ADVANCED OPERATIONAL RISK MODELLING IN BANKS AND INSURANCE COMPANIES

CARLA ANGELA (Sapienza - Università di Roma, carla.angela@uniroma1.it)
ROSSELLA BISIGNANI (Sapienza - Università di Roma, rossella.bisignani@uniroma1.it)
GIOVANNI MASALA (Università di Cagliari, gb.masala@tiscali.it)
MARCO MICOCCI (Corresponding Author, Università di Cagliari, micocci@unica.it)

Abstract

The aim of this paper is to measure Operational Risk in financial institutions when historical data are available starting from a fixed threshold.
To quantify the Operational Risk we apply the Loss Distribution Approach (LDA), a frequency/severity approach widely used in the actuarial models. Risk measures like Value at Risk (VaR) and Expected Shortfall (ES) are used for determining the risk capital necessary to cover the Operational Risk.
The dependence among the events the Operational Risk is arisen from, has been taken into account using the Copula Function while Extreme Value Theory (EVT) has been used to model the right tail of the severity of loss distributions.
The Expectation and Maximization (EM) algorithm has been applied to estimate the parameters of the frequency and severity of loss distributions when only their left truncated distributions are available.
The example concerns an application of the proposed model for evaluating the risk capital for a single financial institution.
To this aim we have used, as empirical observations, the OpData® dataset supplied by OpVantage®.

Keywords: Operational Risk, Extreme Value Theory, Value at Risk, Expected Shortfall, Monte Carlo, Copula Function, EM algorithm.
1. INTRODUCTION

With the regulatory spotlight on Operational Risk management, there has been ever increasing attention devoted to the quantification of Operational Risk.

The Operational Risk potential devastating power has been shown by many large operational losses; some of the best known Operational Risk incidents are the $9 billion loss of Banco National due to credit fraud in 1995, the $2.6 billion loss of Sumimoto Corporation due to unauthorized trading activity in 1996, the $1.7 billion loss and subsequent bankruptcy of Orange County due to unauthorized trading activity in 1998, the $1.3 billions trading loss causing the collapse of Barings Bank in 1995, the $0.75 billion loss of Allied Irish Bank in 2002, the loss of $2 million of Prudential Insurance of America in 2002.

The new regulatory framework in banking sector (Basel II) and the project of the new solvency regime in insurance sector (Solvency II) recognizes the importance of Operational Risk by requiring its explicit treatment with the determination of a specific capital requirement.

There is no generally accepted definition of Operational Risk in the financial community. In this paper we refer to the definition proposed by the Basel Committee on Banking Supervision in 2001: “the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events”. This definition has been adopted for the insurance sector until now.

In this categorization Operational Risk includes the following event types: business disruption and system failures; clients, products and business practice; damage to physical assets; employment practice and workplace safety; execution delivery and process; external fraud; internal fraud.

In this paper, we develop a comprehensive model to quantify the capital charge necessary to cover the Operational Risk in a financial institution.

The proposed model belongs to the class of the “Loss Distribution Approach” (LDA). LDA is a frequency/severity model widely used in many fields of the actuarial practice.

In this model the frequency and severity of loss distributions is determined for each of the events of loss identifying the best fitting distribution of empirical data (see Moscadelli (2004); De Fontnouvelle (2003)); then we apply copula functions to reflect the dependence amongst the different events dealing with operational losses (see for example Di Clemente-Romano (2003), Reshetar (2004)). The Operational Risk capital charge is estimated quantifying Value at Risk and Expected Shortfall of the joint distribution of losses (see Rockafellar-Uryasev (2002), (2003)), estimated using Monte Carlo simulation.

In order to better estimate the eventually fat tails of the severity distributions the Extreme Value Theory (Embrechts et al (1997), (2001), (2003), (2004), (2005)) is also included in the model.

The paper includes an application aimed to the evaluation of Operational Risk capital charge.

The numerical example has been performed using the OpData dataset supplied by OpVantage that is the largest dataset existing on the market concerning operational losses.

In the OpData dataset only Operational Risk losses exceeding $1 million are collected. This limitation may be over-passed using appropriate algorithms to solve the problem of estimating the severity/frequency distribution with truncated data. In the model we adopt the EM algorithm (compare Bee (2005)).

The paper has the following structure:

Section 2 methodology overview; here we describe our model of evaluation. In particular section 2.1 deals with the Loss Distribution Approach and explains the structure of a typical model of this family; introduces Copula Functions to model dependencies amongst different risk types, gives a briefly overview of the risk measures (VaR and Expected Shortfall) chosen to quantify the capital charge able to cover operational loss risk and introduces the Extreme Value Theory.
Section 2.2 anticipates the theme of distribution estimation with truncated data by presenting the EM algorithm that is used in section 3.

Section 3 a model application. The model is developed using the OpData dataset as input data. The methodologies, procedures and algorithms described in section 2 are performed on empirical data to test the model feasibility. The results represent the quantification of the Operational Risk capital charge in a hypothetical financial institution.


In this section the several “technical blocks” that compose our model are briefly described. In particular, we present the main characteristics of the Loss Distribution Approach (LDA), some useful risk measures to quantify the Operational Risk capital charge and some basic elements about copula functions and Extreme Value Theory; we also introduce the EM algorithm that is a useful tool to manage truncated datasets.

2.1 An overview of Loss Distribution Approach, Copula Function and Extreme Value Theory (EVT).

As already said in the introduction, our model may be considered a Loss Distribution Approach (LDA) for the quantification of Operational Risk.

To apply the Loss Distribution Approach we need to determine the loss frequency and the loss severity distributions for each event type the Operational Risk arisen from. The aggregate distribution for each event type is then obtained as the convolution of the frequency and the severity of loss distributions.

Let us denote $N_i$ the loss frequency for event type $i$, whose distribution function is $p$.

The loss frequency distribution is given by the following expression:

$$P(n) = \sum_{k=0}^{n} p(k)$$  \hspace{1cm} (2.1.1)

We denote $Y_{ik}$ the loss severity associated to the $k$-th event for event type $i$ and with $F$ the probability distribution of the aggregate losses.

In the Loss Distribution Approach framework, the aggregate annual loss $Y_i$, for event type $i$, can be obtained as the sum of the stochastic number $N_i$ of events occurred in one year with severity $Y_{ik}$:

$$Y_i = \sum_{k=0}^{N_i} Y_{ik}$$  \hspace{1cm} (2.1.2)

The probability distribution function $G_i(y)$ for the aggregate loss $Y_i$ is the following:

$$G_i(y) = \Pr(Y_i \leq y) = \begin{cases} \sum_{k=1}^{\infty} P(N_i = k) \times F^{k*}(y_i) & y > 0 \\ P(N_i = 0) & y = 0 \end{cases}$$  \hspace{1cm} (2.1.3)

where $F(y_i)$ is the probability that the $k$-th loss is $y_i$ while $F^{k*}$ represents the $k$-th convolution of $F$.

We assume the following hypothesis:

- losses $Y_{ik}$ are i.i.d. random variables;
- loss frequencies and loss severities are independent random variables.
The analytical expression for $G_i(y)$ can generally not be obtained. We must then set up a Monte Carlo simulation in order to generate a high number of simulated aggregated losses.

In order to apply correctly tail risk measures such as Value at Risk (VaR) or Expected Shortfall (ES), we must estimate efficiently the tails of severity distributions. Indeed, severities data have generally heavy tailed distributions. In this situation, Extreme Value Theory (EVT) permits to take into account large losses in a correct way.

We recall here that if we fix a confidence level $\alpha \in (0,1)$, the Value at Risk (VaR) and the Expected Shortfall (ES) for the loss random variable $X$ at probability level $\alpha$ with cdf $F_X(x)$ is defined as:

$$\text{VaR}_\alpha = \min \{ \xi \in R : F_X(\xi) \geq \alpha \}$$

and

$$\text{ES}_\alpha = E[X \mid X > \text{VaR}_\alpha]$$

In the first definition, the VaR turns out to be the left endpoint of the nonempty interval consisting of the values $\xi$ such that $F_X(\xi) = \alpha$. Besides we deduce that the probability that $X \geq \text{VaR}_\alpha$ equals $1 - \alpha$. Consequently, the ES is defined as the conditional expectation of the loss associated with $X$ which are equal to VaR or greater.

In this model we also use the EVT that is an approach useful to correctly estimate heavy tails distributions.

EVT aims to describe the distributions of rare events focusing on the tail of the distribution. EVT is then a powerful tool for managing losses due to rare events and inadequacy of internal controls (Low Frequency High Impact Events).

In general, extreme events can be treated in the two following ways:

- considering the maximum (or the minimum) of the analyzed random variable in consecutive periods (for example months or years) in which the time horizon has been subdivided ("Block Maxima" method);
- considering the value that random variable assumes over a given threshold ("Peaks Over Threshold method", POT).

In the rest of the paper, we will consider exclusively the POT method.

Traditionally, the total capital charge for Operational Risk is obtained by summing capital charges for each event type. This procedure implies a perfect dependence between each event type (in other words, aggregate losses coming from different event types are comonotonic random variables). This unrealistic assumption implies the overestimation of the total capital charge. The consideration of a realistic dependence structure, through Copula Functions produces a lower capital charge (diversification effect).

To this aim we choose the t-Student Copula.

When we consider more than two marginal distributions, the Archimedean family is no more efficient because this copula has only one parameter (thus the same dependence structure between each couple of marginal distributions is assumed). The t-Student Copula is then more appropriate as the tail dependence is captured (through the degrees of freedom). For these reasons, t-Student Copula is widely used in literature.

The main characteristics of the t-Student Copula are here briefly described. Its parameters are the correlation matrix $R$ and the degrees of freedom $\nu$.

Let $Y$ be a vector with a $n$- variate standardized Student’s $t$-distribution with $\nu$ degrees of freedom, and covariance matrix $\frac{\nu}{\nu - 2} R$ (for $\nu > 2$). It can be defined in the following way:

$$X = \frac{\sqrt{\nu}}{\sqrt{S}} Y$$
where $S \sim \chi^2_\nu$ (the chi-square distribution) and the random vector $Y \sim N_\nu(0, \mathbf{R})$ are independent. The copula of vector $Y$ is the t-Student Copula with $\nu$ degrees of freedom. The analytical representation is the following:

$$C_{\nu, \mathbf{R}}(u_1, \ldots, u_n) = t^n_{\nu, \mathbf{R}} \left( t^{-1}_\nu(u_1), \ldots, t^{-1}_\nu(u_n) \right)$$

or equivalently:

$$C_{\nu, \mathbf{R}}(u_1, \ldots, u_n) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \frac{\left( \frac{\nu + n}{2} \right) \left| \mathbf{R} \right|^{1/2}}{\left( \frac{\nu}{2} \right)^{1/2} \left( \nu \cdot \pi \right)^{n/2}} \left( 1 + \frac{1}{\nu} \mathbf{u}^\top \cdot \mathbf{R}^{-1} \cdot \mathbf{u} \right) du_1 \, du_2 \cdots du_n$$

where $\mathbf{R}_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii} \cdot \Sigma_{jj}}}$ for $i, j \in \{1, \ldots, n\}$ are the correlations. We also indicate with $t^n_{\nu, \mathbf{R}}$ the multivariate cumulative distribution function of the random vector $\sqrt{\frac{\nu}{S}} \cdot Y$, where the random variable $S \sim \chi^2_\nu$ and the random vector $Y$ are independent. Besides, $t_{\nu}$ (cdf of the standard univariate Student distribution) denotes the margins of $t^n_{\nu, \mathbf{R}}$.

Finally, the t-Student Copula has the following density:

$$c(u_1, \ldots, u_n; \nu, \mathbf{R}) = \frac{\Gamma \left( \frac{\nu + n}{2} \right) \left[ \frac{\nu}{2} \right]^n \left( 1 + \frac{1}{\nu} \omega^\top \cdot \mathbf{R}^{-1} \cdot \omega \right)^{-\frac{\nu + n}{2}}}{\sqrt{|\mathbf{R}|} \cdot \Gamma \left( \frac{\nu}{2} \right) \left[ \frac{\nu + 1}{2} \right]^n \prod_{i=1}^{n} \left( 1 + \frac{\omega_i^2}{\nu} \right)^{-\frac{\nu}{2}}$$

where $\omega = t^{-1}_\nu(u_i)$.

If we choose a t-Student Copula, the degree of freedom $\nu$ can be evaluated with a log likelihood estimator.

Once the marginal distributions have been chosen, this can be accomplished through a Maximum Likelihood Estimation (MLE).

To generate pseudo-random numbers from the t-Student Copula we use the following algorithm:

- find the Cholesky decomposition $\mathbf{A}$ of the correlation matrix $\mathbf{R}$;
- simulate $n$ independent random variates $\mathbf{z} = (z_1, \ldots, z_n)$ from the standard normal distribution;
- simulate a random variate $s$ from $\chi^2_\nu$ distribution, independent of $\mathbf{z}$;
- determine the vector $\mathbf{y} = \mathbf{A} \times \mathbf{z}$;
- set $\mathbf{x} = \frac{\mathbf{y}}{\sqrt{S}}$;
- determine the components $u_i = t_{\nu}(x_i) \quad i = 1, \ldots, n$. The resultant vector is $(u_1, \ldots, u_n)^\top \sim C_{\nu, \mathbf{R}}^n$. 
2.2 EM algorithm.

The model is designed taking in mind the availability of truncated data. In the Operational Risk losses database we can find in the markets the lower losses are generally not reported. For internal database the threshold is generally $10,000 while for external database this threshold is generally $1,000,000. The problem of determining the parameters of the distribution which represents empirical loss frequencies and loss severities is obviously influenced by the presence of “truncated data”. If the database consist of truncated data the traditional best fitting techniques based on maximum likelihood estimation or percentile matching, can not be applied. In this case it is necessary to apply the so called EM algorithm to evaluate the parameters of the unknown complete distribution.

At this aim the conditional distributions and maximum likelihood estimation techniques involving truncated data are considered.

We assume the standard hypothesis that the severity distribution follows a lognormal distribution. The maximum likelihood estimation in presence of incomplete data can be carried out using the EM algorithm (Expectation-Maximization) (see Dempster et al (1977)). The estimation of the parameters of the complete distribution is done by maximizing the loglikelihood function (see Frachot at al (2003)).

Let \( Y = (N_{NT}, y) \) be the observed incomplete sample, where \( N_{NT} \) denotes the number of observed data and \( y = (y_1, y_2, \ldots, y_{N_{NT}}) \) is the observed data vector. We also denote \( X = (N_T, x) \) the missing data, where \( N_T \) represents the number of missing data and \( x = (x_1, x_2, \ldots, x_{N_T}) \) is the vector of missing data.

The complete data will assume then the following form:

\[
Z = (X, N_T, Y, N_{NT}) \text{ with } \begin{align*}
Y &= \{z_i : z_i \geq s\} ; \\
x &= \{z_i : z_i < s\} ; \\
N_{X_T} &= \# \{z_i : z_i \geq s\} ; \\
N_T &= \# \{z_i : z_i < s\} \text{ and } s \text{ is the truncation threshold. (see Bee (2005))}
\end{align*}
\]

Let us denote:

\( l = l(\theta) \) the loglikelihood function for \( Y \);

\( \theta \) the parameters vector we have to estimate;

\( l_c = l_c(\theta) \) the loglikelihood function for the complete data;

The two steps of the algorithm are:

1) E-step: estimate the expected conditional value of the loglikelihood function \( l_c = l_c(\theta) \) for the observed sample \( y \) and the current value for \( \theta \).

Let denote \( \theta_0 \) the initial value of the vector parameters (determined arbitrarily), we must then estimate:

\[
Q(\theta; \theta_0) = E_{\theta_0} \{l_c(\theta) | y\} \tag{2.2.1}
\]

2) M-step: the previous expression must be maximized with respect to \( \theta \):

\[
Q(\theta; \theta_0) = \max Q(\theta; \theta_0) \tag{2.2.2}
\]

We repeat the procedure by substituting \( \theta_0 \) with \( \theta \).

These two steps are then repeated until convergence occurs.

We describe now the EM algorithm in the hypothesis of lognormal distribution.
Given a sample \( w_1, w_2, \ldots, w_{N_{NT}} \) coming from a lognormal truncated distribution with parameters \( \mu \) and \( \sigma^2 \) (\( N_{NT} \) is the number of observed data) we determine \( y_i = \log(w_i) \approx N(\mu, \sigma^2) \) in order to transform the data into a normal distribution. The complete distribution function is:

\[
L_C(\mu, \sigma^2; z) = \prod_{i=1}^{N} L(\mu, \sigma^2; z_i)
\]

where \( \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+ \) and \( L(\mu, \sigma^2; z_i) \) is the normal distribution function estimated in \( z_i \).

The observed distribution function is:

\[
L_{obs}(\mu, \sigma^2; s, y) = \prod_{i=1}^{N_{NT}} \left[ \frac{1}{1 - \Phi(s^*)} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma}(y_i - \mu)^2} \right]
\]

with \( \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}^+, s^* = \frac{s - \mu}{\sigma} \).

In the implementation of EM algorithm, for calculating:

\[
L_C(\mu, \sigma^2; z) = \prod_{i=1}^{N} L(\mu, \sigma^2; z_i)
\]

we use the fact that the joint distribution of \( Y, N_T \) and \( X \) can be expressed in the following way:

\[
f(y, N_T, x) = f(y) \cdot f(N_T|y) \cdot f(x|N_T, y)
\]

As the function \( f(y) \) is already known, it remains to determine \( N_T|y \) and \( x|N_T, y \).

Given the complete data set \( Z \) from a normal distribution with parameters \( \mu \) and \( \sigma^2 \), it has been proved (McLachlan, Krishman, 1996) that \( N_T|y \) follows a negative binomial distribution with parameters \( N_{NT} \) and \( K = \Pr(Z > s) = 1 - \Phi(s^*) \).

Consequently, the expected conditional value is:

\[
E(N_T|y) = N_{NT} \cdot \frac{1 - K}{K}
\]

The distribution of the missing data \( X \) is a right truncated normal distribution with density function:

\[
f_X(x) = \frac{1}{\Phi(s^*)} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma}(x - \mu)^2}, \ x < s
\]

The estimation of \( E \left[ L(\mu, \sigma^2; s, x, N_T, \mu(o), \sigma(o^2)) \right] \) is rather easy due to the fact that the loglikelihood function is linear in \( X \) and \( X^2 \), so that we have to estimate:

\[
E(\bar{X}[\mu(o), \sigma(o^2), y_1, y_2, \ldots, y_{N_{NT}}])
\]

\[
E(\bar{X^2}[\mu(o), \sigma(o^2), y_1, y_2, \ldots, y_{N_{NT}}])
\]

The estimation of (2.2.8) and (2.2.9) can be performed thanks to the known formulas for the expected value and the variance of the right truncated normal distribution.
We set \( \alpha(s^*) = \frac{\phi(s^*)}{\Phi(s^*)} \) so that:

\[
E\left(X \mid \theta^{(i)}, y\right) = \mu^{(i)} - \sigma^{(i)} \cdot \alpha\left(s^{(i)}\right)
\]

\[
E\left(X^2 \mid \theta^{(i)}, y\right) = \sigma^{2(i)} \cdot \left[1 - s^{(i)} \cdot \alpha\left(s^{(i)}\right) - \left(\alpha\left(s^{(i)}\right)\right) \right]^{2} + \left[ E\left(X \mid \theta^{(i)}, y\right) \right]^{2}
\]

(2.2.10)

(2.2.11)

Finally, the E-step is given by:

\[
E\left(N \mid y\right) = \hat{N}_T = N_{NT} \cdot \frac{1 - K}{K}
\]

\[
E\left(X \mid \theta^{(i)}, y\right) = \mu^{(i)} - \sigma^{(i)} \cdot \alpha\left(s^{(i)}\right)
\]

\[
E\left(X^2 \mid \theta^{(i)}, y\right) = \sigma^{2(i)} \cdot \left[1 - s^{(i)} \cdot \alpha\left(s^{(i)}\right) - \left(\alpha\left(s^{(i)}\right)\right) \right]^{2} + \left[ E\left(X \mid \theta^{(i)}, y\right) \right]^{2}
\]

(2.2.11)

Now the M-step requires the formulas for the maximum likelihood, in the case of normal distribution, for the missing data coming from (2.2.10) and (2.2.11):

\[
\mu^{(i+1)} = \frac{1}{N_{NT} + N_T^{(i)}} \sum_{i=1}^{N_X} y_i + N_T^{(i)} \cdot E\left(X \mid \theta^{(i)}, y\right)
\]

(2.5.12)

\[
\sigma^{2(i+1)} = \frac{1}{N_{NT} + N_T^{(i)}} \sum_{i=1}^{N_X} y_i^2 + N_T^{(i)} \cdot E\left(X^2 \mid \theta^{(i)}, y\right) - \left(\mu^{(i+1)}\right)^{2}
\]

(2.5.13)

The parameters are then obtained by iterating (2.2.10), (2.2.11), (2.2.12) and (2.2.13) until convergence.

3. A Model Application

An application of the model above described has been developed with the aim to calculate the Operational Risk capital charge in a financial institution (bank or insurance company).

The input data derive from OpData, an operational losses database supplied by OpVantage, a division of Fitch Risk Management.

The data are collected from public sources and in the database only losses, whose amounts exceed a truncation threshold of $1 million, are registered on the period 1972-2006.

In this application we consider the database on the period 1994-2006, being the previous data not statistically significant.

In the OpData, operational losses are categorized according the Basel Committee’s event types classification:

1) Business Disruption and System Failures;
2) Clients, Products and Business Practice;
3) Damage to Physical Assets;
4) Employment Practice and Workplace safety;
5) Execution Delivery and Process Management;
6) External Fraud;
7) Internal Fraud.
Due to the lack of the loss data in event types 1) and 3), we have taken in account five event types data.

For each loss the following information are available:
- classification in event type;
- firm name;
- loss event description;
- loss amount in local currency;
- loss amount in dollars;
- loss amount in current value dollars (based on CPI);
- loss data;
- country;
- total assets of the firm.

The model provides the following steps:

- estimation of the parameters of frequency and severity distributions, for each event type applying the EM algorithm. The frequency of loss arising from each event type is assumed to be a Poisson distribution while the lognormal distribution is used to model the severity of loss.
- estimation of the aggregate loss distribution, for each event type, via Monte Carlo simulation;
- quantification of Operational Risk capital charge, for each event type, through risk measure as Value at Risk and Expected Shortfall;
- quantification of the total Operational Risk capital charge in different hypothesis:
  - perfect dependence (comonotonicity) among event types. The total capital charge is then obtained by summing capital charges for each event type and it results overestimated
  - independence between event types. This assumption leads to underestimation of the total capital charge;
  - realistic dependence structure through a t-Student copula.

For efficiently modelling the right tail of the severity distribution we repeat the above steps using Extreme Value Theory (Di Clemente - Romano (2003) Embrechts (2003), (2005) Moscadelli (2004)). Therefore we model the severity distribution using the lognormal distribution (in the left tail and in the centre) and the Generalized Pareto Distribution (GPD) for the right tail.

The database at our disposal is structured as follows: for each year we know the determination $k$ of the random variable $N$ “number of loss events in one year” and consequently we have at our disposal $k$ determinations of the random variable $Y_i$.

Then, for each financial institution, we normalize its loss amount $y_{id}$ dividing it by the its total asset $A$:

$$\hat{y}_{id} = \frac{y_{id}}{A};$$

in this way we shall obtain results that are expressed as percentages of the total assets of the firm.

We apply then the EM algorithm to estimate the parameters of frequency and severity distributions (see table 3.1).
Table 3.1 Parameters Estimation of severity and frequency distributions for each Event Type.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>Severity (lognormal)</th>
<th>Frequency (Poisson)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clients, Products and Business Practices</td>
<td>$\mu = -10.425$ ; $\sigma = 2.286$</td>
<td>$\lambda = 37.130$</td>
</tr>
<tr>
<td>2. Employment Practices and Workplace Safety</td>
<td>$\mu = -12.139$ ; $\sigma = 2.066$</td>
<td>$\lambda = 6.686$</td>
</tr>
<tr>
<td>3. Execution, Delivery and Process Management</td>
<td>$\mu = -11.456$ ; $\sigma = 2.039$</td>
<td>$\lambda = 6.678$</td>
</tr>
<tr>
<td>4. External Fraud</td>
<td>$\mu = -10.824$ ; $\sigma = 1.975$</td>
<td>$\lambda = 13.741$</td>
</tr>
<tr>
<td>5. Internal Fraud</td>
<td>$\mu = -10.692$ ; $\sigma = 2.143$</td>
<td>$\lambda = 18.971$</td>
</tr>
</tbody>
</table>

The aggregate loss distribution, for each event type $i$, is obtained as a convolution of the frequency and severity distributions with the Monte Carlo simulation method.

We can estimate operational losses at firm level, for each event type $i$, dividing aggregated losses by the number of firms that suffered from event type $i$.

We consider Value at Risk and Expected Shortfall as the capital amount needed to cover Operational Risk. Capital charges, at different confidence levels, for each event type are reported in tables 3.2 and 3.3 (Monte Carlo simulation with 100,000 replications).

Table 3.2 Value at Risk for Event Type.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>VaR 95%</th>
<th>VaR 99%</th>
<th>VaR 99.9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clients, Products and Business Practices</td>
<td>0.000564</td>
<td>0.001309</td>
<td>0.004634</td>
</tr>
<tr>
<td>2. Employment Practices and Workplace Safety</td>
<td>0.000045</td>
<td>0.000116</td>
<td>0.000411</td>
</tr>
<tr>
<td>3. Execution, Delivery and Process Management</td>
<td>0.000051</td>
<td>0.000133</td>
<td>0.000477</td>
</tr>
<tr>
<td>4. External Fraud</td>
<td>0.000092</td>
<td>0.000210</td>
<td>0.000637</td>
</tr>
<tr>
<td>5. Internal Fraud</td>
<td>0.000178</td>
<td>0.000431</td>
<td>0.001260</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0.000929</td>
<td>0.002199</td>
<td>0.007419</td>
</tr>
</tbody>
</table>

Table 3.3 Expected Shortfall for Event Type.

<table>
<thead>
<tr>
<th>Event Type</th>
<th>ES 95%</th>
<th>ES 99%</th>
<th>ES 99.9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Clients, Products and Business Practices</td>
<td>0.001209</td>
<td>0.002871</td>
<td>0.009721</td>
</tr>
<tr>
<td>2. Employment Practices and Workplace Safety</td>
<td>0.000103</td>
<td>0.000243</td>
<td>0.000735</td>
</tr>
<tr>
<td>3. Execution, Delivery and Process Management</td>
<td>0.000121</td>
<td>0.000301</td>
<td>0.001043</td>
</tr>
<tr>
<td>4. External Fraud</td>
<td>0.000183</td>
<td>0.000398</td>
<td>0.001127</td>
</tr>
<tr>
<td>5. Internal Fraud</td>
<td>0.000366</td>
<td>0.000807</td>
<td>0.002224</td>
</tr>
<tr>
<td>TOTAL</td>
<td>0.001982</td>
<td>0.004620</td>
<td>0.014850</td>
</tr>
</tbody>
</table>

Finally, we model the dependence structure among event types by using a t-Student Copula. The MLE estimation for the optimal degrees of freedom of the copula gives $\nu = 5$. 


Figure 3.1 The estimation of the optimal degrees of freedom.

To apply the chosen copula we must assess the correlation matrix of the risk events. We cannot use an empirical correlation matrix (i.e. inferred from historical data) due to database truncation; for this reason, we consider a theoretical correlation matrix based on qualitative considerations. The following table shows the correlation matrix.

Table 3.4 Qualitative Correlation Matrix
(medium correlation = 0.35, high correlation = 0.55)

<table>
<thead>
<tr>
<th>Qualitative correlations among events</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 (medium)</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3 (high)</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4 (zero)</td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5 (high)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

The aggregate loss distribution with Copula Function is obtained through algorithms presented in section 2.1.

In the table 3.5 the total Operational Risk capital charge for the t-Student Copula dependence structure together with the Comonotonicity and Independence cases are shown.

Table 3.5 Operational Risk Capital Charge (% of total asset A).

<table>
<thead>
<tr>
<th>OR Capital Charge</th>
<th>VaR 95%</th>
<th>VaR 99%</th>
<th>VaR 99.9%</th>
<th>ES 95%</th>
<th>ES 99%</th>
<th>ES 99.9%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Comonotonicity</td>
<td>0.000929</td>
<td>0.002199</td>
<td>0.007419</td>
<td>0.001982</td>
<td>0.004620</td>
<td>0.014850</td>
</tr>
<tr>
<td>t-Student copula</td>
<td>0.000836</td>
<td>0.001932</td>
<td>0.005991</td>
<td>0.001694</td>
<td>0.003793</td>
<td>0.010647</td>
</tr>
<tr>
<td>Independence</td>
<td>0.000755</td>
<td>0.001566</td>
<td>0.004839</td>
<td>0.001437</td>
<td>0.003139</td>
<td>0.009965</td>
</tr>
</tbody>
</table>

As it is expected for, considering the dependencies among event types in case of application of t-Student Copula, the Operational Risk capital charge, expressed as percentage of total asset, results always between
the minimum values, assumption of independence, and the maximum values obtained in the assumption of
perfect dependence (comonotonicity).
The results obtained may be refined with the Extreme Value Theory; it can be used for modelling
efficiently the right tail of the severity distribution.
A right tail analysis suggests that only for the event type Internal Fraud, the lognormal distribution
underestimates the large loss probability (see figure 3.2).

![Figure 3.2 Lognormal distribution (line), empirical distribution (dots).](image)

We may “correct” the observed underestimation by modelling the right tail with the Generalized Pareto
Distribution (POT method).
In order to fit the GPD on our data we perform the following steps:

- find the appropriate threshold \( k \);
- determine loss excesses, (loss amounts over the threshold minus the threshold);
- estimate GPD parameters \( \xi \) and \( \sigma \) from the excesses.

According to Figure 3.2, the large loss probability is underestimated by lognormal distribution over the
threshold 0.0007174 (95-th percentile).
Maximum Likelihood Method has been used to estimate GPD parameters (see table 3.6).

| Table 3.6 GPD parameters Internal Fraud. |
|-----|-----|-----|
| \( k \) | \( \xi \) | \( \sigma \) |
| 0.0007174 | 0.817819 | 0.001319 |

The best fitting between GPD \((\xi, \sigma)\) and empirical data are reported in figure 3.3

![Figure 3.3 Internal Fraud. Lognormal Distribution (dark line), Empirical Distribution (dots), Generalized Pareto Distribution (clear line).](image)

The results are reported in table 3.7.
Comparing the results of table 3.7 with those in table 3.5, we can appreciate how much is relevant to choose the more appropriate probability distribution to fit the right tails.

### 4. Conclusions.

The application is developed using Operational Risk data of financial institutions, based in different countries, with different business dimension, with different risk profile. For this reason we have to interpret it mainly as an exercise done to check the model feasibility and to understand the difficulties the evaluator may meet in practice. But at the same time the model results let us to underline how is relevant to consider the dependence among the events bringing to Operational Risk and its impact on the Operational Risk capital charge. Analogue comment is valid for the technical consideration of the fat tails of the severity of loss distributions.

In addition we find that the possibility to quantify the risk capital when only left truncated loss distributions are available is relevant from the practical point of view. It is well known that the most common concern of risk managers refers to the lack of empirical data. With this model we intend to suggest that, as first approach, the financial institutions could start to collect Operational Risk data over a fixed threshold. A last consideration: we calculate (table 3.5 and 3.7) the Operational Risk capital charge as a percentage of the total asset at different confidence levels. This means the single financial institution could measure its own Operational Risk capital charge simply multiplying this percentage by its total asset. We could think that this kind of evaluation be carried on for example at the country level and the results utilized as a benchmark.

### References


[75] MARKT, (2001), Note to the Solvency Subcommittee of the Insurance Committee, Solvency II: Presentation of the proposed work, MARK/2027/01 EC DG Internal Market.


[77] MARKT, 2002, Note to the Solvency Subcommittee, Risk Models of Insurance Companies or Groups, MARK/2515/02 EC DG Internal Market.

[78] MARKT, 2002, Discussion note to the Members of the IC Solvency Subcommittee, Current and Future Solvency Work in the IAIS and within the Actuarial Profession from a Solvency II point of View, MARKT/2520/02 EC DG Internal Market.

[79] MARKT, 2003, Design of a Future Prudential Supervisory System in the EU: Recommendations by the Commissions Services, MARKT/2509/03 EC DG Internal Market.


