International Journal of Innovation and Applied Studies
ISSN 2028-9324 Vol. 3 No. 1 May 2013, pp. 249-254

http://www.issr-journals.org/ijias/

Introduction to the Extreme Value Theory applied to operational risk

EL ARIF Fatima Zahra

Department of Economics and Management,
Center of Doctoral Studies,
University Mohammed V, Souissi,
Rabat, Morocco

ABSTRACT: This paper aims to present the main lines of the Extreme Value Theory applied to the operational risk. The idea is to present a methodology which allows to identify a threshold by type of risk, and to feign the losses below the threshold with the classical laws, and the losses above with a Generalized Pareto Distribution (GPD).
The adequacy of the data to the law GPD allows to consider an extreme quantile, as minimal strategy, sensitive to the size of samples, and to plan random costs whose probability of occurrence is very low, but the choice of the threshold beyond of which the observation will be judged extreme, is a point to be handled with precaution, even if we propose a technique to quantify this threshold.
Furthermore, the costs of extreme losses do not lend themselves to modeling; by definition this type of costs is rare, and the forecasts or the estimations must be often established with a big distrust, and outside the available data. The models must be used in a supple way, without believing completely to the limit.
The adoption of this method could allow the risk managers to observe the extreme events with a certain objectivity, to check the hierarchical organization of the classes of operational risks, and in the other hand, establish reserves to face these risks.

KEYWORDS: Extreme Value Theory, operational risk, threshold, GPD, extreme losses, extreme events.

1 INTRODUCTION

The everyday life of the practitioner of the statistics or the econometrist in company is to work on data which are often rebellious to the analysis. Data are missing, atypical; errors of seizure, hostility, coding, bad voluntary or involuntary statements. The location of all these imperfections is probably not the most rewarding aspect of the work, even though the methodological aspects can be completely interesting. Anyway, it is essential to the production of reliable and robust indicators.

The classic Extreme Value Theory, or based on the law of Generalized Pareto does not solve all these difficulties at once, but it provides binoculars through which practitioners can observe the extreme events with a certain objectivity, on the one hand in order to control the hierarchical organization of the classes of risk, and on the other hand, to establish reserves to cope with these extreme risks.

Thus, the Extreme Value Theory provides a number of tools probabilistic and statistical modeling of rare events [1], that we may be useful for modeling serious operational risks. In this paper, we present the main lines of this theory, and we will see how it can be applied to operational risk.

Corresponding Author: EL ARIF Fatima Zahra (elarif.tutorat@gmail.com)

2 PRESENTATION OF THE APPROACHES

There are two main approaches:

- The maxima approach, or Block Method, which consists in observing and in modelling the values of maximal losses. If we argue about the whole sample, we shall have only one maximum.

  To have a number of observations of maximum, It is necessary to cut the sample and to set maximum for every block. We so obtain a number of maxima equal to the number of blocks.

  The theorem of Fisher Tippett supplies us the limit theorem for the maximum of normalized. This law is the law of the Generalized Extreme Value. It is necessary to use then our observed Maxima to calibrate it.

- The approach said « Peaks Over Threshold (POT) », Which consists in considering the values beyond a threshold and either only the maximum. All the difficulty comes in choosing an appropriate threshold, which will allow to apply the Pickands theorem.

  This theorem allows to model the excess by Generalized Pareto Distribution (GPD). This approach is more usually used than the approach of maxima, because she allows to exploit more information supplied by the selected sample.

3 MAXIMA APPROACH

Let $(x_1, x, \dots x_n)$ a given sample representing the cost of losses.

In practice, the maximum sample size $x_{(n)}$ represent only one observation, it is not possible to rely on a single observation to our modeling.

The idea is to split the sample into m samples of size n / m obtaining m maximum values.

Then the Fisher-Tippett theorem gives us the law of maximum.

<u>The Fisher-Tippett theorem :</u>

Let $(X_1, X, \dots X_n)$ be a sequence of independent random variables and identically distributed with distribution function F_X.

Let us note: $X_{n:n} = max\left(X_1, X_1, \dots X_n\right)$.

If there are two sequences of reals $(a_n) \in R\ et\ (b_n) > 0$ and a law Non-degenerate distribution function G such as $\frac{X_{n:n} - a_n}{b_n} \to G$

Then G is necessarily in one of these three forms:

Fréchet	Weibull	Gumbel
$\Phi_\alpha(x) = \begin{cases} 0 & if\ x \le 0 \\ \exp(-x^{-\alpha}) & if\ x > 0 \end{cases}$	$\Psi_\alpha(x) = \begin{cases} \exp(-(-x)^{-\alpha}) & if\ x \le 0 \\ 1 & if\ x > 0 \end{cases}$	$\Lambda(x) = \exp(-e^{-x}) \quad for\ x \in \Re$

These three distributions are the special case of the GEV (Generalized Extreme Value):

$$G_{\mu,\sigma,\xi}(x) = \begin{cases} \exp\left(-(1+\xi.\frac{x-\mu}{\sigma})^{-1/\xi}\right), & if\ (1+\xi.\frac{x-\mu}{\sigma})^{-1/\xi} > 0 \quad and\ \xi \neq 0, \\ exp\left(-exp\left(-\frac{x-\mu}{\sigma}\right)\right), & if\ (1+\xi.\frac{x-\mu}{\sigma})^{-1/\xi} > 0 \quad and\ \xi = 0. \end{cases}$$

- ξ represents the index of extreme values. The more it will be higher in absolute value, the more the weight of extremes in the initial distribution will be important.
- μ the location parameter. It indicates approximately the heart of the distribution.
- σ the scale parameter. It shows the spread of extremes.

Fréchet	Weibull	Gumbel
$\xi = \alpha^{-1} > 0$ $\mu_n = 0$ $\sigma_n = F^{-1}\left(1 - \frac{1}{n}\right)$	$\xi = -\alpha^{-1} < 0$ $\mu_n = x_F$ $\sigma_n = x_F - F^{-1}\left(1 - \frac{1}{n}\right)$	$\xi = 0$ $\mu_n = F^{-1}\left(1 - \frac{1}{n}\right)$ $\sigma_n = \frac{1}{S(\mu_n)} \int_{\mu_n}^{\infty} S(t)dt$

Where $S(X) = 1 - F(X)\ and\ x_F = sup\{x\ as\ \ F_X(x) < 1\}$

The F_X distribution function affects the limit distribution G, that is to say on parameters, it is called in the domain of attraction of G.

The Extreme Value Theory allows us to know the laws that belong the three domains of attractions:

- The laws of Cauchy, Pareto, Student and LogGamma belong to the domain of attraction of the Fréchet distribution : they are laws to « thick tails »,
- The laws of Weibull, Normal, Exponential, Gamma and Log-Normal belong to the domain of attraction of the Gumbel distribution : they are laws to « fine and medium tails »,
- Uniform and Beta laws belong to the domain of attraction of the Weibull distribution : they are laws to « fine tails ».

Thus, if we know the law claims, we can deduce the law of maximum. However, in practice we do not know the law of our theoretical sample and the parameters must be estimated by statistical methods. There are several methods to estimate the parameters: we can distinguish parametric and nonparametric approaches.

Parameter estimation by the method of maximum likelihood:

Let us consider the sample of maxima on the m samples created, noted : $(M_1, M_2 \ldots, M_m)$.

The log-likelihood can be written:

For $\xi \neq 0$,

ln $L(M_1, \ldots, M_m;\ \xi, \mu, \sigma) = \sum_{i=1}^{m} \ln(h_{\xi,\mu,\sigma}(M_i))$

$$= -m\ ln\ \sigma - (\frac{1+\xi}{\xi}) \sum_{i-1}^{m} \ln(1 + \xi \frac{M_i - \mu}{\sigma}) - \sum_{i=1}^{m} \ln(1 + (\xi \frac{M_i - \mu}{\sigma})^{-1/\xi})$$

Where $h_{\xi,\mu,\sigma}$ is the density of the Generalized Extreme Value (GEV).

$$h_{\xi,\mu,\sigma} = \begin{cases} \frac{1}{\sigma}\left[1 + \xi(\frac{x-\mu}{\sigma})\right]^{-\left(\frac{1+\xi}{\xi}\right)} \exp(-(1 + \xi\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}), & if (1 + \xi\left(\frac{x-\mu}{\sigma}\right)^{-\frac{1}{\xi}}) > 0, \xi \neq 0, \\ \frac{1}{\sigma}\exp\left(-\frac{x-\mu}{\sigma} - \exp\left(\frac{x-\mu}{\sigma}\right)\right), & if\ \xi = 0. \end{cases}$$

The solution of this equation can be done by numerical methods (Quasi-Newton algorithm). In practice, we can use the predefined functions of the statistical software such as R[1] software.

Note: When the sample size is small, the estimator of the maximum likelihood provides a biased estimate.

In the case ξ = 0, the log-likelihood can be written :

ln $L(M_1, \dots, M_m;\ \xi, \mu, \sigma)$= -n ln $\sigma - \sum_{i=1}^{m} \exp(-\frac{M_i-\mu}{\sigma}) - \sum_{i=1}^{m} \frac{M_i-\mu}{\sigma}$.

The resolution is still by numerical methods.

Maxima approach is used in cases where we have periodic observations. In the context of operational risk, this method cannot be applied to the extent that we do not necessarily observations every months and not much recoil.

4 « Peaks Over Threshold (POT) » Approach

This method [2] aims to model the distribution of the observations beyond a threshold by the Generalized Pareto distribution through the following theorem:

Pickands theorem-Balkema-de Haan:

Let X be a random variable of distribution function F_X. .

$$\lim_{n\to\infty} P(\frac{X_{n,n}-b_n}{a_n} \le x) = G(x) \Leftrightarrow \lim_{u\to X_F} sup_{[0,X_F]} \left\| F_X^u(x) - G_{\xi,\sigma(u)}(x) \right\| = O$$

With $x_F = sup\{x\ as\ F_X(x) < 1\}, F_X^u(x) = P(X-u \le x | X > u)$ and $G_{\xi,\sigma}$ is the distribution function of the Generalized Pareto Distribution:

$$G_{\xi,\sigma}(x) = \begin{cases} 1-(1+\xi x/\sigma)^{(-1/\xi)} & if\ \xi \neq 0 \\ 1-e^{(-x/\sigma)} & if\ \xi = 0 \end{cases}$$

Note: the tail index involved in the maximum limit is identical to the ξ parameter of the GPD.

For a fixed threshold u, the Excess distribution function over this threshold may be written:

$F_X^u(x) = P(X-u) \le x | X > u)$ = $\frac{P(X\le x+u \cap X>u)}{P(X>u)} = \frac{F_X(x+u)-F_X(u)}{1-F_X(u)}$

u must be sufficiently large so that we can apply the above result, but u should not be too large in order to have enough data to obtain good estimators. Typically, u is determined graphically. For sample sizes less than 500, the number of excess is between 5 and 10% of the sample.

Methods to determine the threshold u:

Let us define the Mean Excess Function, it is about the expectation of excess, knowing that the losses exceed the threshold:

$$e(u) = E(X-u|X>u) = \int_u^{x_F} (y-u) \frac{dF_X(y)}{1-F_X(u)}\ with\ x_F = \{x\ as\ F_x(x) < 1\}$$

[1] R is a software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms.

This function can be estimated by the Mean Excess Function:

$$e_n(u) = \frac{\sum_{i=1}^{n}(X_i - u)}{\sum_{i=1}^{n} I_{\{X_i > u\}}}$$

For $\nu \geq u, E\{X - \nu | X > \nu\} = \frac{\sigma(u) + \xi \times (\nu - u)}{1 - \xi}$ if the excess over the threshold follow a GPD.

The conditional expectation of excess is a linear function of v, where v is greater than the reference threshold u. This provides a way to test if the empirical u threshold chosen for the calculations is sufficiently high, the expectations of the excesses over higher thresholds must be aligned on a straight line of slope $\frac{\xi}{1-\xi}$.

Hill estimator : it is an estimator of the tail index which is only valid for $\xi > 0$, is written :

$$\hat{\xi}_{k,n}^{H} = \frac{1}{k(n)} \sum_{i=n-k(n)+1}^{n} lnX_{(i,n)} - lnX_{(n-k(n)+1,n)}$$

An alternative approach to the « mean excess plot » is to calculate the Hill estimator of the tail index for different thresholds, and look from what threshold this estimator is approximately constant.

It is also possible to calculate, an estimate of the tail index by method of maximum likelihood, for different thresholds, and search from what point this estimator is approximately constant.

Parameter estimation:

When u is determined, the parameters can be estimated by maximum likelihood method.

The density of the Generalized Pareto Distribution can be written:

$$g_{\xi,\sigma}(x) = \begin{cases} \frac{1}{\sigma}(1 + \xi\frac{x}{\sigma})^{-(1+\xi)/\xi} & if\ \xi \neq 0 \\ \frac{1}{\sigma}e^{(-x/\sigma)} & if\ \xi = 0 \end{cases}$$

Parameter estimation by maximum likelihood method :

The log-likelihood is written $\ln L(X_1, X_2, \ldots, X_{n_u}; \xi, \sigma) = \sum_{i=1}^{n_u} \ln(g_{\xi,\sigma}(X_i))$ avec n_u the sample size naked $X_1, X_2, \ldots, X_{n_u}$ which contains the losses above the threshold u.

- For ξ ≠ 0, the log likelihood is equal to:

ln $\ln L(X_1, X_2, \ldots, X_{n_u}; \xi, \sigma) = -n_u \ln(\sigma) - (\frac{1}{\xi} + 1)\sum_{i=1}^{n_u} \ln(1 + \frac{\xi}{\sigma}X_i)$.

In this case, the maximization of the log-likelihood is effected by numerical methods.

- For ξ = 0, the log likelihood is equal to:

$$\ln L(X_1, X_2, \ldots, X_{n_u}; \xi, \sigma) = -n_u \ln(\sigma) - \frac{1}{\sigma}\sum_{i=1}^{n_u} X_i$$

In this case, the maximization gives an analytical form of the estimator $\hat{\sigma}_{n_u} of\ \sigma$:

$$\hat{\sigma}_{n_u} = \sum_{i=1}^{n_u} \frac{X_i}{n_u}$$

There are other methods for estimating the parameters of a GPD. Hosking and Wallis [3] showed that for samples of size less than 500, Methods of Moments and Weighted Moment Method, are more effective than the Maximum Likelihood Method. However, the main problem is the domain of validity of $\xi : \xi < 1/2$.

Jean Diebolt, Armelle Guillou and Imen Rached [4], found a method to extend the domain of validity to : $\xi < 3/2$. It is the Generalized Method of Moments.

For a sufficiently high threshold, the frequency of excess is modeled by the Poisson Distribution.

5 CONCLUSION

This method is very interesting in the context of operational risk. In fact, the idea is to find a threshold type of risk, and simulate losses below threshold with classical laws, and losses above with a GPD. Then to aggregate them by assuming that, the severe losses are independent from attritional losses[2], which is an assumption commonly used. If the number of severe losses is too low by risk, we shall consider our base and shall consider the threshold of severe losses on all the base.

The adequacy of the data to the law GPD allows to consider an extreme quantile, as minimal strategy, sensitive to the size of samples, and to plan random costs whose probability of occurrence is very low, but the choice of the threshold beyond of which the observation will be judged extreme, is a point to be handled with precaution, even if we propose a technique to quantify this threshold. This technique, based on decrease of the variance, seems to be a good empirical compromise between the FEM (Finite Element Method) and that of the GPD.

The costs of extreme losses do not lend themselves to modeling ; by definition this type of costs is rare and the forecasts or the estimations must be often established with a big distrust, and outside the available data. The models must be used in a supple way, without believing completely to the limit. The approach must be opened and multi-form, and in this sense, there is no method for a problem.

REFERENCES

[1] E. Planchet, « Using the theory of extreme values in the context of solvency 2 », Course material, April 2008.

[2] Alexander, McNeil, T. Saladin, « The peaks over thersholds for estimating high quantiles of loss distribution », International ASTIN Colloquium, pp. 70-94, 1997.

[3] J. Wallis, JRM Hosking, « Parameter and quantile estimation for the Generalized Pareto Distribution », Technometrics, Volume 29, Issue 3, pages 339-349, 1987.
DOI: 10.1080/00401706.1987.10488243

[4] J. Diebolt, A. Guillou et I. Rached, “Approximation of the distribution of excesses using a generalized probability weighted moment method”, Comptes rendus - Mathematiques, Volume 340, issue 5, pp. 383-388, 2005.
DOI: 10.1016/j.crma.2005.01.017

[2] The risk of loss events with a low potential, but with a high probability of occurrence.