Expected shortfall of claim amounts: some practical aspects[†]

Frédéric Planchet[*] **Pierre-E. Thérond**[α]

ISFA – Laboratoire SAF [β]
Université de Lyon - Université Claude Bernard Lyon 1
WINTER & Associés [γ]

Abstract

In this paper, we focus on the practical aspects of the methodologies which enable the calculation of the expected shortfall of claim amounts. Although many theoretical results exist on the subject, the operational setting for applications is seldom discussed. One shows that the choice of the methods to adopt requires some particular investigations. Moreover the results obtained are overall not very robust when the only extreme value theory is used (without making an assumption on the structure of the tail distribution).

KEYWORDS: Reserving, extreme values.

Expected shortfall of claim amounts :

some practical aspects

Abstract

This article focuses on the robustness of practical methods used in calculating the expected shortfall of claim amounts. Although many theoretical results exist on this subject, the operational implementations have seldom been discussed. It shows that the choice of methods to be adopted requires special investigations and the obtained results are generally not very robust when we stick to the extreme value theory.

KEYWORDS: Reserving, extreme values.

[†] This paper will be presented in English.

[*] Corresponding author. Contact : fplanchet@winter-associes.fr

[α] Contact : ptherond@winter-associes.fr

[β] Institut de Science Financière et d'Assurances (ISFA) - 50 avenue Tony Garnier - 69366 Lyon Cedex 07.

[γ] WINTER & Associés – 43-47 avenue de la Grande Armée - 75116 Paris et 18 avenue Félix Faure - 69007 Lyon.

1. Introduction

In many situations in insurance, it is essential to have the robust information available for the behaviour of the large numbers of a statistical series.

We can certainly refer to the project Solvability 2 (see directive draft framework published 2007) which have profoundly modified the rules for determining the level of equity in insurance as a criterion by introducing explicit control of overall risk implied by the company, that risk should be quantified through the probability of ruin to one year horizon. The level kept at 99.5%, implies the need to estimate a high order percentile of the distribution of interest (in practice here, that of the excess or the asset/liability margin). The problems raised by such a criterion are studied in Thérond, Planchet (2007). Outside the European Union, the Swiss Solvency Test also uses extreme values using the expected shortfall (expectation of excesses over a threshold) to 99% in calculating the solvency of insurance companies.

One can of course cite the Solvency 2 (see draft framework directive was published 2007), which has been profoundly altered the rules for determining the level of equity in insurance as a criterion by introducing explicit control of the overall risk borne by society, that risk should be quantified through the probability of ruin to one year horizon. The level chosen by 99.5% implies the need to estimate a high order percentile of the distribution of interest (in practice here that surplus, or margin ALM). The problem posed by such a criterion is studied in Thérond, Planchet (2007). Outside the European Union, the Swiss Solvency Test also uses extreme values using the expected shortfall (hope beyond a threshold) to 99% in calculating the solvency of insurance companies.

In a more traditional way, the question arises in non-life insurance when calculating reserves (cf. Partrat et Besson (2005)); then it is essential to be able to effectively determine the estimates of the expected loss exceeding the threshold interpreted as the boundary between the "large loss" and the "standard" claims. Especially since, in number of cases, the contracts are reinsured by policies covering such excess loss (XS).

In this article, after a brief review of methodologies, we examine the behaviour of different estimators from simulated data where the real distribution is known. We then conclude that the approaches exclusively based on the results of extreme value theory are not very robust, particularly the low volume of data available for estimating parameters, but also because the large number of the distribution being studied are not necessarily the extreme value as in the theory.

The parametric approaches are based on a mixed model in which special attention is paid to the modeling of the large numbers of the distribution (cf. Planchet and Thérond (2007) on this subject).

2. Distribution of excesses over a threshold: theoretical review

Given sample random variables $X_1,...,X_n$ (assumed to represent the loss amount), for a fixed threshold u, we are interested in the expectation of the excesses over that threshold. It is defined by the formula

$$e_X(u) = \mathrm{E}\left(X - u \middle| X > u\right).$$

The threshold u is supposed to be large.

Two approaches are *a priori* considered:

- Calculate explicitly or using an approximation method for u sufficiently large, the value of $e_X(u)$ under a known parametric distribution;
- Use a non-parametric approach saving the effort of specifying the distribution of *X*.

The first approach is traditional and expressions of $e_X(u)$ for the classical distributions can be found in many references (cf. for example Partrat et Besson (2005) p. 460). A mixed distribution is presented in Planchet and Thérond (2007) in addressing the problems of estimation and tests. It will not be developed here. We'll focus working on the non-parametric approach in the following.

We consider the following ordered sample $X_{(1)} \leq ... \leq X_{(n)}$; in order to reduce the notations, sometimes we will note $X_{k,n} = X_{(n-k+1)}$, so that the sub-sample of the k largest observations are written as: $X_{k,n} \leq ... \leq X_{1,n}$. We denote $S(x) = 1 - F(x)$ the survival function associated to *F*, the cumulative distribution function of *X*. This function allows us to simply describe the excess distribution, where

$$S_u(x) = \Pr\left(X - u > x \mid X > u\right) = \frac{S(u+x)}{S(u)}$$

This approach was introduced by De Haan and Rootzen (1993). For a class of rather large probability distribution, the Pickands theorem (1975) allows to make sure that there exists a constant of positive normalization $a(u)$ such that:

$$\lim_{u \to \omega(F)} S_u\left(xa(u)\right) = S_{\xi,\beta}(x)$$

with

$$S_{\xi,\beta}(x) = \begin{cases} \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ \exp\left(-\frac{x}{\beta}\right) & \text{si } \xi = 0 \end{cases}$$

and $\beta > 0$. This limit distribution is called Generalized Pareto Distribution (GPD). ξ is the tail index of the associated extremes of the distribution. We note that $\xi = 0$ corresponds to the domain of the attraction of Gumbel, which contains in particular the exponential, normal, log-normal, gamma, Weibull and Benktender distributions and so on. In the practical situations of property and casualty insurance, we will find that in the Fréchet domain, for which $\xi > 0$, with distributions of Pareto, log-gamma, Burr, stables, etc.

The equality below can be reformulated in the following way: when u is large (i.e. close to the terminal point $\omega_F = \sup\left\{y \in \mathbb{R}, F(y) < 1\right\}$), then we can find $\beta(u) = \beta \times a(u)$ such that:

$$S_u(x) \approx S_{\xi,\beta(u)}(x).$$

Moreover, the function $\beta(u)$ is unique to a rare multiplicative constant. In the case where F is in the Gumbel domain of attraction, we can show that we can take $\beta(u) = \frac{1}{S(u)} \int_u^{\omega(F)} S(t)\,dt$. This result can express that the distribution of excesses over a threshold converges to a GPD according to the sample size, and hence the threshold tends to the infinity (at least in the case where the distribution is not bounded, *i.e.* $\omega(F) = +\infty$).

2.1. Properties of Generalized Pareto Distribution (GPD)

Let X be the generalized Pareto random variable with parameter (ξ, β). Then we have the following results (cf. Embrechts et al. (1997))

Property 1: If $\xi < 1$, we have

$$\mathrm{E}\left(1 + \frac{\xi}{\beta} X\right)^{-r} = \frac{1}{1+\xi}, \text{ for } r > -1/\xi,$$

$$\mathrm{E}\left[\ln\left(1 + \frac{\xi}{\beta} X\right)\right]^{k} = \xi^k k!, \text{ for } k \in \mathbb{N},$$

$$\mathrm{E}\left(X\left(1 - G_{\xi,\beta}(X)\right)^r\right) = \frac{\beta}{(r+1-\xi)(r+1)}, \text{ for } (r+1)/\xi > 0.$$

As soon as we are in the context of a heavy-tailed distribution $(\xi > 0)$, the existence of moments in different orders is no longer guaranteed. Precisely, we have:

Property 2: The random variable Y admits the moments up to order $\left[\xi^{-1}\right]$ and we have

$$\mathrm{E}\left(X^r\right)=\frac{\beta^r\Gamma\left(\xi^{-1}-r\right)}{\xi^{r+1}\Gamma\left(\xi^{-1}+1\right)},\ \text{for } r\le\left[\xi^{-1}\right].$$

In particular, $\mathrm{E}\left(X\right)=\dfrac{\beta}{1-\xi}$.

Property 3: (Stability) The random variable $X_u=\left[X-u\mid X>u\right]$ is GPD distributed with parameters $\left(\xi,\beta+\xi u\right)$. We derive if $\xi<1$, then for all $u<\omega_F$,

$$\mathrm{E}\left[X-u\mid X>u\right]=\frac{\beta+\xi\times u}{1-\xi},\ \text{for } \beta+\xi u>0.$$

If $\xi\ge 1$, the expectation doesn't exist any more.

2.2. Estimation of parameters in Generalized Pareto Distribution (GPD)

The estimation of parameters can be effected by the maximum likelihood approximation. The density function of GPD is expressed as:

$$g_{\xi,\beta}(x)=\begin{cases}\beta^{\frac{1}{\xi}}\left(\beta+\xi x\right)^{-\frac{1}{\xi}-1} & \text{si } \xi\ne 0\\ \beta^{-1}\exp\left(-\dfrac{x}{\beta}\right) & \text{si } \xi=0\end{cases}$$

and we derive the log-likelihood from it:

$$\ln L\left(\xi,\beta;X_1,\ldots,X_n\right)=-n\ln\beta-\left(1+\frac{1}{\xi}\right)\sum_{i=1}^{n}\ln\left(1+\frac{\xi}{\beta}X_i\right).$$

Using the reparametrisation $\tau=\xi/\beta$, canceling partial derivative of the log-likelihood leads to system:

$$\begin{cases}\xi=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\ln\left(1+\tau X_i\right)=:\hat{\xi}(\tau),\\ \dfrac{1}{\tau}=\dfrac{1}{n}\left(\dfrac{1}{\xi}+1\right)\displaystyle\sum_{i=1}^{n}\dfrac{X_i}{1+\tau X_i}.\end{cases}$$

The estimator of maximum likelihood of $\left(\xi,\tau\right)$ is $\left(\hat{\xi}=\hat{\xi}(\hat{\tau}),\hat{\tau}\right)$, where $\hat{\tau}$ is the solution of

$$\frac{1}{\tau}=\frac{1}{n}\left(\frac{1}{\hat{\xi}(\tau)}+1\right)\sum_{i=1}^{n}\frac{X_i}{1+\tau X_i}.$$

We solve the last equation numerically with iteration to the best that we dispose an initial value τ_0 not too far from τ. In practice, this initial value can be obtained by the moment generating method (so that it exists up to order 2) or by the quantile method. Once obtained estimator $\hat{\tau}$, we derive immediately $\hat{\xi}=\hat{\xi}(\hat{\tau})$. When $\xi>-1/2$, Hosking and Wallis (1987) have shown the asymptotic normality of the maximum likelihood estimators:

$$n^{1/2}\left(\hat{\xi}_n-\xi,\frac{\hat{\beta}_n}{\beta}-1\right)\xrightarrow[\mathbf{n}n\to\infty]{\mathrm{L}}\mathrm{N}\left[0,\left(1+\xi\right)\begin{pmatrix}1+\xi & -1\\ -1 & 2\end{pmatrix}\right].$$

In particular, this result allows calculating errors in estimating the estimators of maximum likelihood. We can also note that quantile function has an explicit expression, when $\xi>0$:

$$x_p=G_{\xi,\beta}^{-1}\left(p\right)=\frac{\beta}{\xi}\left(\left(1-p\right)^{-\xi}-1\right).$$

That allows proposing an alternative estimator of ξ by the quantile method, noting that:

$$\frac{x_{p_1}}{x_{p_2}} = \frac{\left(1-p_1\right)^{-\xi}-1}{\left(1-p_2\right)^{-\xi}-1}.$$

Once the estimation is completed, we can verify graphically the relevance of the estimations by comparing the estimated Generalized Pareto Distribution with the empirically observed distribution of observations of excesses above the threshold. From a practical point of view, the estimation of parameters $\left(\xi,\beta\right)$ for a fixed threshold *u* leads to the following estimation of distribution of excesses over a threshold: $\hat{S}_u\left(x\right) \approx S_{\hat{\xi},\hat{\beta}(u)}(x)$. We will note that the proposed approximation of parameter β by this formula depends on the selected threshold *u*.

3. Application: the expectation of excesses over a threshold

The expectation of excesses over a threshold *u* (or mean excess) is defined for a distribution *F* by:

$$e_X(u) = \mathrm{E}\left(X-u \middle| X>u\right).$$

It's thus a matter of expectation of conditional distribution introduced below:

$$e_X\left(u\right) = \int_0^{\omega_F} S_u\left(x\right)dx.$$

In the special case where X is distributed by the generalized Pareto distribution with parameters $\left(\xi,\beta\right)$, we have seen in paragraph 2.1 that $e_X\left(u\right) = \frac{\beta+\xi u}{1-\xi}$. This expectation only exists when $\xi<1$, we will assume in the following that this condition is satisfied.

3.1. Estimators of expectation of excesses over a threshold

In practice, if N_u is the number of observations of thresholds *u* above, we have the empirical estimator following $e\left(u\right)$:

$$\hat{e}_1(u) = \frac{1}{N_u}\sum_{j=1}^{n}(x_j-u)1_{\{x>u\}}(x_j).$$

We can equivalently propose as estimator:

$$\hat{e}_2(u) = \frac{\hat{\beta}\left(u\right)}{1-\hat{\xi}},$$

with parameters $\left(\xi,\beta\right)$ estimated by the maximum likelihood (cf. the paragraph 2.2). At last, a third method allows returning to the estimation of a single parameter ξ, in the special case where $\xi>0$. For that we observe that the survival function of the distribution can be set under the form

$$S\left(x\right) = 1-F\left(x\right) = x^{-1/\xi}L\left(x\right),$$

where *L* is a slowly varying function at infinity as in where

$$\lim_{t\to\infty}\frac{L\left(tx\right)}{L\left(t\right)} = 1.$$

With the threshold *u*, we can thus write that

$$S_u\left(x\right) = \frac{S\left(x\right)}{S\left(u\right)} = \frac{L\left(x\right)}{L\left(u\right)}\left(\frac{x}{u}\right)^{-1/\xi}$$

and considering that the ratio of slowly varying function is close to 1 we find:

$$S_u(x) \approx \left(\frac{x}{u}\right)^{-1/\xi}.$$

Since $\xi < 1$ we find that $e_x(u) = \int_0^{+\infty} S_u(u+t)\,dt \approx \frac{\xi u}{1-\xi}$. It's then sufficient to determine an estimator of ξ, for example, the estimator of Hill, in order to propose as estimator of expectation of excesses over u :

$$\hat{e}_3(u) = \frac{\hat{\xi} u}{1-\hat{\xi}}.$$

$\hat{\xi}$ can be chosen to estimate for a fixed threshold $u_0 \le u$, where re-estimated for every value of u, which thus leads to $\hat{\xi} = \hat{\xi}(u)$.

3.2. Illustration : the Pareto distribution case

We illustrate the implementation of estimators in the example of a Pareto distribution with cumulative distribution function $F(x) = 1 - \left(\frac{x_m}{x}\right)^{\alpha}$ for $x \ge x_m$. In this configuration, the GDP approximation is in fact exact:

$$S_u(x-u) = \Pr\left(X > x \middle| X > u\right) = \left(\frac{x}{u}\right)^{-\alpha} = \left(1 + \frac{\xi}{\beta}(x-u)\right)^{-1/\xi}$$

posing $\xi = \frac{1}{\alpha}$ et $\beta(u) = \xi u$. In the particular case, from the fact of invariance by the change of threshold in the Pareto distribution, the estimation is in fact returned to the estimation of a threshold parameter $\xi = \frac{1}{\alpha}$.

3.3. Determination of a threshold

We derive from property 3, for a fixed threshold u such that $S_u(x) \approx S_{\xi,\beta(u)}(x)$ and if we consider $v \ge u$, then denoting $Y = \left[X - u \middle| X > u\right]$, we have the equality of distribution

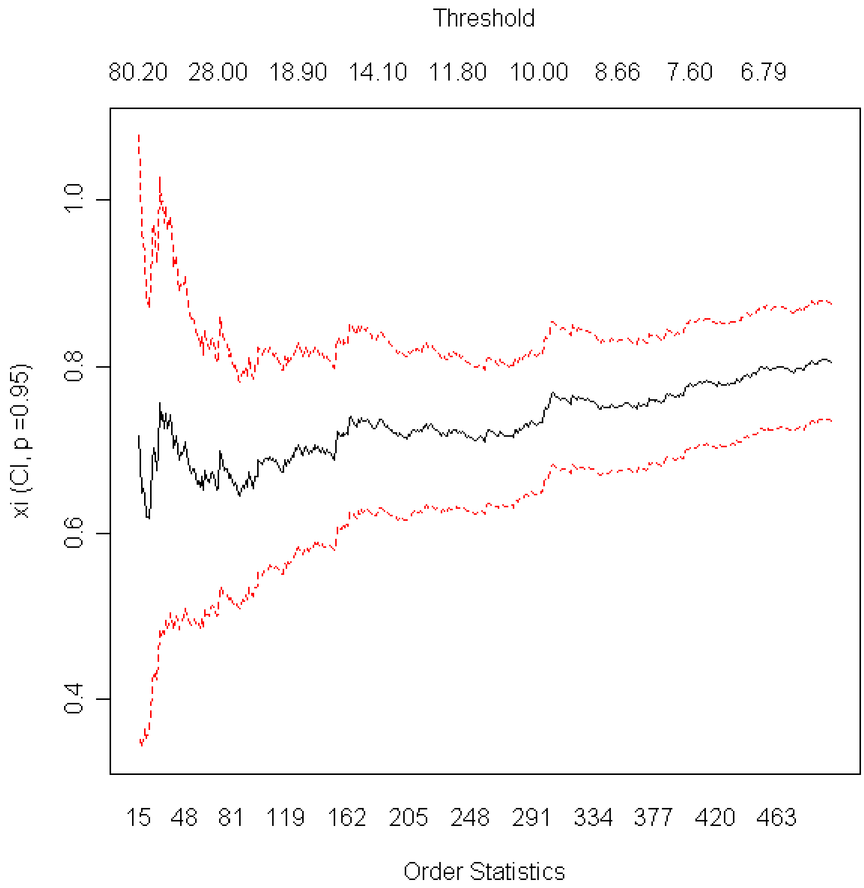
$$\left[X - v \middle| X > v\right] = \left[Y - (v-u) \middle| Y > v - u\right],$$

and since Y has a generalized Pareto distribution $(\alpha, \beta(u))$:

$$\mathrm{E}\left[X - v \mid X > v\right] = \frac{\beta(u) + \xi \times (v-u)}{1-\xi}.$$

Hence, the conditional expectation of the excesses is an affine function of v, when v is superior to the threshold of reference u. That gives an empirical mean to test if the only u chosen for the calculation is sufficiently high: the highest expectation of excesses over a threshold should be aligned on the slope $\frac{\xi}{1-\xi}$.

An alternative approach consists of calculating the estimator of Hill of index tail for different threshold and to find from such a threshold this estimator is approximately constant:



4. Numerical applications

4.1. Methodology

The applications presented here are realized with the software programme R[1]. The objective is to compare the three proposed estimators:

$$\hat{e}_1(u) = \frac{1}{N_u}\sum_{j=1}^{n}(x_j - u)1_{\{x>u\}}(x_j),$$

$$\hat{e}_2(u) = \frac{\hat{\beta}_{GPD}(u)}{1-\hat{\xi}_{GPD}}, \text{ et}$$

$$\hat{e}_3(u) = \frac{\hat{\xi}_{HILL}}{1-\hat{\xi}_{HILL}}u.$$

The adopted approach consists of simulating a sampling of data according to a known distribution, then to compare, for different thresholds, the theoretical value $e(u)$ with the estimation results. In practice this technique will be applied in two contexts:

- when the underlying distribution is in the Fréchet domain: $\xi > 0$ and the estimators $\hat{e}_2$ and $\hat{e}_3$ are theoretically justified;
- when the underlying distribution is in the Gumbel domain: $\xi = 0$, and in this case the estimators $\hat{e}_2$ and $\hat{e}_3$ don't make any sense; they will be calculated in order to bring an opinion on the robustness of these techniques.

[1] The code R is available on the site http://www.isfa.info/r.

The selection of $\xi > 0$ or $\xi = 0$ is complex and needs investigations beyond the scope of this article. We should pay particular attention to the log-normal distribution which is slow variation and sub-exponential at the same time, which gives an intermediate behaviour between the light-tailed distribution and heavy-tailed distribution.

4.2. Implementation in R programme

The detailed of functions used in R are listed below. These functions assume that a sample X is available. The calculations are conducted for all the thresholds associated from 90 % to 99.5 % quantile, with one without 0.5 %. The thresholds are obtained *via* the following R function:

```
seuils=function(X){
# Renvoie un tableau avec les 20 quantiles de 90% à 99,5%
      seuils=as.vector(1:20)
      for (i in 1:20){
            seuils[i]=quantile(X,.895+i/200)
      }
      seuils
}
```

Thus they are the empirical quantiles.

4.2.1. Empirical estimator

```
e_emp=function(X,seuils){
# Ajustement empirique de l'espérance au-delà d'un seuil
# Retourne un tableau avec l'estimation du mean XS pour les quantiles
définis dans "seuils"

      e_est=as.vector(1:length(seuils))
      for (i in 1:length(seuils)){
            e_est[i]=mean(X[X>seuils[i]])
      }
      plot(seuils, e_est)
      res=data.frame(e_est)
      res
}
```

4.2.2. GPD estimator

```
e_gpd=function(X,seuils){
# Ajustement GPD de l'espérance au-delà d'un seuil
# Retourne un tableau avec l'estimation du mean XS pour les quantiles
définis dans "seuils"

      xi_est=as.vector(1:length(seuils))
      beta_est=as.vector(1:length(seuils))
      e_gpd=as.vector(1:length(seuils))
      library(evir)
      for (i in 1:length(seuils)){
            w=gpd(X,seuils[i])
            fitgpd=w$par.ests
            xi_est[i]=fitgpd[1]
            beta_est[i]=fitgpd[2]
      }

      e_est=beta_est/(1-xi_est)
      plot(seuils,xi_est)
```

```
        res=data.frame(e_est,xi_est,beta_est)
        res
}
```

4.2.3. Hill estimator

```
e_hill=function(X,seuils){
# Ajustement Hill de l'espérance au-delà d'un seuil
# Retourne un tableau avec l'estimation du mean XS pour les quantiles
définis dans "seuils"

        xi_est=as.vector(1:length(seuils))
        e_est=as.vector(1:length(seuils))
        library(evir)
        #On ne travaille que sur les 10% de valeurs les plus élevées de
        X
        p=.1
        u=quantile(X,1-p)
        Xu=X[X>u]
        Xs=rev(sort(Xu))
        esthill=hill(Xs,"xi")
        for (i in 1:length(seuils)){
              k=trunc((.895+i/(length(seuils)*10))*length(Xs))
              xi_est[i]=esthill$y[esthill$x==k]
              }
        e_est=xi_est*seuils/(1-xi_est)
        plot(seuils,xi_est)
        res=data.frame(e_est,xi_est)
        res
}
```

In the code above, we limit the estimation to the sub-sample consisting of 10% of the highest values.

4.2.4. Theoretical values

The theoretical values used as references are provided by the function below:

```
e_th=function(dist,param,seuils){

        #Loi exponentielle
        if (dist=="dexp"){
              e_th=1/param[1]*(seuils/seuils)
        }
        #Loi normale
        if (dist=="dnorm"){
              e_th=1
        }
        #Loi log-normale
        if (dist=="dlnorm"){
              s=param[1]
              e_th=s^2*seuils/log(seuils)
        }
        #Loi de Weibull (Fréchet)
        if (dist=="dweibull"){
              xi=param[1]
              e_th=xi*seuils^(1-1/xi)
        }
        #Loi GPD
        if (dist=="dgpd"){
              xi=param[1]
```

```
                beta=param[2]
                e_th=(beta+xi*seuils)/(1-xi)
          }
          res=data.frame(e_th)
          res
    }
```

The thresholds passed as a parameter to this function are not the empirical thresholds, but rather the theoretical thresholds, calculated from the functions qlnorm, qgpd, etc.

4.3. Results

The results presented here are intended to illustrate the process. They must carry out supplementary simulations within the context of each particular situation (in terms of underlying distributions as well as the sample size) to retain the most suitable method.

We use the following R code (for the case of GPD distribution, the adaptation to other distribution is simple):

```
library(evir)
xi_th=.75
beta_th=1
n_obs=10^6
X=rgpd(n_obs,xi_th,0,beta_th)
source("C:\\Logiciel R\\EVT\\seuils.r")
seuils=seuils(X)
source("C:\\Logiciel R\\EVT\\e_gpd.r")
res=e_gpd(X,seuils)
source("C:\\Logiciel R\\EVT\\e_hill.r")
res2=e_hill(X,seuils)
source("C:\\Logiciel R\\EVT\\e_emp.r")
res3=e_emp(X,seuils)
source("C:\\Logiciel R\\EVT\\e_th.r")
p=as.vector(1:2)
p[1]=xi_th
p[2]=beta_th
for (i in 1:20){
      seuils[i]=.895+i/200
}
seuils=qgpd(seuils,xi_th,0,beta_th)
res4=e_th("dgpd",p,seuils)
```

Two sets of simulations are presented below: one with 10 ^ 6 trials, effectiveness not available in practice, and the other with 5,000 trials, which is close to the size of the available sample data.

4.3.1. Samples from a Generalized Pareto Distribution

Table 1 – Results based on a sample of 10^6 trials of GPD

	Estimateur GPD			Estimateur Hill		Estimateur empirique	Valeur théorique
Quantile	e_est	xi_est	beta_est	e_est	xi_est	e_est	e_th
90,0%	22,82	0,75	5,66	30,55	0,83	26,79	22,49
90,5%	23,70	0,75	5,88	32,03	0,83	27,87	23,38
91,0%	24,77	0,75	6,11	33,68	0,83	29,05	24,34
91,5%	25,90	0,75	6,37	35,48	0,83	30,35	25,41
92,0%	27,08	0,75	6,68	37,56	0,83	31,79	26,59
92,5%	28,49	0,75	7,00	39,95	0,83	33,39	27,91
93,0%	29,91	0,75	7,40	42,64	0,83	35,19	29,39
93,5%	32,02	0,76	7,76	45,79	0,83	37,22	31,07
94,0%	34,30	0,76	8,21	49,22	0,84	39,54	32,99
94,5%	36,22	0,76	8,82	53,10	0,84	42,21	35,22
95,0%	38,65	0,75	9,52	57,71	0,84	45,35	37,83
95,5%	41,45	0,75	10,37	63,08	0,84	49,08	40,94
96,0%	44,97	0,75	11,38	69,59	0,84	53,59	44,72
96,5%	49,88	0,75	12,55	78,22	0,84	59,20	49,43
97,0%	55,57	0,74	14,21	89,10	0,84	66,37	55,49
97,5%	63,05	0,74	16,29	103,73	0,84	75,94	63,62
98,0%	73,62	0,74	19,32	124,31	0,84	89,47	75,21
98,5%	90,40	0,73	24,02	156,94	0,84	110,38	93,32
99,0%	119,43	0,73	32,73	216,18	0,84	148,05	126,49
99,5%	200,26	0,73	53,76	371,16	0,84	242,86	212,73

Table 2 – Results based on a sample of 5,000 trials of GPD

	Estimateur GPD			Estimateur Hill		Estimateur empirique	Valeur théorique
Quantile	e_est	xi_est	beta_est	e_est	xi_est	e_est	e_th
90,0%	21,33	0,78	4,63	20,56	0,78	29,46	22,49
90,5%	23,60	0,80	4,66	21,70	0,78	30,70	23,38
91,0%	24,83	0,80	4,85	22,18	0,78	32,06	24,34
91,5%	32,37	0,86	4,66	23,54	0,78	33,56	25,41
92,0%	41,01	0,89	4,68	25,04	0,78	35,23	26,59
92,5%	39,54	0,87	5,07	25,95	0,78	37,09	27,91
93,0%	56,79	0,91	5,06	27,68	0,78	39,20	29,39
93,5%	101,55	0,95	5,09	29,85	0,78	41,58	31,07
94,0%	87,50	0,94	5,62	31,26	0,78	44,33	32,99
94,5%	85,19	0,93	6,16	33,58	0,78	47,55	35,22
95,0%	-707,40	1,01	5,96	36,21	0,78	51,33	37,83
95,5%	207,88	0,97	7,07	38,17	0,78	55,88	40,94
96,0%	-188,70	1,04	7,12	42,67	0,78	61,47	44,72
96,5%	-47,04	1,15	7,02	46,96	0,78	68,47	49,43
97,0%	-80,05	1,11	8,83	52,71	0,79	77,63	55,49
97,5%	455,07	0,97	13,12	58,45	0,79	90,23	63,62
98,0%	-237,98	1,06	14,13	74,28	0,79	108,57	75,21
98,5%	-133,02	1,13	17,46	94,40	0,80	137,47	93,32
99,0%	-234,57	1,12	28,61	125,14	0,79	192,11	126,49
99,5%	582,39	0,84	90,45	220,16	0,79	341,64	212,73

First of all we notice the poor quality of empirical estimator, and also that the sample size is small. Indeed, this estimator is far too volatile and sensitive to large values of the sample. The results for the GPD and Hill estimators are not intuitive:

- In the case of a large sample, the GPD estimator dominates the Hill estimator,
- The situation is reversed in the case of a ‘small’ sample.

We may also note that the estimate of the tail parameter is particularly delicate and we can have a correct estimate of the expectation of excesses over a threshold with a bad estimate of ξ .

4.3.2. Sample from a log-normal distribution

In this case, in any discipline, the estimators GPD and Hill do not apply, since we $\xi = 0$; however we use the proposed algorithms mechanically to test their robustness. This gives the following results:

Table 3 – Results based on a sample of 10^6 trials of log-normal distribution

	Estimateur GPD			Estimateur Hill		Estimateur empirique	Valeur théorique
Quantile	e_est	xi_est	Beta_est	e_est	xi_est	e_est	e_th
90,0%	2,82	0,26	2,09	3,13	0,46	6,42	2,81
90,5%	2,87	0,26	2,13	3,23	0,47	6,57	2,83
91,0%	2,91	0,26	2,17	3,33	0,47	6,72	2,85
91,5%	2,95	0,25	2,20	3,45	0,47	6,89	2,87
92,0%	3,00	0,25	2,24	3,57	0,47	7,07	2,90
92,5%	3,05	0,25	2,28	3,70	0,47	7,27	2,93
93,0%	3,11	0,25	2,33	3,85	0,47	7,48	2,96
93,5%	3,16	0,25	2,37	4,01	0,47	7,71	3,00
94,0%	3,23	0,25	2,43	4,19	0,47	7,96	3,04
94,5%	3,30	0,25	2,48	4,38	0,47	8,25	3,09
95,0%	3,38	0,25	2,55	4,59	0,47	8,57	3,15
95,5%	3,48	0,24	2,62	4,84	0,47	8,93	3,21
96,0%	3,58	0,24	2,72	5,12	0,47	9,34	3,29
96,5%	3,70	0,24	2,80	5,46	0,47	9,83	3,38
97,0%	3,84	0,24	2,92	5,85	0,47	10,41	3,49
97,5%	4,01	0,24	3,03	6,36	0,47	11,12	3,62
98,0%	4,24	0,24	3,24	6,98	0,47	12,04	3,80
98,5%	4,54	0,24	3,45	7,86	0,47	13,30	4,04
99,0%	5,00	0,23	3,83	9,19	0,47	15,23	4,40
99,5%	5,92	0,22	4,62	11,76	0,47	18,98	5,10

Table 4 – Results based on a sample of 5,000 trials of the log-normal distribution

	Estimateur GPD			Estimateur Hill		Estimateur empirique	Valeur théorique
Quantile	e_est	xi_est	beta_est	e_est	xi_est	e_est	e_th
90,0%	2,52	0,20	2,02	2,78	0,44	6,09	2,81
90,5%	2,54	0,20	2,05	2,85	0,44	6,22	2,83
91,0%	2,58	0,19	2,09	2,91	0,44	6,36	2,85
91,5%	2,63	0,17	2,18	2,97	0,43	6,51	2,87
92,0%	2,66	0,17	2,21	3,09	0,44	6,67	2,90
92,5%	2,68	0,18	2,21	3,24	0,44	6,84	2,93
93,0%	2,70	0,19	2,19	3,38	0,44	7,02	2,96
93,5%	2,77	0,16	2,31	3,48	0,44	7,23	3,00
94,0%	2,81	0,16	2,35	3,67	0,44	7,45	3,04
94,5%	2,86	0,15	2,42	3,80	0,44	7,70	3,09
95,0%	2,85	0,19	2,32	4,04	0,44	7,97	3,15
95,5%	2,87	0,21	2,27	4,25	0,44	8,27	3,21
96,0%	2,96	0,20	2,35	4,43	0,44	8,61	3,29
96,5%	3,06	0,19	2,47	4,71	0,44	9,01	3,38
97,0%	3,19	0,16	2,66	4,99	0,44	9,49	3,49
97,5%	3,40	0,11	3,03	5,31	0,44	10,10	3,62
98,0%	3,44	0,13	2,99	5,91	0,44	10,85	3,80
98,5%	3,46	0,18	2,84	6,67	0,44	11,85	4,04
99,0%	3,68	0,20	2,93	7,63	0,44	13,30	4,40
99,5%	4,29	0,49	2,19	9,68	0,44	15,93	5,10

We find the very poor results of empirical estimator. In a quite surprising way, the GPD estimator gives good results for the two sample sizes. The results are unconvincing for the Hill estimator.

Overall, the obtained results on a sample of 5,000 values can be quite variable, however, tend to show that the GPD estimator is more robust. The estimator Hill proves to be a delicate quotation to use.

5. Conclusion

The robust estimation of the expected shortfall is a delicate exercise at least where the volume of data available is small. The use of an empirical estimator must be rejected due to its vagueness and volatility.

In a non-parametric context, the GPD estimator provides a generally satisfactory result.

But whatever the retained estimator is (GPD or Hill), its use poses a relatively important risk on a sample of about 5,000 values. In this context, we can consider turning to a parametric model of the general distribution using an explicit expression of the expected shortfall which can then be calculated (at least using an approximation).

The log-normal distribution, Benktender, or the mixed distribution of log-normal / Pareto are the interested models to this topic. Moreover, the adopted model can be calibrated in comparison with the rates of reinsurance policies in surplus.

Bibliography

Danielson J., De Vries C.G. (1997) « Extreme returns, tail estimator and Value-at-Risk », Working paper, London School of Economics.

De Haan L., Rotzen H. (1993) « On the estimation of high quantiles », *Journal of Statistical Planning and Inference* 35 (1), 1-13.

Embrechts P., Kluppelberg C., Mikosch T. (1997) *Modelling extremal events*, Berlin: Springer Verlag.

Garrido M. (2002) *Modélisation des évènements rares et des quantiles extrêmes, méthodes de sélection de modèle pour les queues de distribution*, Thèse de doctorat, Université J. Fourier.

Hosking J. R., Wallis J. R. (1987) « Parameter and quantile estimation for the generalized pareto distribution », *Technometrics* 29, 339-49.

Partrat C., Besson J.L. (2005) *Assurance non-vie, modélisation, simulation*, Paris : Economica.

Pickands J. (1975) « Statistical inference using extreme order statistics », *Annals of Statistics* 3, 119-31.

Planchet F., Thérond P.E. (2007) « Model risk and determination of economic capital in the Solvency 2 project », *Proceedngs of the 17th AFIR colloquium*.

Resnick S., Starica C. (1997) « Smoothing the Hill estimator », *Advanced Applied Probabilities* 29, 271-93

Thérond P.E., Planchet F. (2007) « Provisions techniques et capital de solvabilité d'une compagnie d'assurance : méthodologie d'utilisation de Value-at-Risk », *Assurances et gestion des risques* 74 (4), 533-63.