lambda}} - 1 \qquad (9)$$

Simulating the Kw-GEP random variable is straight forward. Let U be a uniform variate on the unit interval $(0, 1)$. Thus, by means of the inverse transformation method, we consider the random variable X given by:

$$X = \left\{1 - \left(1 - (1-u)^{\frac{1}{b}}\right)^{\frac{1}{a\theta}}\right\}^{\frac{-1}{\lambda}} - 1$$

which KW, i.e. $X \sim KW - GEP\ (a, b, \theta, \lambda)$

4- Skewness and Kurtosis

The short comings of the classical kurtosis measure are well-known. There are many heavy tailed distributions for which this measure is infinite. So, it becomes uninformative precisely when it needs to be. Indeed, our motivation to use quantile-based measures stemmed from the non-existence of classical kurtosis for many of the Kw distributions.

The Bowley's skewness (see Kenney and Keeping [15]) is based on quartiles:

$$S_k = \frac{Q_{3/4} - 2Q_{1/2} + Q_{1/4}}{Q_{3/4} - Q_{1/4}}$$

and the Moors' kurtosis (see Moors (28)) is based on cortiles:

$$K_u = \frac{Q_{7/8} - Q_{5/8} - Q_{3/8} + Q_{1/8}}{Q_{6/8} - Q_{2/8}}$$

where $Q(\cdot)$ represents the quantile function

5- Estimation and information matrix

In this section, we discuss maximum likelihood estimation and inference for the Kw-GEP distribution. Let $X_1, X_2, \dots, X_n$ be a random sample from $X \sim KW - GEP\left(\underline{\xi}\right)$ where $\underline{\xi} = (a, b, \theta, \lambda)$ be the vector of the model parameters, the log-likelihood function for $\underline{\xi}$ reduces to:

$$\ell\left(\underline{\xi}\right) = n \log a + n \log b + n \log \theta + n \log \lambda$$

$$-(\lambda+1)\sum_{i=1}^{\infty} \log(1+x_i) + (\theta a - 1)\sum_{i=1}^{\infty} \log\left(1-(1+x_i)^{-\lambda}\right)$$

$$+(b-1)\sum_{i=1}^{\infty} \log\left\{1-\left[1-(1+x_i)^{-\lambda}\right]^{\theta a}\right\} \qquad (10)$$

The score vector $U\left(\underline{\xi}\right) = (\partial\ell/\partial a,\ \partial\ell/\partial b,\ \partial\ell/\partial\theta,\ \partial\ell/\partial\lambda)^T$, where the components corresponding to the parameters in $\underline{\xi}$ are given by differentiating (10). By setting $z_i = 1-(1+x_i)^{-\lambda}$

$$\frac{\partial\ell}{\partial a} = \frac{n}{a} + \theta\sum_{i=1}^{n} \log z_i - \theta(b-1)\sum_{i=1}^{n} \frac{{z_i}^{\theta a} \log z_i}{1-{z_i}^{\theta a}}$$

$$\frac{\partial \ell}{\partial b} = \frac{n}{b} + \sum_{i=1}^{n} \log\left(1 - {z_i}^{\theta a}\right)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{n}{\theta} + a\sum_{i=1}^{n} \log z_i - a(b-1)\sum_{i=1}^{n} \frac{{z_i}^{\theta a} \log z_i}{1 - {z_i}^{\theta a}}$$

And

$$\frac{\partial \ell}{\partial \lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} \log(1+x) - (\theta a - 1)\sum_{i=1}^{n} \frac{(1+x_i)^{-\lambda} log(1+x_i)}{z_i}$$

$$-\theta a(b-1)\sum_{i=1}^{n} \frac{{z_i}^{\theta a}(1+x_i)^{-\lambda} log(1+x_i)}{1 - {z_i}^{\theta a}}$$

The maximum likelihood estimates (MLEs) of the parameters are the solutions of the nonlinear equations $\nabla\ell = 0$, which are solved iteratively. The observed information matrix given

$$J_n\left(\underline{\xi}\right) = n\begin{pmatrix} J_{aa} & J_{ab} & J_{a\theta} & J_{a\lambda} \\ J_{ba} & J_{bb} & J_{b\theta} & J_{b\lambda} \\ J_{\theta a} & J_{\theta b} & J_{\theta\theta} & J_{\theta\lambda} \\ J_{\lambda a} & J_{\lambda b} & J_{\lambda\theta} & J_{\lambda\lambda} \end{pmatrix}$$

whose elements are,

$$J_{aa} = -\frac{n}{a^2} - \theta(b-1)\sum_{i=1}^{n} \frac{\theta z_i^{\theta a} \log^2 z_i}{(1 - {z_i}^{\theta a})^2}$$

$$J_{ab} = -\theta \sum_{i=1}^{n} \frac{z_i^{\theta a} \log z_i}{1 - z_i{}^{\theta a}}$$

$$J_{a\theta} = \sum_{i=1}^{n} \log z_i - (b-1) \sum_{i=1}^{n} \frac{z_i^{\theta a} \log z_i \left(a\theta \log z_i - z_i^{\theta a} + 1\right)}{(1 - z_i{}^{\theta a})^2}$$

$$J_{a\lambda} = \theta \sum_{i=1}^{n} \frac{(1+x_i)^{-\lambda} log(1+x_i)}{z_i} - \theta(b-1) \sum_{i=1}^{n} \frac{z_i^{\theta a-1}(1+x_i)^{-\lambda} log(1+x_i)\left[1 + a\theta \log z_i - z_i^{\theta a}\right]}{(1 - z_i{}^{\theta a})^2}$$

$$J_{bb} = -\frac{n}{b^2} J_{b\theta} = -a \sum_{i=1}^{n} \frac{z_i^{\theta a} \log z_i}{1 - z_i{}^{\theta a}}$$

$$J_{b\lambda} = -\theta a \sum_{i=1}^{n} \frac{z_i^{\theta a-1}(1+x_i)^{-\lambda} log(1+x_i)}{1 - z_i{}^{\theta a}}$$

$$J_{\theta\theta} = -\frac{n}{\theta^2} - a(b-1) \sum_{i=1}^{n} \frac{a z_i^{\theta a} \log^2 z_i}{(1 - z_i{}^{\theta a})^2}$$

$$J_{\theta\lambda} = a \sum_{i=1}^{n} \frac{(1+x_i)^{-\lambda} log(1+x_i)}{z_i} - a(b-1) \sum_{i=1}^{n} \frac{z_i^{\theta a-1}(1+x_i)^{-\lambda} log(1+x_i)\left[1 + \theta a - z_i^{\theta a}\right]}{(1 - z_i{}^{\theta a})^2}$$

$$J_{\lambda\lambda} = -\frac{n}{\lambda^2} + (\theta a - 1)\sum_{i=1}^{n}\frac{(1+x_i)^{-\lambda}log^2(1+x_i)}{z_i{}^2}$$

$$-\theta a(b-1)\sum_{i=1}^{n}\frac{z_i^{\theta a}(1+x_i)^{-\lambda}\,log^2(1+x_i)\left[z_i^{\theta a}+\theta a\frac{(1+x_i)^{-\lambda}}{z_i}-1\right]}{[(1-z_i{}^{-\alpha})^{-(a-1)}-(1-z_i{}^{-\alpha})]^2}$$

6 -Application

Here, we use a real data set to compare the fits of the Kum-GEP distribution and those of other sub-models, i.e., the Exponentiated Pareto (EP) and Pareto distributions. We make a results comparison of the models fit. We consider an uncensored data set corresponding an uncensored data set from Nichols and Padgett (2006) consisting of 100 observations on breaking stress of carbon fibers (in Gba): 3.7, 2.74, 2.73, 2.5, 3.6, 3.11, 3.27, 2.87, 1.47, 3.11,4.42, 2.41, 3.19, 3.22, 1.69, 3.28, 3.09, 1.87, 3.15, 4.9, 3.75, 2.43, 2.95, 2.97, 3.39, 2.96, 2.53,2.67, 2.93, 3.22, 3.39, 2.81, 4.2, 3.33, 2.55, 3.31, 3.31, 2.85, 2.56, 3.56, 3.15, 2.35, 2.55, 2.59,2.38, 2.81, 2.77, 2.17, 2.83, 1.92, 1.41, 3.68, 2.97, 1.36, 0.98, 2.76, 4.91, 3.68, 1.84, 1.59, 3.19,1.57, 0.81, 5.56, 1.73, 1.59, 2, 1.22, 1.12, 1.71, 2.17, 1.17, 5.08, 2.48, 1.18, 3.51, 2.17, 1.69,1.25, 4.38, 1.84, 0.39, 3.68, 2.48, 0.85, 1.61, 2.79, 4.7, 2.03, 1.8, 1.57, 1.08, 2.03, 1.61, 2.12,1.89, 2.88, 2.82, 2.05, 3.65 .These data are used here only for illustrative purposes. The required numerical evaluations are carried out using the Package of Mathcad software.

Table 1 provide the MLEs (with corresponding standard errors in parentheses) of the model parameters. The model selection is carried out using the AIC (Akaike information criterion), the BIC (Bayesian information criterion) and the CAIC (consistent Akaike information criteria):

$$AIC = -2\ell(\hat{\theta}) + 2q,$$

$$BIC = -2\ell(\hat{\theta}) + q\log(n),\; CAIC = -2\ell(\hat{\theta}) + \frac{2qn}{n-q-1}$$

where $\ell(\hat{\theta})$ denotes the log-likelihood function evaluated at the maximum likelihood estimates, q is the number of parameters, and n is the sample size.

Table1. MLEs of the model parameters, the corresponding SEs (given in parentheses) and the statistics AIC, BIC and CAIC

Model	Estimates				Statistic		
	$\hat{a}$	$\hat{b}$	$\hat{\theta}$	$\hat{\lambda}$	AIC	BIC	$CAIC$
$Kw-GE$	1	4.368	4.134	3.142	257.762	257.762	257.665
EP	1.519	---	6.041	4.361	245.137	245.137	256.04
P			7.423	3.985	276.106	267.106	258.009

Since the values of the AIC, BIC and CAIC are smaller for the Kum-GP distribution compared with those values of the other models, the new distribution seems to be a very competitive model to these data.

7- Concluding Remarks

The well-known two-parameter exponentiated Pareto distribution, introduced by Shawky, Hanaa, Abu-Zinadah, (2009), is extended by introducing two extra shape parameters, thus defining the KW-G exponentiated Pareto (KW-GEP) distribution having a broader class of hazard rate and density functions. This is achieved by taking (1) as the base line cumulative distribution of the generalized class of KW-G distributions defined by Cordeiro and de Castro (2010). A detailed study on the mathematical properties of the new distribution is presented. The new model includes as special sub-models the Pareto, exponentiated Pareto (EP) (Gupta et al., 1998) and Pareto distributions. We obtain the quantile function, skewness and kurtosis. The estimation of the model parameters is approached by maximum likelihood and the observed information matrix is obtained. An application to a real data set indicates that the fit of the new model is superior to the fits of its principal sub-models. We hope that the proposed model may be interesting for a wider range of statistical research.

References

(1) Arnald, B.C. and Press, S.J (1983), "Bayesian inference for Pareto populations" Journal of Econometrics, 21, 287-306.

(2) Barrett, J. H. (1992), An extreme value analysis of the flow of Burbage Brook. Stochastic Hydrol. Hydraul.6, 151-165.

(3) Cordeiro G. M. and de Castro, M. (2011), A new family of generalized distributions, Journal of Statistical Computation and Simulation, 81, 883–898.

(4) Cordeiro, G.M., Ortega, E.M.M. and Nadarajah, S. (2010), The Kumaraswamy Weibull distribution with application to failure data, Journal of the Franklin Institute, 347, 1399–1429.

(5) Davison, A. C. & Smith, R. L. (1990), Models for exceedances over high thresholds. /. Roy. Statist. Soc. B 52(3), 393-442.

(6) Davison, A. C. (1984a), Modeling excesses over high thresholds, with an application. In: Statistical Extremes and Applications, ed. J. Tiago de Oliveira, 461-482. Reidel, Dordrecht, the Netherlands.

(7) Davison, A. C. (1984b), A statistical model for contamination due to long-range atmospheric transport of radionuclide's. PhD thesis, Department of Mathematics, Imperial College of Science and Technology, London, UK.

(8) DuMouchel, W. (1983), Estimating the stable index α in order to measure tail thickness. Ann. Statist. 11,1019-1036.
(9) Eugene, N., Lee, C. and Famoye, F. (2002), Beta-normal distribution and its applications. *Communications in Statistics*: Theory andMethods 31, 497–512.
(10) Joe, H. (1987), Estimation of quantile-s of the maximum of N observations. Biometrika 74, 347-354.
(11) Jones, M. C. (2009), Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*, 6, 70-81.
(12) Kumaraswamy, P. (1980), A generalized probability density function for double-bounded random processes. Journal of Hydrology, 46, 79-88.
(13) Moors, J.J. (1998), A quantile alternative for kurtosis. Journal of the Royal Statistical Society D, 37, 25–32.
(14) Nadarajah, S., Cordeiro, G.M. and Ortega, E.M.M. (2011), General results for the Kumaraswamy-G distribution. Journal of Statistical Computation and Simulation. DOI:10.1080/00949655.2011.562504.
(15) Pareto, V. (1986), “cours economic politique” , Lausanne and Paris; Range and Cie.
(16) Pickands, J. (1975), Statistical inference using extreme order statistics. Ann. Statist. 3, 119-131.
(17) Rosbjerg, D., Madsen, H. & Rasmussen, P. F. (1992), Prediction in partial duration series with generalized

Pareto-distributed exceedances. Wat. Resour. Res. 28(11), 3001-3010.

(18) Saksena, S. K. & Johnson, A. M. (1984). Best unbiased estimators for the parameters of a two-parameter Pareto distribution. Biometrika 31, 77-83.

(19) Shawky A.I., Hanaa H. and Abu-Zinadah. (2009), Exponentiated Pareto Distribution: Different Method of Estimations" Int, J. Contemp. Math. Sciences, Vol. 4, 2009, No. 14, 677 - 693.

(20) Smith, J. A. (1991), Estimating the upper tail of flood frequency distributions. Wat. Resour. Res. 23(8),1657-1666.

(21) Smith, R. L. (1984), Threshold methods for sample extremes, In: Statistical Extremes and Applications ed.J. Trago de Oliveira, 621-638. Reidel, Dordrecht, the Netherlands.

(22) Smith, R. L. (1987), Estimating tails of probability distributions. Ann. Statist., 15, 1174-1207.

(23) Smith, R. L. (1989), "Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone", Statistical Science, Vol. 4, pp.367-393.

(24) Van Montfort, M. A. J. & Otten, A. (1991), The first and the second e of the extreme value distribution, EV1. Stochastic Hydrol. Hydraul. 5, 69-76.

(25) Van Montfort, M. A. J. & Witter, J. V. (1985), Testing exponentiality against generalized Pareto distribution. Hydrol. 78, 305-315.

(26) Van Montfort, M. A. J. & Witter, J. V. (1986) The generalized pareto distribution applied to rainfall depths. Hydrol. Sci. J. 31(2), 151-162.

(27) Wang, Q. J. (1991), The POT model described by the generalized Pareto distribution with Poisson arrival rate. X Hydro. 129, 263-280.

توزيع كومارسوامى المعمم لتوزيع باريتو المعمم

دكتور/ طارق محمد شمس الدين(*)

ملخص البحث:

قدم Cordeiro and de Castro(2010) لأول مرة عائلة جديدة من التوزيعات تسمي توزيعات كومارسوامى المعممة وأوضح أن هذه التوزيعات هي إمتداد جديد للتوزيع الطبيعي وتوزيع وايبل وجاما وجامبل وتوزيع جاوس العكسي وأوضح أن هذه التوزيعات لها مميزات عن توزيعات عائلة بيتا المعممة التي قدمها Eugene, et al. (2002) فعلى سبيل المثال إنها لا تحتوي على دوال خاصة في توزيعاتها مثل عائلة بيتا المعممة وكذلك يمكن حساب خواصها بسهولة وبشكل أبسط من نظيرتها في توزيعات عائلة بيتا المعممة وقد أعطى Cordeiro and de Castro(2010) الشكل العام لعزوم عائلة توزيعات كومارسوامى المعممة وكذلك الشكل العام للإحصاءات الترتيبية والعزوم الخطية وإستخدام طريقة الإمكان الأكبر لتقدير معالم عائلة كومارسوامى المعممة.

إن الميزة الرئيسية من استخدام عائلة توزيعات كومارسوامي المعممة هي إتباع فكرة عائلة توزيعات بيتا المعممة مع الاستفادة من بساطة توزيع كومارسوامي وخلوه من الدوال الخاصة، لذلك كان الهدف الرئيس من هذا البحث هو حساب الخواص الرئيسية لعائلة كومارسوامي المعممة، واستخدام طريقة تقدير دالة الإمكان الأكبر لتقدير معالم التوزيع وتطبيق ذلك على توزيع باريتو المعمم الذي قدمه Shawky, Hanaa, Abu-Zinadah. (2009)

(*) الأستاذ المساعد بقسم الإحصاء، كلية التجارة بنين جامعة الأزهر.

يتكون هذا البحث من عدة أقسام وتشمل:

القسم الأول يحتوى على دراسة مرجعية لتوزيع كومارسوامى وتوزيع باريتو المعمم المشار عنه سابقاً والقسم الثاني يشتمل على توزيع كومارسوامى مطبقاً على توزيع باريتو المعمم، الخصائص الرياضية لهذا التوزيع مقدمه في القسم الثالث والرابع ومن ضمنها المحاكاة، الدالة المولدة للبيانات العشوائية و الالتواء والتفرطح. وبالنسبة لدالة الإمكان الأكبر لتقدير معالم النموذج المقترح وأيضاً مصفوفة التباينات والتغايرات لمقدرات هذه المعالم تم عرضهم في القسم الخامس وأخيراً تم عرض بعض الاستنتاجات الإحصائية على نتائج التقدير والمراجع المستخدمة.