Pakistan Journal of Applied Sciences 3(2):85-88, 2003

Anderson Darling and Modified Anderson Darling Tests for Generalized Pareto Distribution

[1]M. Arshad, [2]M.T.Rasool and [3]M.I. Ahmad
[1]University College of Administrative Sciences, Kotli, Azad Kashmir
[2]Government College, Samundri, Pakistan
[3]Department of Mathematics and Statistics, College of Science, Sultan Qaboos University Masqat, Oman

Abstract: The Anderson Darling and Modified Anderson Darling test statistics are considered for testing the goodness of fit of the three parameters Generalized Pareto Distribution. The test statistics for testing the goodness of fit of the completely specified distributions are modified by replacing the Generalized Pareto distribution by their probability weighted moment estimates. The tables for critical values are derived for these empirical distribution function tests for various sample sizes.

Key Words: Intensities, Peaks Over Threshold, Precipitation, Probability Weighted Moments

Introduction

Uncertainty in the rainfall intensities and in the flood frequency analysis creeps in because a meteorologist and a hydrologist can never be sure about a fitted distribution being the same as nature has used to generate flood flows and also the data sample may not truly reflect the complete characteristics of the population. It is unlikely that a consensus will ever be reached regarding the selection of a universal parent distribution of flood flows. In literature to get over this uncertainty many candidate distributions with different parameters and parameter estimation techniques are examined, various goodness of fit criteria are used to evaluate the performance of candidate distributions. Decision making values usually lie at the tail region as compared to a two parameter distribution. So the Generalized Pareto (GP) distribution which is a three parameter distribution is selected.

Generalized Pareto Distribution (GP) was introduced by Pickands (1975) and has since been further studied by Davison (1984); Smith (1984 and 1985) and Van Montfort and Witter (1985). Van Montfort and Witter (1986) has demonstrated its applications to the distribution of peaks over threshold of rainfall series using the maximum likelihood method of estimation. Hosking and Wallis (1987) used the method of moments and the method of probability weighted moments for the estimation of the GP distribution. Using Montecarlo simulations, they concluded that the estimates obtained by the Method of Moments (MOM) and Probability Weighted Moments (PWM) were more reliable than the Maximum Likelihood Estimates (MLE). They also observed that the GP distribution gave the better fit to large peaks, thus suggesting its use for modeling Peaks Over a Threshold (POT). The GP distribution was also used by Wang (1991) for the comparison of POT and Annual Maximum (AM) models by PWM method. A comparative study for the estimates of the GP distribution was made by Moharram, Gosian and Kapoor (1993) and concluded that PWM method is best when the value of the shape parameter is less than zero and particularly if shape parameter might be less than -0.2, then PWM estimates will probably be preferred because of their low bias.

The GP distribution has applications in a number of fields, including reliability studies, in the modeling of large insurance claims, as a failure time distribution. Its application include use in the analysis of extreme events e.g. for the analysis of the precipitation data. in the flood frequency analysis, in the analysis of the data of greatest wave heights or sea levels, maximum winds loads on buildings, in the maximum rainfall analysis, in the analysis of greatest values of yearly floods, breaking strength of materials, air craft loads etc. The GP distribution has been quite popular not only for flood frequency analysis but for fitting the distribution of extreme natural events in general.

Once a distribution function is assumed or selected for study at hand, it remains to estimate its parameters and when the parameters of the model are estimated, it is then desirable to access how well the distribution fits the observed data. Goodness of fit tests are often essential to reveal departures from the assumed model. In this study the parameters are estimated by PWM method and critical points are derived for the Anderson Darling and Modified Anderson Darling tests for the Generalized Pareto distribution.

Materials and Methods

As defined by Van-Montfort and Witter (1985) a random variable X is said to be distributed as Generalized Pareto (GP) distribution:

$$X(F) = b + a/c\,[1 - (1-F)^c] \qquad c \neq 0$$
$$= b - a\,\log(1-F) \qquad c = 0$$

where a = scale parameter,
b = location parameter and
c = shape parameter

Method of Estimation: The possible methods for estimating the parameters of the GP distribution are methods of moments, method of maximum likelihood, ordinary least squares, Probability Weighted Moments (PWM) method, Generalized moment methods etc. As a comparative study of the estimation of GP distribution by Moharram *et al.,* (1993) and Hosking and Wallis (1987) that the PWM estimates will probably preferred because of their low bias, so we have used the PWM method.

Greenwood (1979) defined the probability weighted moments of X as:

$$M_{p,r,s} = E[X^p\{F(x)\}^r\{1-F(x)\}^s]$$
$$= \int X^p\,\{F(x)\}^r\,\{1-F(x)\}^s\,dF(x)$$
$$= \int_0^1 \{X(F\}^p F^r\{1\text{-}F\}^s\,dF$$

where p, r, s are real numbers and X(F) is the quantile function of X.

Goodness of Fit Technique: Goodness of fit techniques means the methods of examining how well a sample of data agrees with a given distribution as its population. The important goodness of fit techniques are:

i. Tests of chi-square types
ii. Moment ratio techniques
iii. Tests based on correlation
iv. Tests based on empirical distribution function

Most of these test statistics suffer from serious limitations. In general test of chi-square type have less power due to loss of information caused by grouping. The distribution theory of Chi-square statistics is a large sample theory. The higher order moments are usually under estimated and this fact prevents the use of moment ratio techniques and so would be the case with correlation type tests.

Several power studies have revealed Empirical Distribution Function (EDF) tests to be more powerful than other tests of fit for a wide range of sample sizes (Stephen, 1974 and 1977). Recently, satisfactory use of EDF tests has been difficult due to lack of readily available tables of significance points for the case where the parameters of the assumed distribution have to be estimated from the sample data. This case is referred to by Stephens (1974 and 1977) as case 3. The significant points that have been available are appropriate to the case where the parameters of the distribution are known. This case is referred to by Stephens (1974 and 1977) as zero case. Such tables are of limited value in practice because the parameters of the distribution are seldom known. When the parameters are estimated, the critical values are considerably smaller than for the specified parameter case. Thus the use of these critical values which are for specified parameters case to assess the agreement of a theoretical distribution when parameters are estimated from the data may result in accepting fitted distribution that ought to be rejected.

The important EDF tests are Kolmogorov-Smirnov test (KS or D), Anderson Darling test (AD or A^2_n) or modified Anderson Darling (MAD or AU^2_n) and Cramar Von Mises test (CVM or W^2). Stephens (1976) has shown that in a wide variety of situations AD is the most powerful EDF test followed by CVM and KS is rather weak. It is important to understand the behaviour of the upper tail of the distribution than it is to fit the entire distribution. Although a particular model may adequately describe most of the maximum rainfall distribution, it would be useless for predicting maximum or extreme values if the model breaks down for the upper percentiles. For assessing the behaviour of the upper tail of a distribution, a modified Anderson Darling test is included in this study.

Empirical Distribution Function: Suppose a given random sample of size n i.e., $X_1, X_2, \ldots, X_n$ and let $X_{(1)}<X_{(2)}<\ldots<X_{(n)}$ be the order statistic and also suppose that the cumulative distribution function of X is F(x). The EDF is defined as:

$$F_n(X) = \frac{\text{No. of observations} \le x}{n}$$

More specifically the definition is

$$F_n(X) = 0 \qquad x < X_{(i)}$$
$$F_n(X) = i/n \qquad X_{(i)} \le x < X_{(i+1)} \qquad i = 1,2,\ldots,n-1$$
$$F_n(X) = 1 \qquad X_{(n)} \le x$$

Thus $F_n(x)$ is a step function which shows the proportion of observations less than or equal to x for any x. While F(x) is the probability of an observation less than or equal to x.$[(F(x)=P(X\le x)]$.

EDF Statistics: A statistic measuring the difference between $F_n(x)$ and F(x) will be called an EDF statistic. Anderson and Darling (1954) proposed an EDF test as:

$$A^2_n = n \int_{-\infty}^{\infty} \{Fn(x) - F(x)\}^2 \psi(x)\, dF(x)$$

where $\psi(x) = [\{F(X)\{1-F(X)\}]^{-1}$

Sinclair *et al.*, (1990) proposed a modified form of the Anderson Darling test using the weight function:

$\psi(x) = [1 - F(x)]^{-1}$

$$AU^2_n = n \int_{-\infty}^{\infty} [Fn(x) - F(x)]^2 \psi^2(x)\, dF(x)$$

For computational purposes the Anderson Darling and Modified Anderson Darling statistics are:

$$A^2_n = -n- \frac{1}{n} \Sigma(2i-1) \{\log Z_{(i)} + \log(1-Z_{(n+1-i)})\}$$

$$AU^2_n = \frac{n}{2} - 2\sum_{i=1}^{n} Z_{(i)} - \sum_{i=1}^{n} [2-\frac{(2i-1)}{n}]\log[1-Z_{(i)}]$$

where $Z_i = F(X_i)$, i=1,2,...,n

Results and Discussion

The main objective of the research reported here in was to derive the significance points for the distribution of Anderson Darling and Modified Anderson Darling Tests for GP distribution when the parameters are estimated by probability weighted moments method. For this purpose a simulation procedure was used to approximate the distribution of Anderson Darling and Modified Anderson Darling tests. The GP random variable X was generated by taking for each of the fixed random sample sizes 10(5)40, 50 and 100 and for each value of the shape parameter -0.2, -0.1, 0.1 and 0.2. For each sample size the location and the scale parameters (b, a) were set respectively at zero and unity without loss of generality. It was expected that the distribution of test statistic would be affected more profoundly by the sample size in the range of interest than the values of the scale and shape parameters of the distribution. David and Johnson (1948) imply that the distributions of Kolmogorov Smirnov, Crammer Von Mises and Anderson Darling do not depend on the values of location and scale parameters. Each sample was used to calculate PWM estimates. By using these parameter estimates, the Anderson Darling test statistic values and Modified Anderson Darling test statistic values were calculated. This procedure was repeated 1000 times for each value of the shape parameter, thus generating 4000 Pseudo-independent values of Anderson Darling and Modified Anderson Darling test statistics. These 4000 values were then ranked and the 50th, 75th, 85th, 90th and 99th percentiles were found. A computer program was developed in MINITAB, a statistical computer package and the entire process was performed five times and the critical values of these test statistics for GP distribution are developed by computing the means of the five percentile estimates.

Table 1: Critical Values for the Anderson Darling Test for GP-distribution when the Parameters are Estimated (case 3)

n/P	0.50	0.25	0.15	0.10	0.05	0.01
10	0.27728	0.37187	0.44760	0.50029	0.59692	0.74982
15	0.28560	0.39257	0.47889	0.53025	0.60644	0.86288
20	0.29214	0.40527	0.49856	0.55300	0.61466	0.92691
25	0.29982	0.41646	0.50625	0.56828	0.62331	0.99783
30	0.30483	0.42740	0.51656	0.58442	0.63985	1.10657
35	0.30901	0.43887	0.55529	0.64405	0.66443	1.28695
40	0.32299	0.45689	0.59098	0.65815	0.69582	1.58493
50	0.37126	0.52598	0.62397	0.70945	0.84550	1.91943
100	0.51268	0.76689	0.87670	0.91608	1.91479	2.68975

Table 2: Critical Values for the Modified Anderson Darling Test for GP-distribution when the Parameters are Estimated (case 3)

			AU^2n			
n/P	0.50	0.25	0.15	0.10	0.05	0.01
10	0.14062	0.20017	0.25963	0.30864	0.48918	0.55847
15	0.14475	0.20759	0.26520	0.31478	0.50516	0.61738
20	0.14666	0.21293	0.26827	0.31726	0.51093	0.63498
25	0.14843	0.21686	0.27103	0.31940	0.51273	0.65861
30	0.15084	0.21969	0.27369	0.32083	0.51425	0.68741
35	0.15259	0.22514	0.28028	0.34354	0.52005	0.75888
40	0.16541	0.23839	0.29589	0.39166	0.54237	0.79617
50	0.18209	0.27561	0.32647	0.42559	0.58912	0.87543
100	0.20632	0.32240	0.39500	0.59577	0.76245	1.28305

Table 3: Critical Values for the Statistics A^2n and AU^2n when F(x) is Completely Known (case 0)

Statistics	Modified forms	Upper percentage points in modified forms					
		0.25	0.15	0.10	0.05	0.025	0.01
A^2_n	For all $n \geq 5$	1.248	1.610	1.933	2.492	3.020	3.857
AU^2_n	For all n	0.620	0.746	0.998	1.303	1.623	2.060

Various probabilities of the upper tail areas of the distributions of Anderson Darling and Modified Anderson Darling statistics for the GP distribution when the parameters are estimated (case 3) are presented in Table 1 and 2, respectively.

The critical values of these test statistics for the completely specified distribution are adopted from D'Agostino and Stephens (1986) and Sinclair *et al.*, (1990) and are presented in Table 3. In case 3 the dependence of the upper tail area probability P on the sample size is investigated. It is found that the distributions for case 3 for these EDF statistics depend on the sample size. This is in line with Stephens (1977) and Sinclair *et al.*, (1990) results.

It has been seen that the significance points for the case 0 are larger than the case when the parameters of the distribution are estimated from the sample. So by using these critical values of case 0 to assess the agreement of a theoretical distribution when parameters are estimated from the data may result in accepting fitted distribution that ought to be rejected.

Conclusion

Until recently satisfactory use of EDF tests has been difficult due to lack of readily available tables of significance points for the case where the parameters of the distribution have to be estimated from the sample, then the EDF tests no longer apply at least not using the critical points which are for the specified case. Thus the use of these critical values which are for a specified parameters case to assess the agreement of a theoretical distribution when the parameters are estimated from the data may result in accepting fitted distribution that ought to be rejected. In general, when unknown parameters have to be estimated particularly for finite sample sizes, percentage points can only be obtained by simulation. So in this study the tail area probabilities are derived by simulation for the Anderson Darling and the Modified Anderson Darling tests for the Generalized Pareto distribution, when the parameters are estimated by probability weighted moments method.

References

Anderson, T.W. and D. A. Darling, 1954. A test of goodness of fit. J. Am. Stat. Assoc. 49: 765-769.

Arshad, M., M. T. Rasool and M. I. Ahmed, 2002. Kolmogorov Smirnov test for Generalized Pareto Distribution. Pak. J. Appl. Sci. 2: 488-490.

David, F.W. and N.L. Johnson, 1948. The probability integral transformation when parameters are estimated from the sample. Biometrika, 35: 182-190.

Davison, A. C., 1984. Modeling excesses over high threshold with an application. In: J. Tiago de Oliveria (Editor), Statistical Extremes and Applications. Reidel, Dordrecht, 461-482.

Hosking, J. R. M. and J.R. Wallis, 1987. Parameter and quantile estimators for Generalized Pareto distribution. Technometrics, 29: 339-349.

Moharram, S. H., A. K. Gosain and P.N. Kapoor, 1993. A comparative study for the estimators of the Generalized Pareto distribution. J. Hydrolog, 150: 169-185.

Pickands, J., 1975. Statistical inference using extreme order statistics. Ann. Stat., 3: 119-131.

Sinclair, C.D., B.D. Spurr and M.I. Ahmad, 1990. Modified Anderson Darling test; Commun. Statist. Theory and Math. 19: 3677-3686.

Smith, R. L. 1984. Threshold methods for sample extremes. In: J. Tigao de Oliveria (Editor), Statistical Extremes and Applications. Reidel. Dordrecht,. 621-638.

Smith, R.L. 1985. Maximum likelihood estimation in a class of non regular cases. Biometrika, 72: 67-90.

Stephens, M. A., 1974. EDF statistics for goodness of fit and some comparisons. J. Amer. Stat. Assoc. 69: 730-737.

Stephens, M. A., 1976. Asymptotic power of EDF statistic for exponentiality against Gamma and Weibull alternatives. Tech. Report No.297, Dept. of Stat. Stanford Univ.

Stephens, M. A., 1977. Goodness of fit for the extreme value distribution. Biometrika, 64: 585-588.

Van Montfort, M. A. J. and J. V. Witter, 1985. Testing exponentiality against Generalized Pareto distribution. J. Hydrol., 78: 305-315.

Wang, Q. J., 1991. The Pot model described by the Generalized Pareto distribution with Poisson arrival rate. J. Hydrol., 129: 263-280.