Global Journals LaTeX JournalKaleidoscope™

Artificial Intelligence formulated this projection for compatibility purposes from the original article published at Global Journals. However, this technology is currently in beta. *Therefore, kindly ignore odd layouts, missed formulae, text, tables, or figures.*

On Fitting of Generalized Pareto Distribution

Dr. T .A. Raja

Received: 11 December 2012 Accepted: 3 January 2013 Published: 15 January 2013

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.

Index terms— pareto distribution, picklands (1975) generalized pareto distribution, a new generalized Pareto distribution, Income data set, Goodness of fit.

1 Introduction a) Pareto Distribution (PD)

he 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. **??**ingo (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 ?, ?>0 and x?0

Where ? is a scale parameter and ? is a shape parameter.f(x, ?, ?, ?)= ?/ ?) (1) 1 (+ ? ? ? ? ? ? ? ? ? + ? ? ? x = 0 otherwise

where ?<x< ?, ?>0, ?>0

? is a scale parameter, ? is a shape parameter and ? is the location.

2 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 asf(x, ?, ?)= 1/ ? 1) 1 1 (? ? ? ? ? ? ? ? ? ? ? ? x = 0 otherwise

The range of x is 0?x<? for ??0 and 0?x? ?/? for ?>0

The GPD is heavy tailed, skewed and is used to model extreme values as investigated by Hoking and Well (1987), Smith (1989Smith (, 1990)), Davison and Smith (1990). Smith (1990) Author ? ? : Division of Agricultural Statistics, Skuast-K Shalimar. (J & K). E-mail : tariqaraja@rediffmail.com f(x, ?, ?)=) / ( ) 1 (+ ? ? ? ? ? ? (1.1) (2.1) (1.2)

The probability density function (p. d. f) of three parameter Pareto distribution is defined as ? is defined as the gamma function.( ) ? ? ? = ? = ? ? ? ? ? µ µ µ dx x f x x E r r r). , , , ; (.) (= ? ? ? ? ? ? ? ? ? ? ?

? + ? + ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? + ? ? ? ? ? ? ? ? ? ? ? =) 1 ( ) 1 ( ). (.) ( ) 1 1 ( ). 1 (.) 1 (0 ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ??

Here r can take any value r=1, 2, 3?., there fore mean and variance of x can be defined as,) ( ) 1 1 ( ). 1 (? ? ? ? ? ? ? ? ? ? µ + + ? = and = 2 ? + ? ? + 2 2) ( ) 1 1 ( ). 1 ( ) ( ) 2 1 ( ). 2 (? ? ? ? ? ? ? ? ? ? ? ? ? ? ? ? d)

3 Derivations of Some Distributions

Many distributions can be derived from 4parameter generalized Pareto distribution for different choices of the parameters. f(x, ?, ?, ?) =

4 c) A Model of Generalized Pareto Distribution

In this paper a model of generalized Pareto distribution as given by Abd Elfattab etal (2007) by introducing one more shape parameter "?" 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 asf(x; ?, ?, ?, ?)= ?/ ? .) (1 1) 1 (? + ? + ? ? ? ? ? ? ? ? x x

where ?<x< ?, ?>0, ?>0 and ?>0 ? and ? are shape parameter and ? is the location and ? is a scale parameter To prove that f(x) is a probability density function, following conditions are to be satisfied.i. f(x) ? 0 and ii. ? ? ? ? = 1). (dx x f

?<x< ?, ?>0, ?>0 and ?>0 are the parameters of the model. f(x) ? 0 for all x which proves (i) clearly f(x) ? 0 establishes condition (ii) as? ? ? ? = 1). (dx x f

The rth moment about mean of the generalized Pareto distribution is **??**3.1) where ?<x< ?, ?>0, ?>0 ? is a scale parameter, ? is a shape parameter and ? is the location Similarly many more distributions can be derived for suitable choice of parameters.

5 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. 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 pvalue 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 nonsignificance in both the cases. Thus encouraging that the new generalized Pareto distribution also provides a good fit to the real life data set.

6 Global

[1] [2] [3] [4]

[1] 9 Year 2013 © 2013 Global Journals Inc. (US)

[2] Year 2013 © 2013 Global Journals Inc. (US) On Fitting of Generalized Pareto Distribution

[3] © 2013 Global Journals Inc. (US)

[4] Year 2013On Fitting of Generalized Pareto Distribution

Figure 1:

1

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

Figure 2: Table 1 :

[Elfattah et al. (2007)] 'A new generalized Pareto distribution'. Abd Elfattah , A M Elsherpieny , E Hussein , E . *J. of Interstat* # 2007. Dec-2007. 001.

[Abdel-Ghaly et al. ()] 'Estimation of The Parameters of Pareto Distribution and The Reliability Function Using Accelerated Life Testing with Censoring'. A A Abdel-Ghaly , A F Attia , H M Aly . *Communications in Statistics* 1998. 27 (2) p. . (Simulation and Computation)

[Smith ()] 'Extreme value analysis of time Series, An application to trend detection in groundlevel ozone'. R Smith . *Statistical Sciences* 1990. 4 p. .

[Chaoulakian and Stephens ()] 'Goodness of fit tests on the generalized Pareto distribution'. V Chaoulakian , M Stephens . *Technometrics* 2001. 43 (1) p. .

[Wingo ()] 'Maximum Likelihood methods for fitting the Burr Type XII distribution to the life test data'. D Wingo . *J of Biometrical* 1983. 25 p. .

[Davison and Smith ()] 'Model for exceedances over high thresholds'. A Davison , R Smith . *Journal of the Royal Statistical Society, Ser.B* 1990. 52 p. . (with Comments)

[Samia and Mohamed ()] *Modified moment estimators for the three parameters Pareto distribution*, A S Samia , M M Mohamed . 1993. ISSR, Cairo University

[Arnald and Press ()] 'On a general system of distribution: I its curved-shape characteristics, II. The sample median'. B Arnald , S Press . *J. Amer. Statistics. Association* 1983. 63 p. .

[Hosking and Walls ()] *Parameter and quantile estimation for the generalized Pareto distribution Technometrics*, J R Hosking , Walls . 1987. 29 p. .

[Picklands ()] 'Statistical inference using extreme under statistics'. J Picklands . *Annals of statistics* 1975. 3 p. .

[Smith (ed.) ()] *Thresholds methods for sample extremes in Statistical extremes and applications*, R Smith . J. Tiago De Oliveira (ed.) 1989. Dordrecht. p. .