Global Journal of human social science
Economics
Volume 13 Issue 2 Version 1.0 Year 2013
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-460X & Print ISSN: 0975-587X

On Fitting of Generalized Pareto Distribution

By T. A. Raja & A. H. Mir
Division of Agricultural Statistics, Skuast-K Shalimar. (J&K)

Abstract - The Pareto distribution is to model the income data set of a society. The distribution is appropriate to the situations in which an equilibrium exists in distribution of small to large. There exists many generalization approaches to the distribution. In this paper an effort has been made to compare the applicability of generalized Pareto distribution with Picklands (1975) by using a real life income data set. The model has provided considerable a good fit to the data set. Some well known distributions has been derived as a special case of this model for suitable choice of parameters.

Keywords : *pareto distribution, picklands (1975) generalized pareto distribution, a new generalized Pareto distribution, Income data set, Goodness of fit.*

GJHSS-E Classification : FOR Code: 140201

ONFITTINGOFGENERALIZEDPARETODISTRIBUTION

Strictly as per the compliance and regulations of:

On Fitting of Generalized Pareto Distribution

T. A. Raja [α] & A. H. Mir [σ]

Abstract - The Pareto distribution is to model the income data set of a society. The distribution is appropriate to the situations in which an equilibrium exists in distribution of small to large. There exists many generalization approaches to the distribution. In this paper an effort has been made to compare the applicability of generalized Pareto distribution with Picklands (1975) by using a real life income data set. The model has provided considerable a good fit to the data set. Some well known distributions has been derived as a special case of this model for suitable choice of parameters.

Keywords : pareto distribution, picklands (1975) generalized pareto distribution, a new generalized Pareto distribution, Income data set, Goodness of fit.

I. Introduction

a) Pareto Distribution (PD)

The Pareto distribution was proposed by an Italian born Swiss economist named Vilfredo Pareto (1897) as a model for the distribution of income. It is a skewed, heavy tailed distribution and is some times referred as Bradford distribution. Pareto used this distribution to describe the allocation of wealth among individuals. A large portion of wealth of many societies is owned by a smaller percentage of the people in that society. This distribution s sometimes expressed more simple as the Pareto principle or The "80-20" rule which says that 20% of the population owns 80% of the wealth. This distribution is not limited to describing wealth or income distribution, but to many situations in which an equilibrium is found in the distribution of the "small" to the "large". It is widely used and has played a very important role in explaining population occurrence, natural resources, insurance risk, business failures and has recently been used to study the ozone levels in the upper atmosphere. Wingo (1982) discussed the unimodility of the conditional likelihood function of the Pareto distribution using multi censored samples. Arnold and Press (1983) gave an extensive historical survey of its use in the content of income distribution.

The probability density function (p. d. f) of two parameter Pareto distribution is defined as

$$f(x, \alpha, \beta)= \frac{\alpha}{\beta}(\alpha / \beta)^{-(\alpha+1)} \quad (1.1)$$

$\alpha, \beta>0$ and $x\geq0$

Where β is a scale parameter and α is a shape parameter.

The probability density function (p. d. f) of three parameter Pareto distribution is defined as

$$f(x, \alpha, \beta, \lambda)= \alpha/ \beta \left[1+\frac{(x-\lambda)}{\beta}\right]^{-(\alpha+1)} \quad (1.2)$$

$= \quad 0 \quad$ otherwise

where $\lambda<x< \alpha, \quad \beta>0, \alpha>0$

β is a scale parameter, α is a shape parameter and λ is the location.

b) Generalized Pareto distribution (GPD)

Like other distributions the Pareto distribution was generalized. The Generalized Pareto distribution (GPD) was introduced by Picklands (1975). The probability density function (p.d.f) is defined as

$$f(x, \alpha, \beta)= 1/ \beta \left[1-\frac{x\alpha}{\beta}\right]^{(\frac{1}{\alpha}-1)} \quad (2.1)$$

$= \quad 0 \quad$ otherwise

The range of x is $0\leq x<\alpha$ for $\alpha\leq0$ and $0\leq x\leq \beta/\alpha$ for $\alpha>0$

The GPD is heavy tailed, skewed and is used to model extreme values as investigated by Hoking and Well (1987), Smith (1989, 1990), Davison and Smith (1990). Smith (1990) gave a review of two most widely used methods based on generalized extreme value distribution. Samia and Mohammad (1993) used five

Author α σ : Division of Agricultural Statistics, Skuast-K Shalimar. (J & K). E-mail : tariqaraja@rediffmail.com

modifications of moments to estimate the parameters of Pareto distribution. Abdel Ghaly etal (1998) obtained the prediction of the shape parameters. Choulakin and Stephens (2001) obtained the goodness of fit for the generalized Pareto distribution. Abd Elfattah etal (2007) obtained a new generalized Pareto distribution and derived some well known distributions as special cases.

c) A Model of Generalized Pareto Distribution

In this paper a model of generalized Pareto distribution as given by Abd Elfattah etal (2007) by introducing one more shape parameter "γ" is applied it to real life data set regarding family income sample from Kashmir (Jammu and Kashmir)-India.

The probability density function of the new generalized Pareto distribution is as

$$f(x;\alpha,\gamma,\beta,\lambda)=\alpha/\beta\left[\left(1+\frac{(x-\lambda)}{\beta}\right)^{\gamma}\right]^{-(\alpha+1)}.\left(\frac{x-\lambda}{\beta}\right)^{\gamma-1} \quad (3.1)$$

where $\lambda<x<\alpha$, $\beta>0$, $\alpha>0$ and $\gamma>0$

α and γ are shape parameter and λ is the location and β is a scale parameter

To prove that f(x) is a probability density function, following conditions are to be satisfied.

i. $f(x)\geq 0$ and

ii. $\int_{-\infty}^{\infty} f(x).dx=1$

$\lambda<x<\alpha$, $\beta>0$, $\alpha>0$ and $\gamma>0$ are the parameters of the model.

$f(x)\geq 0$ for all x which proves (i)

clearly $f(x)\geq 0$ establishes condition (ii) as $\int_{-\infty}^{\infty} f(x).dx=1$ The rth moment about mean of the generalized Pareto distribution is

$$\mu_r=E(x-\mu)^r=\int_{\lambda}^{\infty}(x-\mu)^r.f(x;\alpha,\beta,\gamma,\lambda).dx$$

$$=\alpha\beta^r\sum_{j=0}^{r}\binom{r}{j}(-1)^j.\left[\frac{\tau(\alpha-\frac{1}{\gamma}).\tau(1+\frac{1}{\gamma})}{\tau(\alpha)}\right]^j.\left[\frac{\tau(\alpha-\frac{r-j}{\gamma}).\tau(1+\frac{r-j}{\gamma})}{\tau(\alpha+1)}\right]$$

Where

$\tau(.)$ is defined as the gamma function.

Here r can take any value r=1, 2, 3…., there fore mean and variance of x can be defined as

$$\mu=\beta\frac{\tau(\alpha-\frac{1}{\gamma}).\tau(1+\frac{1}{\gamma})}{\tau(\alpha)}+\lambda, \quad \text{and}$$

$$\sigma^2=\beta^2\left[\frac{\tau(\alpha+\frac{2}{\gamma}).\tau(1-\frac{2}{\gamma})}{\tau(\alpha)}-\left(\frac{\tau(\alpha+\frac{1}{\gamma}).\tau(1-\frac{1}{\gamma})}{\tau(\alpha)}\right)^2\right]$$

d) Derivations of Some Distributions

Many distributions can be derived from 4-parameter generalized Pareto distribution for different choices of the parameters.

(i) For γ = 1 the four-parameter Pareto distribution reduces to three-parameter Pareto distribution with p. d. f as

$$f(x,\alpha,\beta,\lambda)=\left[1+\frac{(x-\lambda)}{\beta}\right]^{-(\alpha+1)}$$

where $\lambda<x<\alpha$, $\beta>0$, $\alpha>0$

β is a scale parameter, α is a shape parameter and λ is the location

(ii) For $\gamma = 1$ and $\lambda = 0$ it reduces to two parameter Lomax distribution with p. d. f as

$$f(x, \alpha, \beta) = \alpha/\beta \left[1+\left(\frac{x}{\beta}\right)\right]^{-(\alpha+1)}$$

Where $x > 0$ and, $\beta > 0$, $\alpha > 0$

(iii) For ($\gamma = \beta = 1$ and $\lambda = 0$) it reduces to Beta type II distribution with p. d. f as

$$f(x, \alpha,) = \alpha(1+x)^{-(\alpha+1)}$$

Where $x > 0$ and, $\alpha > 0$

Similarly many more distributions can be derived for suitable choice of parameters.

e) Goodness of Fit

As Pareto distribution provides a good fit to income data a lot of work has been done on it. In this paper the new generalized Pareto distribution has been fitted to income data along with Picklands (1975) generalized Pareto distribution to three hundred families from Kashmir valley of Jammu and Kashmir-India. The sample has been selected at random and stratified random sampling procedure involving all the six districts of Kashmir valley has been adopted for the purpose.

Table 1 : Fitting of New Generalized Pareto Distribution

Class	Income (Rs) Xi	Observed frequency (Oi)	Expected Frequency (Ei) by Picklands (1975)	Expected Frequency (Ei)
1	< 10,000	202	196	186
2	10,000-20,000	65	69	74
3	20,000-30,000	18	23	25
4	30,000-40,000	9	7	13
5	40,000-50,000	4	3	1
6	50,000 and above	2	2	1
Total	-	300	300	300

The mean and standard deviation of the above data set has been found as mean=16565.75 and standard deviation= 18850.40. The Chi-square statistics for new generalized Pareto distribution referred by its p-value is ($p=0.387$) and Chi-square statistics for Picklands (1975) generalized Pareto distribution referred by its p-value is ($p=0.843$) reveals clear non-significance in both the cases. Thus encouraging that the new generalized Pareto distribution also provides a good fit to the real life data set.

References Références Referencias

1. Abdel-Ghaly, A. A., Attia, A. F. and Aly, H. M. (1998), "Estimation of The Parameters of Pareto Distribution and The Reliability Function Using Accelerated Life Testing with Censoring", Communications in Statistics, Simulation and Computation, 27(2), 469-484.
2. Abd Elfattah, A. M. Elsherpieny, E.A and Hussein, E.A (2007). "A new generalized Pareto distribution." J. of Interstat #001,Dec-2007.
3. Arnald, B.C and press, S.J (1983). "On a general system of distribution: I its curved-shape characteristics, II. The sample median". J. Amer. Statistics. Association; 63,627-635.
4. Chaoulakian, V. and Stephens, M.A (2001). Goodness of fit tests on the generalized Pareto distribution, Technometrics, 43(1), 478-484.
5. Davison, A.C and Smith, R.L (1990).Model for exceedances over high thresholds (with Comments), Journal of the Royal Statistical Society, Ser.B, 52,393-442.
6. Hosking, J.R.M and walls (1987).Parameter and quantile estimation for the generalized Pareto distribution Technometrics, 29,339-349.
7. Picklands, J. (1975).Statistical inference using extreme under statistics, Annals of statistics, 3, 119-131.
8. Samia A. S. and Mohamed M.M. (1993). Modified moment estimators for the three parameters Pareto distribution. ISSR, Cairo University, Vol.(28), Part (2).
9. Smith R.L (1989).Thresholds methods for sample extremes in Statistical extremes and applications, Ed. J. Tiago De Oliveira, Dordrecht; Riedel 6211-6638.
10. Smith, R. L (1990).Extreme value analysis of time Series, An application to trend detection in ground-level ozone, Statistical Sciences, 4,367-393.
11. Wingo, D.R (1983). Maximum Likelihood methods for fitting the Burr Type XII distribution to the life test data. J of Biometrical, 25, 77-84.

This page is intentionally left blank