Estimating parameters Gumbel Pareto Distribution

Mahdi Wahhab Namah Nasrallah

Department of Statistics - College of Administration and Economic - University of Karbala- Iraq

Mahdi_na2002@yahoo.com

Received: 27 March 2017 **Accepted: 26 September 2017**

Abstract

The proposed method for generating a new distribution, depends on the Cumulative Distribution Function (CDF) of two distributions, namely, Gumbel distribution and pareto distribution.We obtain a new compound, which is called (Gumble – Pareto distribution GPD). In this research we work on deriving the formula for the new distribution, and all other additional properties as well as introducing different methods to estimate the four parameters ($\mu^*, \theta , \alpha , B$) By different Method like Maximum likelihood, and method of Moments estimator and also, we derive Percentiles estimator and least squares and also weighted least square. Then the Comparison is done through simulation.

Keywords: Gumbel distribution, Pareto distribution, MLE, Moments, LS, WLS, PE.

DIYALA JOURNAL FOR PURE SCIENCES

Estimating parameters Gumbel Pareto Distribution

Mahdi Wahhab Namah Nasrallah

تقدير معالم توزيع كامبل باريتو

مهدي وهاب نعمة نصرالله

قسم الاحصاء - كلية الاداره والاقتصاد - جامعة كربلاء - العراق

ألخلاصة

الطريقة المقترحة لتوليد توزيع جديد تعتمد على الدالة التراكمية (CDF) لتوزيعين هما توزيع كامبل وتوزيع باريتو، لذلك نحتاج الى تحويل بصيغة احتمالية جديدة من خلال بناء توزيع جديد يعتمد على التوزيعين المذكورين، نحصل على توزيع كامبل باريتو (Gumbel pareto distribution GPD). في هذا البحث نحن نعمل على اشتقاق صيغ التوزيع الجديد وخصائصه. فضلاً عن ادخال أساليب مختلفة لتقدير المعالم الأربع (θ, α, μ, B) بطرق مختلفة مثلاً الإمكان الأعظم، طريقة العزوم المقدرة، وأيضاً نستخلص النسب المئوية المقدرة ، والمربعات الصغرى، وأيضاً المربعات الصغرى الموزونة، يتم ذلك من خلال المحاكاة.

ألكلمات المفتاحية: توزيع كامبل، توزيع باريتو، الامكان الأعظم، العزوم، النسب المئوية، المربعات الصغرى، المربعات الصغرى الموزونة.

Introduction

The Gumbel Pareto is well known as probability distribution for its ability to model different types of data. and also has many applications in risk analysis and quality. Many papers on the distribution of extremes appeared in the literature. Gumbel (1958) gave detailed results on extremes value theory in his book statistics of Extremes. Furthermore, Gumbel has been agree with Johnson et al. (1995), as the first to bring attention to the possibility of using the Gumbel distribution to model extreme value of random data. Kotz and Nadarajah (2000), and Beirlant et al, (2006). Alzaatrch, Lee and Famoye (2013) proposed a method for generating new distribution, Al-Aqtash et al (2014) proposed some properties of the Gumbel – Weibull distribution the mean deviations and modes are studied. Tahir et al (2015) proposed the introduction of a new four-parameter model named the Gumbel – Lomax distribution, stand up from the Gumbel –x generator recently which was proposed by Al-Aqtash (2013).

For a proposed method for generating new distribution namely. The cumulative Gumbel distribution function CDF is distinct as:

$$F(x) = exp\left[-e^{-\frac{(x-\mu)}{B}}\right] \qquad \ldots(1)$$

Additionally, the probability density function (pdf) is distinct as:

$$f(x) = \frac{1}{B} e^{-\frac{(x-\mu)}{B}} exp\left[-e^{-\frac{(x-\mu)}{B}}\right] \quad \text{... (2)}$$

Also, x to be Pareto with pdf $f(x) = \frac{\alpha}{x}(\frac{\theta}{x})^{\alpha},\ for\ x > \theta$ ... (3)

And CDF of preto F(x)= 1- $(\frac{\theta}{x})^{\alpha}$ ... (4)

From now on, we can omit the dependence on the model parameters, the CDF and pdf of the four parameters Gumble-Pareto distribution are given by:

$$G(x) = \int_{-\infty}^{ln\frac{F(x)}{1-F(x)}} r(t)dt = R\left[\ln(\frac{F(x)}{1-F(x)})\right], \quad \text{...(5)}$$

Then

$$= exp\left[-e^{-\left[\frac{\left\{\left(Ln\left((\frac{\theta}{x})^{-\alpha}-1\right)-\mu\right\}\right.}{B}\right]}\right]$$

$$G(x) = exp\left[-\frac{\mu^*}{B}\left\{(\frac{\theta}{x})^{-\alpha} - 1\right\}^{-1}\right] \quad \text{... (6)}$$

and let $\mu^* = Lne^{\mu}$

Where $0 < \mu^* < \infty$, θ ,α , B > 0 , x > 0 .The corresponding pdf is

$$g(x) = \frac{\mu^*\theta\alpha}{x^2B} (\frac{\theta}{x})^{-(\alpha+1)} \left((\frac{\theta}{x})^{-\alpha} - 1\right)^{-2} exp\left[-\frac{\mu^*}{B}\left\{(\frac{\theta}{x})^{-\alpha} - 1\right\}^{-1}\right] \quad \text{...(7)}$$

In this article, derived some properties of GPD such as moments and additional properties including the mean deviations and modality are studied.

uch as moments and additional properties including the mean deviations and modality are studied.

Estimation of parameters

For the four parameters of GPD defined in equation (2), which are two shape parameters (α,θ) and two scale parameters (B, μ^*), these four parameters are estimated using different methods maximum likelihood method, and moments method. also we derived to obtain Central Moment and apply numerical method on the statistical measure like mean, variance.

Methods of Estimation

However, we introduce four methods of estimation for the four Parameters (θ , α ,B , μ^*). For the generated Gumbal pareto distribution. afher explaining the estimators briefly, the result of Simulation procedure also explained.

Maximum Likelihood Method

IF $X_1, X_2, \ldots\ldots, X_n$ is a r.s. from GP distribution, then the Likelihood function is:

$$\mathbf{L} = \left(\frac{\mu^*\theta\alpha}{B}\right)^n \prod_{i=1}^{n}\frac{1}{x_i^2}\ \prod_{i=1}^{n}\left(\frac{\theta}{x_i}\right)^{-(\alpha+1)}\ \prod_{i=1}^{n}\left(\left(\frac{\theta}{x_i}\right)^{-\alpha}-1\right)^{-2} e^{-\frac{\mu^*}{B}\sum_{i=1}^{n}\left(\left(\frac{\theta}{x_i}\right)^{-\alpha}-1\right)^{-1}} \quad \ldots(8)$$

So, the log likelihood function is

$$\log L = n\mathrm{Ln}\left(\frac{\mu^*\theta\alpha}{B}\right) + \sum \mathrm{Ln}\left(\frac{1}{x_i^2}\right) - (\alpha+1)\sum \mathrm{Ln}\left(\frac{\theta}{x_i}\right) - 2\sum \mathrm{Ln}\left(\left(\frac{\theta}{x_i}\right)^{-\alpha}-1\right) - \frac{\mu^*}{B}\sum \mu^* \quad \ldots(9)$$

By assuming that θ is known and derived for $\alpha,\ \mu^*$ and B , We obtain the equations as follows:

$$\frac{\partial L}{\partial \alpha} = \frac{n}{\alpha} - (\alpha+1)\sum_{i=1}^{n}\left(\frac{\theta}{x}\right)^{-(\alpha+2)} + 2\alpha\sum\left(\frac{\theta}{x}\right)^{-(\alpha+1)}\left(\left(\frac{\theta}{x}\right)^{-\alpha}-1\right)^{-3} - \frac{\mu^*\alpha}{B}\sum\left(\frac{\theta}{x}\right)^{-(\alpha+1)}\left(\left(\frac{\theta}{x}\right)^{-\alpha}-1\right)^{-2} = 0 \quad \ldots(10)$$

$$\frac{\partial L}{\partial \mu^*} = \frac{n}{\mu^*} - \frac{1}{B}\left(\left(\frac{\theta}{x_i}\right)^{-\alpha}-1\right)^{-1} = 0 \quad \ldots(11)$$

$$\frac{\partial L}{\partial B} = \frac{n}{B} - \frac{\mu^*}{B}\left(\left(\frac{\theta}{x_i}\right)^{-\alpha}-1\right)^{-1} = 0 \quad \ldots(12)$$

The above equations can be equal to zero, and then the numerical solution can be used to solve their simultaneously. Consequently, we obtain $\alpha^{\wedge}_{MLE}$ and $B^{\wedge}_{MLE}$ $\mu^{*\wedge}_{MLE}$ as M.L.E. estimators of α, B, μ^*, respectively.

DOI: http://dx.doi.org/10.24237/djps.1402.372A
P-ISSN: 2222-8373
E-ISSN: 2518-9255

Moments method estimator

The technique of moments, which is used for constructing estimator of the parameters, is based on the identification between the sample moments with the corresponding distribution moment. Here, we provide the m.o.m. estimators of α, B and μ^* Parameters of GP distribution when both are unknown. IF X follows GP (α , B , μ^* , θ) , Assuming θ is known, the r th moment of X is:

$$E\,X^r = \int_\theta^\infty X^r\ g(x)\,dx$$

$$= \frac{\mu^*\theta\alpha}{B}\int_\theta^\infty X^r\,(\frac{\theta}{x})^{-(\alpha+1)\alpha}\left[(\frac{\theta}{x})^{-\alpha}-1\right]^{-2} e^{-\frac{\mu^*}{B}}\left[(\frac{\theta}{x})^{-\alpha}-1\right]^{-1} dx \quad \ldots(13)$$

By equating the first three moments with their corresponding sample moments and solve the resulting equations numerically, we obtain the Method of moments estimators $\alpha^{\wedge}_{MME}$, $B^{\wedge}_{MME}$ and $\mu^{*\wedge}_{MME}$ for α, B, μ^* respectively.

Percentiles Estimation (PE)

Kao in (1959) initially discovered this method via the graphical approximation to the best linear unbiased estimators. The estimators might be found by fitting a straight line to the theoretical points determined from the distribution function, and the sample percentile points. In the case of a GP distribution, it is probable to use the same idea to determine the estimators of α, B, and μ^* based on PE, because of the structure of its distribution function. Since G(x) distinct in (1). First of all, we find numerically the value of x where x= G^{-1}(x, α, B, μ^*), since P_i is the estimate of G ($X_{(i)}$, α , B , μ^*) .

$\mu^{*\wedge}_{PE}$, $\alpha^{\wedge}_{PE}$, $B^{\wedge}_{PE}$ can be determined by minimizing

$\sum_{i=1}^{n}\left[x_{(i)} - G^{-1}(P_i, , \alpha , B , \mu^*)\right]^2$ with respect to α, B, μ^*, where

$E\left[G(x_{(i)}\right] = P_i = \frac{i}{n+1}$ is the most used estimator of G($X_{(i)}$).

Least squares Estimation (LSE) and Weighted least squares estimators (WLSE)

This method was initially proposed by Swain et al in 1988, to estimate the parameters of Beta distribution. However, suppose x_1, x_2, ….x_n is a r.s. of size n with distribution function G(x), uses the distribution of G($x_{(i)}$). For a sample of size (n) we have in [5].

E(G($x_{(i)}$) = $\frac{i}{n+1}$, var (G($x_{(i)}$) = j(n-j+1) / [$(n+1)^2$ (n+2)] and

Cov $[(G(x_{(i)}), (G(x_{(k)}))] = j(n-k+1) / [(n+1)^2 (n+2)]$ for $j < k$

So, one can obtain the LS estimators of α, B, μ^*say $\mu^{*\wedge}_{LSE}$, $\alpha^{\wedge}_{LSE}$, $B^{\wedge}_{LSE}$ by minimizing,

$\sum_{i=1}^{n}(G(x_{(i)} - \frac{i}{n+1})^2$ with respect to the unknown parameters α, B, μ^*.

The WLSE of α, B, μ^*say $\mu^{*\wedge}_{WLSE}$, $\alpha^{\wedge}_{WLSE}$, $B^{\wedge}_{WLSE}$ respectively, can be determined by minimizing,

$\sum_{j=1}^{n} w_j(G(x_{(i)} - \frac{i}{n+1})^2$ with respect α, B, μ^*, where

$$w_j = \frac{1}{var\left(G\left(x_{(i)}\right)\right)} = (n+1)^2 \frac{(n+2)}{j(n-j+1)}$$

Table (1) Empirical MSE to Estimate parameters of GP Distribution with Different Sample Size Methods of Estimation and Different Values of parameters α, B, μ

Table 1: Empirical MSE to Estimate parameters of GP Distribution with Different Sample Size Methods of Estimation and Different Values of parameters α, B, μ

case		1			2					
parameters		α	B	μ^*	α	B	μ^*	α	B	μ^*
Sample Size	The method	0.5	o.7	0.9	0.7	0.9	0.5	0,9	0.5	0.7
20	MLE	0.784	3.213	6.217	0.863968	5.21477	7.4604	1.06232	3.06068	6.177052
	Mom	0.915	3.662	5.464	1.00833	5.93976	6.5568	1.23982	3.48622	9.397847
	LSE	1.312	4.824	8.313	1.445824	7.82615	9.9756	1.77776	4.5934	9.180791
	WLSE	1.216	4.617	8.121	1.340032	7.48877	9.7452	1.64768	4.39538	7.282681
	PE	0.969	3.912	6.442	1.067838	6.34526	7.7304	1.31299	3.72422	0.722299
50	MLE	0.722299	2.96198	5.727722	0.795974	4.804331	6.873267	0.978715	2.819804	6.47519
	Mom	0.84299	3.373801	5.033983	0.928974	5.472305	6.04078	1.142251	3.211858	5.690918
	LSE	1.208746	4.445273	7.652767	1.332038	7.210232	9.19052	1.63785	4.231899	8.658236
	WLSE	1.120301	4.253642	7.481877	1.2345710.9	6.899407	8.978253	1.518008	4.049467	8.458262
	PE	0.89274	3.604126	5.935015	83799	5.845892	7.122018	1.209662	3.431128	6.709534
100	MLE	0.632218	2.592576	5.013389	0.696704	4.205158	6.016067	0.856655	2.468132	5.667636
	Mom	0.737256	2.953037	4.40617	0.813117	4.789826	5.287404	0.999795	2.811291	4.981175
	LSE	1.057997	3.89088	6.703603	1.165912	6.311007	8.044324	1.433586	3.704118	7.578423
	WLSE	0.980582	3.723149	6.548774	1.080602	6.038947	7.858529	1.328689	3.544438	7.403389
	PE	0.781402	3.154637	5.194829	0.861105	5.116821	6.233795	1.058799	3.003214	5.872754

We achieved extensive simulations to contrast the performances of the various methods, which were stated in section (3), mainly with respect to their MSE for various sample sizes, and for various parameters values.

The experiments were conducted according to run size R=1000. We reported the results for n=20 (small sample), n=50 (moderate sample), and n= 100 (large sample) and for the following different values of α, B, μ^* .

α	0.5	0.7	0.9
B	0.7	0.9	0.5
μ^*	0.9	0.5	0.7

The results are presented in Table 1. From the table, we observe that:

1. The MSE's decrease as sample size increases in all methods of estimation. It verifies the asymptotic unbiasedness and consistency of all the estimators.
2. It can be said that the estimation of shape parameters are more accurate for the smaller values of those parameters whereas the estimation of scale parameters are more accurate for the larger values of those parameters .in other words, MSE's increase as shape parameter increases whereas MSE's increase as scale parameter decreases.
3. The values of MLE, Mom and PE are related to their order.
4. small sample size (n=20) and moderate sample size (n=50), it is detected that MLE can be the best for the three parameters, while the second best one is Mom, and the third is PE.
5. With large sample size (n=50, 100), it is found that MLE, Mom, PE works the greatest from all other methods to estimate.

Summary and Conclusions

A new Gumbel Pareto Distribution in statistical analysis is represented a vital method. Different methods to estimate distribution parameters are studied, MLE, Mom, PE, LSE and WLSE. An empirical study was acieved to compare among these methods. It was found that the MLS is the best one for small and moderate samples. Moreover, it is the best to estimate for large samples, while the PE is the best to estimate the parameter for large samples.

References

1. Abid S (2014). The fréchet stress-strength model. International Journal of Applied Mathematical Research, 3 (3) (2014) 207-213.
2. Abid S (2016). Propertices of Doubly-Truncated Fréchet Distribution American Journal of Applied Mathematics and Statistics. Vol. 4, No. 1, 2016, pp 9-15.
3. Abid S, Hassan H (2015). Some Additive Failure Rate Models Related with MOEU Distribution. American Journal of Systems Science 2015, 4(1): 1-10.
4. Alzaatreh, A., Famoye, F. & Lee, C. (2013). Weibull-Pareto distribution and its applications, Communications in Statistics − Theory and Methods, 42, 1673- 1691.
5. Alzaatreh, A., Lee, C. & Famoye, F. (2013). A new method for generating families of continuous distributions, Metron, 71(1), 63-79.
6. Al-Aqtash, Raid; Lee, Carl; and Famoye, Felix (2014) "Gumbel-Weibull Distribution: Properties and Applications," Journal of Modern Applied Statistical Methods: Vol. 13: Iss. 2, Article 11.
7. Eugene, N., Lee, C. & Famoye, F. (2002). Beta-normal distribution and its applications. Communications in Statistics − Theory and Methods, 31(4), 497- 512.
8. Gumbel, E. J. (1958). Statistics of Extremes. New York: Columbia University Press.
9. Shawky Y.A.I. and Hanah Abu – Zinadah, (2009), "Exponentiated Pareto distribution: different methods of estimations", Int. J. Contemp.Math. Sciences, Vol. 4, No. 14, 677 – 693.
10. Raid Al-Aqtash, Carl Lee, Felix Famoye,(2014) Gumbel-Weibull Distribution: Properties and Applications, Journal of Modern Applied Statistical Methods, Volume 13 | Issue 2.
11. The_Gumbel- Lomax_ Distribution_Properties_and_Applications, (2015), https://www.researchgate.net/publication/280125764.