Bulletin of the *Transilvania* University of Braşov • Vol 4(53), No. 2 - 2011
Series III: Mathematics, Informatics, Physics, 73-78

A STUDY ABOUT ESTIMATION OF PARAMETERS FOR EXTREME VALUE DISTRIBUTION

Alexandrina Maria PROCA[1]

Abstract

In this paper, starting from the clasical method of moments, we realize an estimation of the three parameters for extreme value distribution. These results are applied to data sets related with exchange rate returns.

2000 *Mathematics Subject Classification:* 60G70, 62P05.
Key words: Extreme value theory, statistic of extreme value, extreme index, method of moments.

1 Introduction

Extreme value theory has emerged as one of the most important statistical disciplines for the applied sciences over the last 50 years. The theory has its origin in some papers of Fréchet (1927), R. Fisher and L. Tipped (1928) and also R. von Mises (1936). In 1943 Gnedenko continued these studies and realized a complete characterization of nondegenerate limit distributions of the sequence of normalized and centered maxima and gave a partial characterization of max-domain of attraction. A complete characterization of max-domain of attraction has been given by L. Haan in 1970. In 1975, J. Pickands III was the first who studied the theory from a statistical point of view.

Definition 1. *(Jenkinson and R. von Mises) The generalized standard distribution of extreme value is defined by*

$$H_\gamma(x) = \begin{cases} \exp\left\{-(1+\gamma x)^{-1/\gamma}\right\}, & \text{for } \gamma \neq 0 \text{ and } 1+\gamma x > 0 \\ \exp\left\{-\exp\left\{-x\right\}\right\}, & \text{for } \gamma = 0. \end{cases}$$

Recall that other three parameters of the generalized standard distribution of extreme value which is useful in statistical applications is

$$H_{\gamma,\mu,\phi}(x) = \begin{cases} \exp\left\{-(1+\gamma\frac{x-\mu}{\phi})^{-1/\gamma}\right\}, & \text{for } \gamma \neq 0 \text{ and } 1+\gamma\frac{x-\mu}{\phi} > 0 \\ \exp\left\{-\exp\left\{-x\right\}\right\}, & \text{for } \gamma = 0. \end{cases}$$

[1]Faculty of Mathematics and Informatics, *Transilvania* University of Braşov, Romania, e-mail: alexproca@unitbv.ro

The parameter $\theta = (\gamma, \mu, \phi) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+$ consists of a shape parameter γ, location parameter μ and scale parameter ϕ.

An important result of probability theory with fundamental impact in non-parametric statistics is given by the following theorem.

Theorem 1 (Glivenko-Cantelli)**.** *Let $X_1, \ldots, X_n$ be a sequence of independent and identically distributed (i.i.d.) non-degenerate random variables, defined on a probability space $(\Omega, \mathcal{K}, P)$ with common distribution function F. Denote by*

$$F_n(x) = \frac{1}{n}\sum_{i=1}^{n} I_{\{X_i \le x\}},\ x \in \mathbb{R}$$

the empirical distribution functions of the i.i.d. sample $X_1, \ldots, X_n$.

Then

$$F_n(x) \xrightarrow{a.s.} EI_{\{X \le x\}} = F(x),\ x \in \mathbb{R}$$

where I_A denotes the indicator function of the set A.

Let $F^{\leftarrow}(t) = \inf\{x \in \mathbb{R} : F(x) \ge t\}$, $0 < t < 1$ be the quantile function of the distribution function F, and let $F_n^{\leftarrow}$ be the empirical quantile function for a sample X_1, $\ldots$, X_n.

If F is a continuous function, we may assume (see [5]) that $X_{n,n} < \cdots < X_{1,n}$, where $X_{n,n}, \ldots, X_{1,n}$ are the order statistics, and $F_n^{\leftarrow}(t) = X_{k,n}$, for $1 - k/n < t \le 1 - (k-1)/n$.

An important problem of Statistics is the estimation of the distribution parameters. One of the classical approaches, also used in the statistical extreme values is the method of moments, which goes back to Hosking, J.R., Wallis, J.R., Wood, E.F., (see [3]).

Consider the moment of order r

$$w_r(\theta) = E(XH_\theta^r(X))\ ,\ r \in \mathbb{N} \tag{1}$$

where H_θ is the generalized extreme value distribution and X has H_θ as distribution function with the parameter $\theta = (\gamma, \mu, \phi)$. In the sequence we will consider only the case $\gamma < 1$, $\gamma \ne 0$.

Define the empirical analogue of (1), $\hat{w}_r(\theta) = \int_{\mathbb{R}} xH_\theta^r(x)dF_n(x)\ ,\ r \in \mathbb{N}$, where F_n is the empirical distribution function corresponding to the data $X_1, \ldots, X_n$. To estimate θ we solve the equations

$$w_r(\theta) = \hat{w}_r(\theta)\ ,\ r = 0, 1, 2. \tag{2}$$

Recall that the quantile transformation is given by

$$(H_\theta(X_{n,n}), \ldots, H_\theta(X_{1,n})) \overset{\mathrm{d}}{=} (U_{n,n}, \ldots, U_{1,n}),$$

(see [5]) where $U_{n,n} \le \ldots \le U_{1,n}$ are the order statistics of an i.i.d. sequence $U_1, \ldots, U_n$ uniformly distributed on $(0, 1)$.

With this interpretation we have:

$$\hat{w}_r(\theta) = \frac{1}{n}\sum_{j=1}^{n} X_{j,n}U_{j,n}^r. \tag{3}$$

2 Main results

An important problem in the studies regarding statistic of extreme value is the estimation of parameters. In this context there exist numerous case studies which intend to obtain future predictions (see [4] and [7]).

In this paper, starting from the clasical method of moments which was presented at the beginning, we obtain two estimations for the three dimensional parameter θ by replacing $U^r_{j,n}$ with $p_{j,n} = \dfrac{n-j+0.5}{n}$ and with

$$EU^r_{j,n} = \frac{(n-j)(n-j-1)\dots(n-j-r+1)}{(n-1)(n-2)\dots(n-r)}\ ,\ r = 1,2$$

respectively.

The replacement of $U^r_{j,n}$ by $EU^r_{j,n}$ is motivated by the Glivenko-Cantelly theorem and the quantile transformation.

A widely used nonparametric statistical based approach for data analysis is the quantile plot. There are different variants of quantile plot of the type $\{(X_{k,n}, F^{\leftarrow}(p_{k,n})), k = 1, \dots, n\}$ where $p_{k,n}$ is a certain plotting position. Typical choices are

$$p_{k,n} = (n-k+\delta_k)/(n+\gamma_k),$$

with (δ_k, γ_k) appropiately chosen to allow some continuity correction.

A good choice for $p_{k,n}$ is $p_{k,n} = (n-k+0.5)/n$ (see [5]).

It is known that the moment of order r of the extreme value distribution is

$$w_r(\theta) = \frac{1}{r+1}\left\{\mu - \frac{\phi}{\gamma}(1-\Gamma(1-\gamma)(1+r)^{\gamma})\right\}$$

(see [9]), where Γ is the usual Gamma function.

The estimation $\hat{\gamma}$ of γ is obtained by solving the equations (2) for $r \in \{0,1,2\}$ using the quantile transformation mentioned before.

For $r \in \{0,1,2\}$ we have:

$$w_0(\theta) = \mu - \frac{\phi}{\gamma}(1-\Gamma(1-\gamma)) \tag{4}$$

$$2w_1(\theta) = \mu - \frac{\phi}{\gamma}(1-\Gamma(1-\gamma)2^{\gamma}) \tag{5}$$

$$3w_2(\theta) = \mu - \frac{\phi}{\gamma}(1-\Gamma(1-\gamma)3^{\gamma}) \tag{6}$$

By substracting (4) from (5) and also from (6), we obtain

$$2w_1(\theta) - w_0(\theta) = \frac{\phi}{\gamma}(1-\gamma)(2^{\gamma}-1) \tag{7}$$

and

$$3w_2(\theta) - w_0(\theta) = \frac{\phi}{\gamma}(1-\gamma)(3^\gamma - 1), \tag{8}$$

respectively.

Thus

$$\frac{3w_2(\theta) - w_0(\theta)}{2w_1(\theta) - w_0(\theta)} = \frac{3^\gamma - 1}{2^\gamma - 1}. \tag{9}$$

The left hand side term of (9) is computed by using the empirical moments from (2). The unique value of $\hat{\gamma}$ can be found by numerical computations.

From equations (7) and (8), by making use of $\hat{\gamma}$ we obtain the other two parameters

$$\hat{\phi} = \frac{(2\hat{w}_1 - \hat{w}_0)\hat{\gamma}}{\Gamma(1-\hat{\gamma})(2^{\hat{\gamma}} - 1)} \tag{10}$$

$$\hat{\mu} = \hat{w}_0 + \frac{\hat{\phi}}{\hat{\gamma}}(1 - \Gamma(1-\hat{\gamma})). \tag{11}$$

The steps of the algorithm which estimates the parameters γ, μ and ϕ are:

Step 1 Read the n input data.

Step 2 Process the data using the maximum method with a range of 7 days ([5]). We only consider the trading days (see [6]) and take the logarithm of the maximum exchange rate returns for 7 consecutive trading days. It is known that the above mentioned logarithmic values have the martingale property (see [1]).

Step 3 Compute the first three empirical moments of $X_1, \ldots, X_n$ as in equation (3) and get a first estimation by substituting $U_{j,n}$ by $p_{j,n}$. A second estimation is obtained by replacing $U_{j,n}^r$ by $EU_{j,n}^r$, where:

$$EU_{j,n}^0 = 1$$

$$EU_{j,n}^1 = \frac{n-j}{n-1}$$

$$EU_{j,n}^2 = \frac{(n-j)(n-j-1)}{(n-1)(n-2)}.$$

Step 4 Compute the quantity $\frac{3\hat{w}_2(\theta) - \hat{w}_0(\theta)}{2\hat{w}_1(\theta) - \hat{w}_0(\theta)}$.

Step 5 The solution of (9) is obtained through a numerical package (*e.g.* Matlab).

Step 6 Compute $\hat{\phi}$, $\hat{\mu}$ as in equations (10) and (11).

An implementation of this algorithm can be found in [8].

For 996 dates obtained as in Step 2 for exchange rate returns of BNR for Euro, we find the following two parameter estimations: $\hat{\gamma} = 0.573629$, $\hat{\phi} = 1.340297$, $\hat{\mu} = -0.148514$ and $\hat{\gamma} = -4.231888$, $\hat{\phi} = 0.309595$, $\hat{\mu} = 4.802198$, respectively.

The algorithm acts well on simulated data. For real data sometimes it produces inaccurate predictions. That happens especially when the quantity $\frac{3\hat{w}_2(\theta) - \hat{w}_0(\theta)}{2\hat{w}_1(\theta) - \hat{w}_0(\theta)}$ is outside the range $(1, 2)$. This is due to the initial restriction $\gamma < 1$, $\gamma \neq 0$.

3 Conclusion

This paper describes an algorithm based on the classical method of moments for parameter estimation. The proposed method is applied for exchange rate returns of BNR for Euro.

In a future study we intend to solve the critics from the end of the previous section, *e.g.* by searching convenient values for $p_{j,n}$. Another idea can be found in ([2]) which propose a different approach based on sorting ascending of the order statistics.

These estimations can be improved because in reality there are numerous situations in which recorded data have dependencies (global warming trend, inflations, and season effects). Details about the studies of trends elimination can be found in [6]. This also deserves further investigation.

References

[1] Bachelier, L., *Theorie de la speculation*, Paris, Volume 27,Gauthier-Villars, 1900.

[2] Lurie, D., Hartley, H.O., *Machine-generation of order statistics for Monte Carlo computations*,The American Statistician, Volume 26, Number 1, 1972.

[3] Hosking, J.R., Wallis, J.R., Wood, E.F., *Estimation of the Generalized Extreme Value Distribution by the Method of Probability Weighted Moments*, Tehnometrics, Volume 27, Number 3, 251-261, 1985.

[4] Smith, R.L., *Extreme Value Analysis of Environmental Time Series: An Application to Trend Detection in Ground-Level Ozone*, Statistical Science,Volume 4, Number 4, 367-393, 1989.

[5] Embrechts, P., Klppelberg, C., Mikosh, T., *Modelling Extremal Events for Insurance and Finance*, Berlin, Springer, 1997.

[6] Reiss, R.D., Thomas, M., *Statistical Analysis of Extreme Value with Applications to Insurance, Finance, Hydrology and Other Fields* , 3th edition, Berlin, Birkhauser, 2007.

[7] Proca, A.M., *Extreme value theory applications in data set study*, Preceedings of the 8th European Conference E-Comm-Line 2007, Bucharest, 71-76, 2007.

[8] Proca, A.M., *Extremal properties of sequences of independent and identically distributed random variables and applications*, Ph.D. Thesis, Braşov, 2010.

[9] Forbes, C., Evans, M., Hastings, N., Peacock, B., *Statistical Distributions*, 4th edition, Wiley, 2011.