



Practical Calculation of Expected and Unexpected Losses in Operational Risk by Simulation Methods

Enrique Navarrete ¹

Abstract: This paper surveys the main difficulties involved with the quantitative measurement of operational risk and proposes simulation methods as a practical solution for obtaining the aggregate loss distribution. An example that calculates both expected and unexpected losses as well as operational risk VAR is provided.

Keywords: Operational Risk, Loss Distribution, Operational Risk VAR, Simulation Methods, Basel II.

Introduction

One of the most difficult tasks in risk management is to set the appropriate level of capital to cover unexpected losses in banks and other financial institutions. Whereas expected losses can be described as the “usual” or average losses that an institution incurs in its natural course of business, unexpected losses are deviations from the average that may put an institution’s stability at risk. Not only risk managers are worried about these types of losses, but also regulators and financial supervisors, hence international standards are being continuously developed and improved to prevent institutions from going bankrupt due to these large potential losses. The most widespread of these standards is *The New Basel Capital Accord*, also known as Basel II.²

Whereas some standard models exist in credit risk to calculate unexpected losses (see, for example, Navarrete 2005a), in operational risk the calculation of unexpected loss is more difficult since one usually takes into account the *frequency* and *severity* components of the loss distribution separately. *Frequency* refers to how often a loss event happens, and is measured in terms of number of events per time units. It is described by a *discrete* distribution. *Severity* depends on the monetary impact of the event, and is described by a *continuous* distribution. In operational risk both components have to be considered separately, since there exist loss events with *low frequency* but *high severity* (eg. catastrophes, damage to physical assets); on the other hand, there are plenty of *high-frequency, low severity* events (eg. small credit frauds, accounting errors, etc.). Obviously, there is a huge range of events in-between.

¹ CEO, Scalar Consulting. Actively involved as risk advisor and consultant to private and public banks, he also collaborates closely with financial regulators. He can be reached at enavarrete@gruposcalar.com

² BCBS (2004).

Calculation of Aggregate Loss Distributions and Capital Requirements for Operational Risk

To establish the appropriate level of capital to cover unexpected losses due to operational risk one first has to establish an adequate *confidence level*. A confidence level is a statistical concept which corresponds, intuitively, to the probability that an institution will not go bankrupt or fail in some business line due to extreme losses. Obviously, one would like to establish confidence levels close to 100 %. In practice, however, this is not possible since loss distributions are never perfectly identified using (usually incomplete) historical data, and even if we could perfectly identify these loss distributions, the level of capital required would be too high (and costly). Nevertheless, the confidence levels used in risk management usually lie in the range from 95 % to 99 % and higher.³

Once we have defined the confidence level at which we would like to cover unexpected losses, the calculation of the corresponding amount of capital involves the following steps:

- i) Frequency and severity distributions are identified from the data;
- ii) Both distributions are combined to obtain an *aggregate loss distribution*;
- iii) *Operational Value at Risk (VAR)* is obtained by taking the *percentile* of the aggregate loss distribution at the desired confidence level.

Unexpected loss is the difference between VAR and expected loss, as Figure 1 shows. This is the amount of capital that the institution should establish to cover unexpected losses for operational risk corresponding to the desired confidence level.⁴

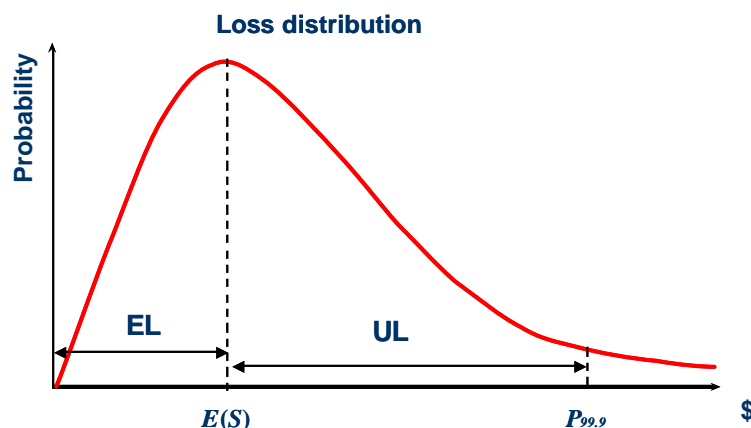


Figure 1: Expected Loss (EL), Unexpected Loss (UL) and VAR ($P_{99,9}$) at the 99,9 % confidence level

³ Basel II suggests 99,9 %. This is equivalent to taking the 10th largest loss in 10,000 losses obtained by running a simulation, as will be discussed below.

⁴ Note that in Figure 1 *expected loss* is simply the *expected value* (mean) of the distribution, $E(S)$, VAR at the 99,9 % level is the *percentile* at such level, $P_{99,9}$, and unexpected loss (UL) is the difference between them.

We should note that one usually establishes a prudential level of capital not for the bank as a whole but for specific types of *loss events* (such as internal fraud, external fraud, etc.) and for its different *business lines*. The example that will be provided below (using simulation) will calculate the amount of capital needed to cover potential losses due to *external fraud* in the *credit card* business line of a small banking institution.⁵

The main difficulty of the procedure described above, however, lies in step (ii), ie. in the “combination” or aggregation of the frequency and severity distributions obtained from the data. As mentioned above, both distributions consist of a completely different nature, since the first is a *discrete* distribution, expressed in terms of number of events per time units (eg. number of frauds per month), while the second is a *continuous* distribution, expressed in monetary units (eg. dollars). Hence both distributions are not directly additive or multiplicative.

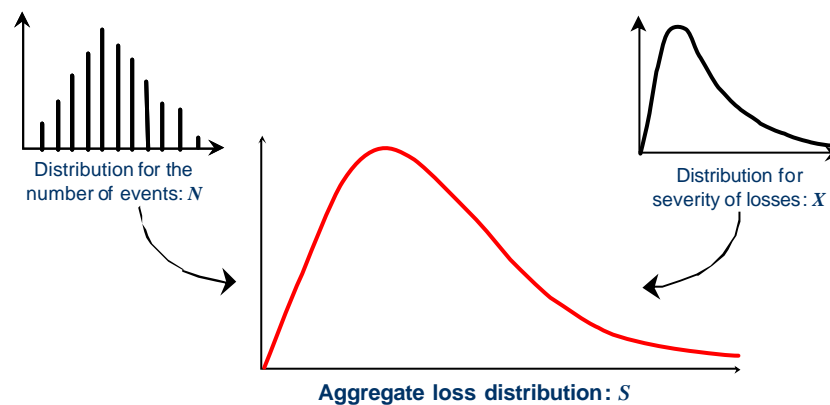


Figure 2: Aggregate Loss Distribution (S)

To combine both types of distributions there are basically two approaches: *closed form* and *open form* solutions.

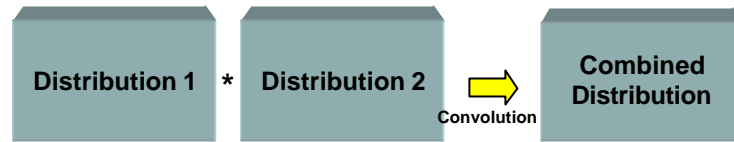
Closed form solutions involve solving analytical formulas. For the problem at hand the most straightforward closed form solution is to combine distributions by means of a (mostly theoretical) mathematical operation, called *convolution*, represented by the * (star) symbol. This operation usually involves solving complicated integrals.

An alternative method to combine both distributions (still closed-form) is not to deal with them directly, but to take some transformation that allows to manipulate the distributions more efficiently. Such a transformation is the *Fourier transform*, which operates on the *frequency domain*. This approach involves dealing with trigonometric functions (such as sines and cosines), and complex numbers. Since Fourier transforms *are* multiplicative, once we obtain the transforms of the distributions we obtain their product (an easier operation than convolution !) To obtain the aggregate loss distribution we take the *inverse Fourier transform* of this product.

⁵ See Navarrete (2002) for a thorough discussion of types of loss events, business lines and database structure.

Both procedures are shown in Figure 3.⁶

1) Convolution:



2) Fourier Transform:

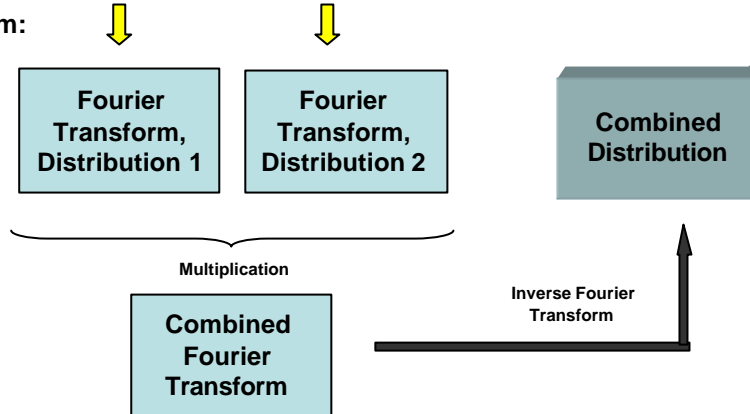


Figure 3: Comparison of Convolution and Fourier Transforms

As it may be expected, this way of obtaining the aggregate loss distribution may be daunting (and usually frustrating) for medium and small financial institutions that lack statisticians, mathematicians, or similar professionals in their risk departments. Even though better algorithms (such as the *Fast Fourier Transform*) have been applied to deal with these mostly theoretical problems, the calculations involved would place the critical calculation of prudential capital requirements beyond the reach of most financial institutions.⁷

Simulation Methods

In contrast to closed-form solutions that involve solving theoretical formulas and equations, an alternative way to obtain the aggregate loss distribution is by means of *open form solutions*, in which an algorithm is cleverly implemented in a computer and *it* does the job.

Monte Carlo simulation is one of these methods. Using simulation we can produce different scenarios for frequency and severity of losses by generating random numbers using each type of distribution (identified using actual loss data). The aggregation issue is straightforward since for the different scenarios each potential loss is generated according to a simulation that uses the frequency distribution identified from the data.

⁶ This figure illustrates the case for 2 distributions, but it can naturally be extended to the n -case. This type of procedures that involve taking some transform of the distributions and then taking an inverse are usually referred to as *inversion methods*.

⁷ Another method to handle joint loss distributions consists of the use of *copulas*. Despite its recent popularity, this method is also closed form and therefore its usefulness has been very limited; practical implementation algorithms are being currently developed.

Within the realm of simulation there exist some efficient methods, such as *Latin Hypercube*, in which random numbers are generated according to the frequency implied by the shape of the probability distribution ie. more random numbers are generated at the regions of the distribution where there exists more probability, hence random numbers are not “wasted”.⁸

The best way to understand to understand simulation is by means of an actual example.

Suppose we have the following data from external frauds committed in the credit card business line at a small banking institution ($N = 100$ data points).

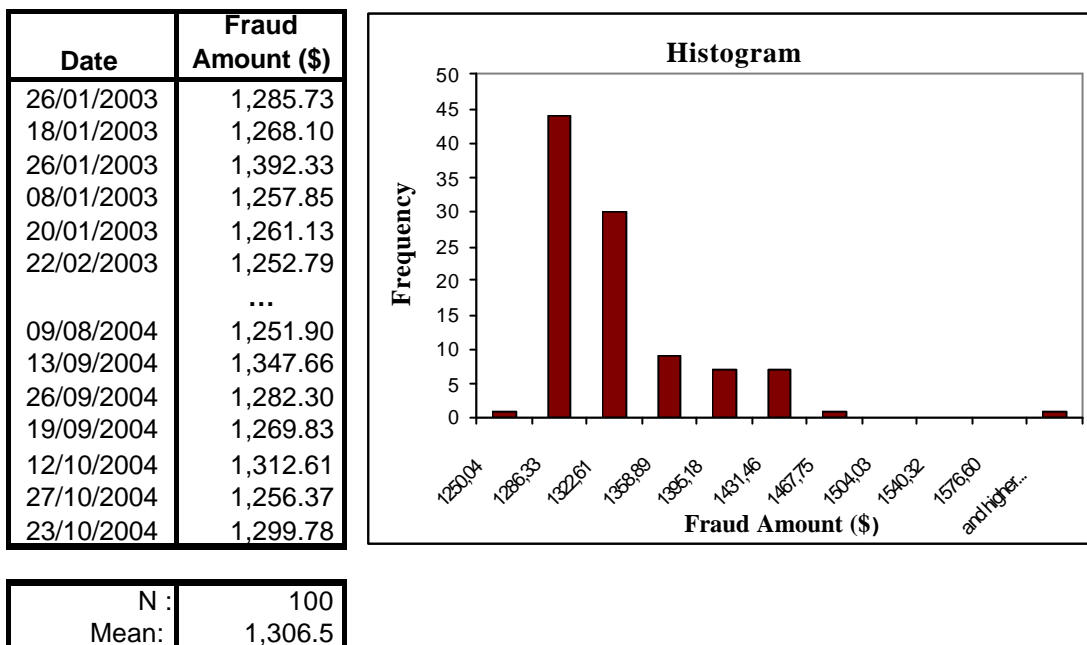


Figure 4: Credit card fraud database

These small data sets are usually found in practice for some types of events (which are either rare or not well documented, but nonetheless important !) Note also that loss data may be *highly skewed* or *truncated* due to credit line limits and other constraints.⁹

Since the data in Figure 4 are in monetary units (dollars) and represent the *severity* of losses, we will fit a *continuous* distribution to them. It turns out that the continuous distribution that best fits the data is the *Pareto distribution*, with parameters $\alpha = 23$, $a = 1250$. Hence this is the distribution that will be used to simulate the severity of potential losses.¹⁰

⁸ For further technical explanations on simulation and other open form solutions (such as Panjer’s algorithm), see Navarrete (2006b).

⁹ Loss figures for different business lines and types of events will differ greatly; see Navarrete 2002.

¹⁰ Each distribution has its own set of parameters; we used the “Best Fit” feature of the @Risk simulation package to find the distribution that best fits the data. Other choices of distributions for this data set (ranked by the Chi-square statistic) were *lognormal*, *exponential*, and *log-logistic* distributions.

In a similar fashion, a *discrete* distribution may be fitted to the *frequency* data shown in Figure 5 below.

Frequency Table	
Events per Month	# Months
k	$n(k)$
0	0
1	0
2	4
3	4
4	4
5	3
6	3
7	1
8	3
9 +	0
Total:	100 22

Figure 5: Frequency database

# Events	100
# Months	22
1	4,55

Figure 6: Estimation of the Poisson Parameter (?)

Figure 5 shows how the 100 events occurred in the time span of 22 months (ie. there were 4 months with 2 fraud events, 4 months with 3 fraud events, etc.). This information is needed to fit a discrete distribution to the data. Since in this example we just want to fit a *Poisson distribution* to the data (as is usually done in practice), we only need to know that the 100 events occurred in 22 months, yielding an average rate of 4,55 events per month (Figure 6). Thus $\lambda = 4,55$ is the parameter of the Poisson distribution that will be used to simulate the frequency or *arrivals* of potential losses.

Simulation Results

Once both types of distributions have been identified, we use Monte Carlo simulation to generate different scenarios for frequency and severity of losses. The obtained aggregate loss distribution is shown below.¹¹

¹¹ A total of 10,000 loss scenarios were generated using the simulation package @Risk using the Latin hypercube method described above.

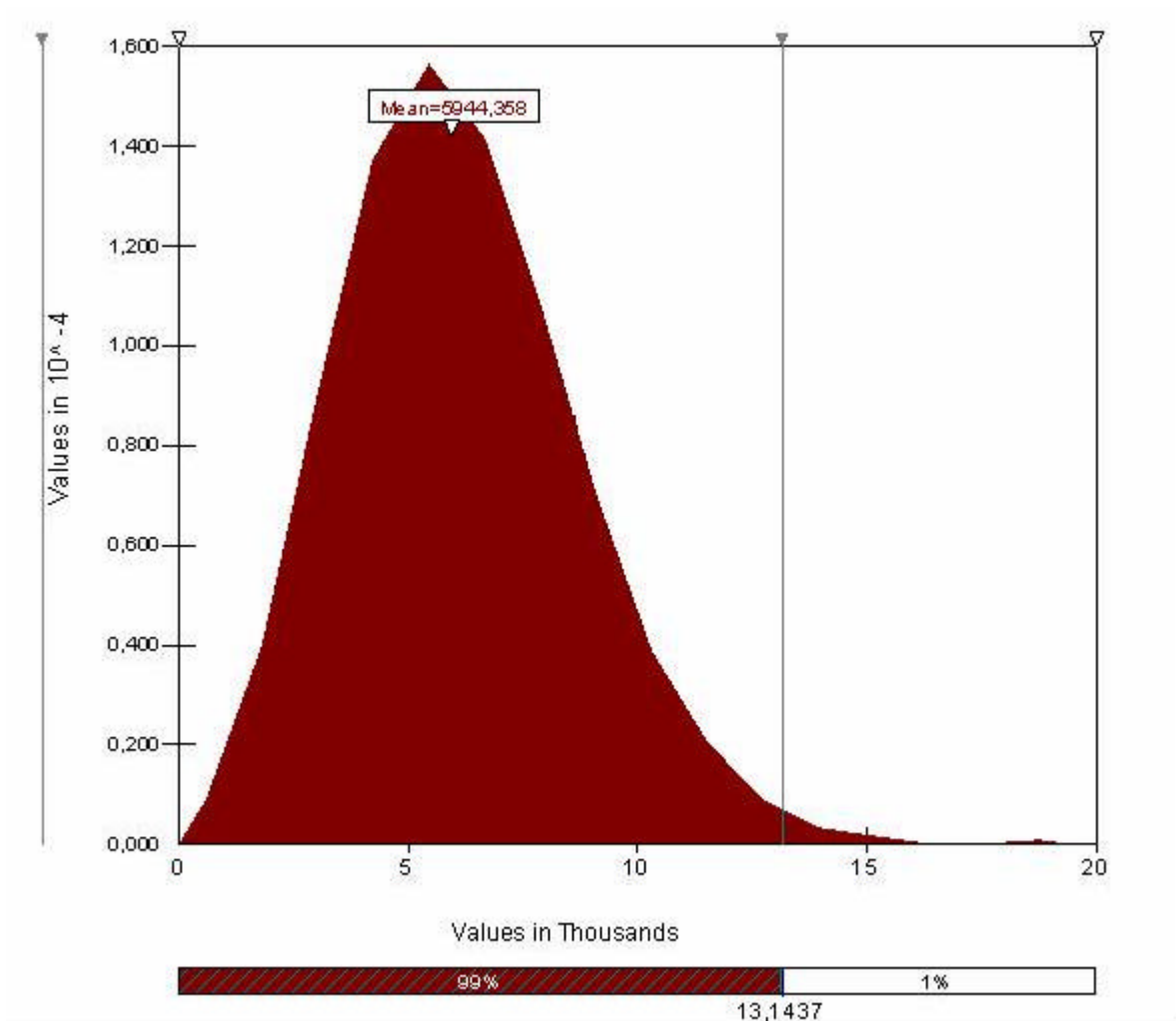


Figure 7: Aggregate loss distribution obtained by simulation

We first note from Figure 7 that the mean of the aggregate loss distribution, \$ 5,944.36, is equal to the product of the means of the Pareto and Poisson distributions (ie. \$ 1,306.5 x 4,55 events/month), as the theory dictates.¹²

Keeping track of units, this means an expected loss of \$ 5,944 per month. This is the **provisioning level** that this bank should establish to cover “mean” or “average” losses incurred in its usual course of business due to external fraud in the credit card business line.

To establish adequate **capital reserves** to cover **unexpected losses**, we simply consider percentiles of the aggregate loss distribution at different confidence levels (ie. VAR), and subtract the expected loss of \$ 5,944.36 (note that the expected loss remains the *same* for

¹² In terms of Figure 2, if N is the Poisson (frequency) distribution, X is the Pareto (severity) distribution, and S is the aggregate loss distribution, we have that $E(S) = E(N)E(X)$.



each confidence level). The figure below shows the results obtained for several confidence levels.¹³

Business Line: Credit Card
Event Type: External Fraud

Percentiles		Losses		Provisions and Capital	
Confidence Level	VAR (Percentile of Loss Distribution)	Expected Loss (EL)	Unexpected Loss (UL)	Provisions	Capital Requirement
95.0%	10,546.90	5,944.36	4,602.54	5,944.36	4,602.54
99.0%	13,143.73	5,944.36	7,199.37	5,944.36	7,199.37
99.9%	16,826.33	5,944.36	10,881.97	5,944.36	10,881.97

Figure 8: Provisions and capital requirements at different confidence levels

Therefore, if this bank wants to cover average losses incurred in its usual course of business due to external fraud in its credit card unit, it should keep \$ 5,944 as (one-month) provisions. However, if the institution also wants to protect the stability of its credit card business by establishing capital to serve as buffer against potential severe losses, it should keep an additional \$ 10,882 in capital reserves (to be adequately covered at the 99,9 % level).

In other words, if these reserves are not established, it could happen that in a single month in which there occur losses which combined yield a total amount of \$ 16,826, this amount would have to be sent directly to P&L, affecting shareholders and equity.¹⁴

On Confidence Levels and Maximum Loss

It turns out that simulation is also an excellent tool to get a better understanding of confidence levels. How are the losses at the 95 %, 99 % and 99,9 % confidence levels obtained via simulation ?

The answer is straightforward: if, for example, 10,000 loss scenarios are generated, the loss amount corresponding to the 95 % confidence level is the 500th largest, when ordered from largest to lowest.¹⁵

Figure 9 provides the order (rank) of losses at different confidence levels.

¹³ Note that Figure 7 shows VAR at the 99 % level; capital requirement at the 99,9 % confidence level is the one suggested by Basel II.

¹⁴ Note that the confidence level indicates that there is only a 0,1% chance of this happening. This is why many authors consider the Basel standard too strict and prefer confidence levels in the 95 % - 99 % range.

¹⁵ ie. 10,000 - 10,000 (95 %).

Confidence Level	Rank
99.9%	10
99%	100
95%	500

Figure 9: Loss ranks at different confidence levels (for 10,000 loss scenarios)

Note from Figure 9 that there *is* substantial difference between losses at the 99 % and 99,9 % confidence levels. Whereas the loss at the 99 % confidence level is the 100th largest, the one at the 99,9 % level is the 10th largest. That is, there are 90 loss amounts in-between generated by the simulation. This could mean a difference of several thousand or millions of dollars in some cases.

This brings us naturally to the question of the *maximum* loss. Whereas the estimation of this critical figure is a difficult theoretical issue (solved by methods such as those proposed by *Extreme Value Theory*), simulation readily offers an estimate, as the largest figure produced by the 10,000 loss scenarios. In our example, the estimate is \$ 19,364 (see Figure 7).¹⁶ Even though this amount should not be taken as the *exact* figure for the maximum, it does offer a very valuable estimate.¹⁷

Conclusion

As we have seen, loss simulation results are transparent and easy to interpret. Percentiles at different confidence levels are also very easy to visualize. Furthermore, simulation methods also provide estimates for the maximum loss, which are very difficult to obtain by traditional methods.

In summary, simulation is a very convenient tool to generate the aggregate loss distribution and hence to obtain all the relevant statistics that are of critical importance for risk managers and regulators alike.

¹⁶ The @Risk simulation program also yields information on the *minimum* loss, and on other useful indicators such as variance, skewness and kurtosis, that could serve, for instance, to test the degree of non-normality of the aggregate loss distribution.

¹⁷ To obtain more estimates on the maximum loss, one would run more simulations (say, 5,000) to obtain a complete *distribution* for this parameter. The distributions obtained this way for the maximum are skewed and leptokurtic, as the theory dictates (see Navarrete 2006a).



References

Basel Committee on Banking Supervision, BCBS (2004), *“International Convergence of Capital Measurement and Capital Standards. A Revised Framework”*, June 2004.

Navarrete, E. (2002), *“Getting started in Operational Risk Measurement: How to define types of loss events, business lines and adequate database structures”*, Internal Document, Scalar Consulting.

Navarrete, E. (2005a), *“Practical estimation of capital requirements for credit portfolios by simulation methods”*, Internal Document, Scalar Consulting.

Navarrete, E. (2006a), *“Comparison of Extreme Value Theory (EVT) and simulation analysis for predicted maxima”*, Internal Document, Scalar Consulting.

Navarrete, E. (2006b), *“Technical issues in simulation methods: convergence, stability, and sensitivity analysis”*, Internal Document, Scalar Consulting.