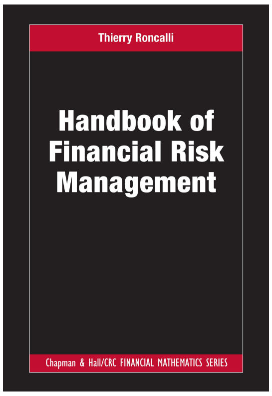

Handbook of Financial Risk Management

Thierry Roncalli

Operational Risk

Publication details
https://test.routledgehandbooks.com/doi/10.1201/9781315144597-5
Thierry Roncalli
Published online on: 20 Apr 2020

Chapter 5

Operational Risk

The integration of operational risk into the Basel II Accord was a long process because of the hostile reaction from the banking sector. At the end of the 1990s, the risk of operational losses was perceived as relatively minor. However, some events had shown that it was not the case. The most famous example was the bankruptcy of the Barings Bank in 1995. The loss of $1.3 bn was due to a huge position of the trader Nick Leeson in futures contracts without authorization. Other examples included the money laundering in Banco Ambrosiano Vatican Bank (1983), the rogue trading in Sumitomo Bank (1996), the headquarter fire of Crédit Lyonnais (1996), etc. Since the publication of the CP2 in January 2001, the position of banks has significantly changed and operational risk is today perceived as a major risk for the banking industry. Management of operational risk has been strengthened, with the creation of dedicated risk management units, the appointment of compliance officers and the launch of anti-money laundering programs.

5.1 Definition of operational risk

The Basel Committee defines the operational risk in the following way:

> "*Operational risk is defined as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. This definition includes legal risk, but excludes strategic and reputational risk*" (BCBS, 2006, page 144).

The operational risk covers then all the losses of the bank that cannot be attributed to market and credit risk. Nevertheless, losses that result from strategic decisions are not taken into account. An example is the purchase of a software or an information system, which is not relevant for the firm. Losses due to reputational risk are also excluded from the definition of operational risk. They are generally caused by an event, which is related to another risk. The difficulty is to measure the indirect loss of such events in terms of business. For instance, if we consider the diesel emissions scandal of Volkswagen, we can estimate the losses due to the recall of cars, class action lawsuits and potential fines. However, it is impossible to know what the impact of this event will be on the future sales and the market share of Volkswagen.

In order to better understand the concept of operational risk, we give here the loss even type classification adopted by the Basel Committee:

1. Internal fraud ("*losses due to acts of a type intended to defraud, misappropriate property or circumvent regulations, the law or company policy, excluding diversity/discrimination events, which involves at least one internal party*")
    (a) Unauthorized activity
    (b) Theft and fraud

2. External fraud ("*losses due to acts of a type intended to defraud, misappropriate property or circumvent the law, by a third party*")
   (a) Theft and fraud
   (b) Systems security
3. Employment practices and workplace safety ("*losses arising from acts inconsistent with employment, health or safety laws or agreements, from payment of personal injury claims, or from diversity/discrimination events*")
   (a) Employee relations
   (b) Safe environment
   (c) Diversity & discrimination
4. Clients, products & business practices ("*losses arising from an unintentional or negligent failure to meet a professional obligation to specific clients (including fiduciary and suitability requirements), or from the nature or design of a product*")
   (a) Suitability, disclosure & fiduciary
   (b) Improper business or market practices
   (c) Product flaws
   (d) Selection, sponsorship & exposure
   (e) Advisory activities
5. Damage to physical assets ("*losses arising from loss or damage to physical assets from natural disaster or other events*")
   (a) Disasters and other events
6. Business disruption and system failures ("*losses arising from disruption of business or system failures*")
   (a) Systems
7. Execution, delivery & process management ("*losses from failed transaction processing or process management, from relations with trade counterparties and vendors*")
   (a) Transaction capture, execution & maintenance
   (b) Monitoring and reporting
   (c) Customer intake and documentation
   (d) Customer/client account management
   (e) Trade counterparties
   (f) Vendors & suppliers

This is a long list of loss types, because the banking industry has been a fertile ground for operational risks. We have already cited some well-know operational losses before the crisis.

In 2009, the Basel Committee has published the results of a loss data collection exercise. For this LDCE, 119 banks submitted a total of 10.6 million internal losses with an overall loss amount of €59.6 bn. The largest 20 losses represented a total of €17.6 bn. In Table 5.1, we have reported statistics of the loss data, when the loss is larger than €20 000. For each year, we indicate the number n_L of losses, the total loss amount L and the number n_B of reporting banks. Each bank experienced more than 300 losses larger than €20 000 per year on average. We also notice that these losses represented about 90% of the overall loss amount.

TABLE 5.1: Internal losses larger than €20 000 per year

Year	pre 2002	2002	2003	2004	2005	2006	2007
n_L	14 017	10 216	13 691	22 152	33 216	36 386	36 622
L (in € bn)	3.8	12.1	4.6	7.2	9.7	7.4	7.9
n_B	24	35	55	68	108	115	117

Source: BCBS (2009d).

Since 2008, operational risk has dramatically increased. For instance, rogue trading has impacted many banks and the magnitude of these unauthorized trading losses is much higher than before[1]. The Libor interest rate manipulation scandal led to very large fines ($2.5 bn for Deutsche Bank, $1 bn for Rabobank, $545 mn for UBS, etc.). In May 2015, six banks (Bank of America, Barclays, Citigroup, J.P. Morgan, UBS and RBS) agreed to pay fines totaling $5.6 bn in the case of the forex scandal[2]. The anti-money laundering controls led BNP Paribas to pay a fine of $8.9 bn in June 2014 to the US federal government. In this context, operational risk, and more specifically compliance and legal risk, is a major concern for banks. It is an expansive risk, because of the direct losses, but also because of the indirect costs induced by the proliferation of internal controls and security infrastructure[3].

Remark 60 *Operational risk is not limited to the banking sector. Other financial sectors have been impacted by such risk. The most famous example is the Ponzi scheme organized by Bernard Madoff, which caused a loss of* $50 *bn to his investors.*

5.2 Basel approaches for calculating the regulatory capital

In this section, we present the three approaches described in the Basel II framework in order to calculate the capital charge for operational risk:

1. the basic indicator approach (BIA);
2. the standardized approach (TSA);
3. and advanced measurement approaches (AMA).

[1]We can cite Société Générale in 2008 ($7.2 bn), Morgan Stanley in 2008 ($9.0 bn), BPCE in 2008 ($1.1 bn), UBS in 2011 ($2 bn) and JPMorgan Chase in 2012 ($5.8 bn).

[2]The Libor scandal was a series of fraudulent actions connected to the Libor (London Interbank Offered Rate), while the forex scandal concerns several banks, which have manipulated exchange rates via electronic chatrooms in which traders discussed the trades they planned to do.

[3]A typical example of expansive cost is the risk of cyber attacks.

We also present the Basel III framework of the standardized approach for measuring operational risk capital with effect from January 2022.

5.2.1 The basic indicator approach

The basic indicator approach is the simplest method for calculating the operational risk capital requirement. In this case, the capital charge is a fixed percentage of annual gross income:

$$\mathcal{K} = \alpha \cdot \overline{\mathrm{GI}}$$

where α is set equal to 15% and $\overline{\mathrm{GI}}$ is the average of the positive gross income over the previous three years:

$$\overline{\mathrm{GI}} = \frac{\max\left(\mathrm{GI}_{t-1}, 0\right) + \max\left(\mathrm{GI}_{t-2}, 0\right) + \max\left(\mathrm{GI}_{t-3}, 0\right)}{\sum_{k=1}^{3} \mathbb{1}\left\{\mathrm{GI}_{t-k} > 0\right\}}$$

In this approach, the capital charge is related to the financial results of the bank, but not to its risk exposure.

5.2.2 The standardized approach

The standardized approach is an extended version of the previous method. In this case, the bank is divided into eight business lines, which are given in Table 5.2. The bank then calculates the capital charge for each business line:

$$\mathcal{K}_{j,t} = \beta_j \cdot \mathrm{GI}_{j,t}$$

where β_j and $\mathrm{GI}_{j,t}$ are a fixed percentage[4] and the gross income corresponding to the j^{th} business line. The total capital charge is the three-year average of the sum of all the capital charges:

$$\mathcal{K} = \frac{1}{3} \sum_{k=1}^{3} \max\left(\sum_{j=1}^{8} \mathcal{K}_{j,t-k}, 0\right)$$

We notice that a negative capital charge in one business line may offset positive capital charges in other business lines. If the values of gross income are all positive, the total capital charge becomes:

$$\begin{aligned}
\mathcal{K} &= \frac{1}{3} \sum_{k=1}^{3} \sum_{j=1}^{8} \beta_j \cdot \mathrm{GI}_{j,t-k} \\
&= \sum_{j=1}^{8} \beta_j \cdot \overline{\mathrm{GI}}_j
\end{aligned}$$

where $\overline{\mathrm{GI}}_j$ is the average gross income over the previous three years of the j^{th} business line.

Example 51 *We consider Bank A, whose activity is mainly driven by retail banking and asset management. We compare it with Bank B, which is more focused on corporate finance. We assume that the two banks are only composed of four business lines: corporate finance,*

[4]The values taken by the beta coefficient are reported in Table 5.2.

TABLE 5.2: Mapping of business lines for operational risk

Level 1	Level 2	β_j
Corporate Finance†	Corporate Finance Municipal/Government Finance Merchant Banking Advisory Services	18%
Trading & Sales‡	Sales Market Making Proprietary Positions Treasury	18%
Retail Banking	Retail Banking Private Banking Card Services	12%
Commercial Banking♮	Commercial Banking	12%
Payment & Settlement	External Clients	18%
Agency Services	Custody Corporate Agency Corporate Trust	15%
Asset Management	Discretionary Fund Management Non-Discretionary Fund Management	12%
Retail Brokerage	Retail Brokerage	12%

†Mergers and acquisitions, underwriting, securitization, syndications, IPO, debt placements.
‡Buying and selling of securities and derivatives, own position securities, lending and repos, brokerage.
♮Project finance, real estate, export finance, trade finance, factoring, leasing, lending, guarantees, bills of exchange.

retail banking, agency services and asset management. The gross income expressed in \$ mn for the last three years is given below:

Business line	*Bank A*			*Bank B*		
	$t-1$	$t-2$	$t-3$	$t-1$	$t-2$	$t-3$
Corporate finance	10	15	−30	200	300	150
Retail banking	250	230	205	50	45	−30
Agency services	10	10	12			
Asset management	70	65	72	12	8	−4

For Bank A, we obtain $\mathrm{GI}_{t-1} = 340$, $\mathrm{GI}_{t-2} = 320$ and $\mathrm{GI}_{t-3} = 259$. The average gross income is then equal to 306.33, implying that the BIA capital charge $\mathcal{K}_A^{\mathrm{BIA}}$ is equal to \$45.95 mn. If we consider Bank B, the required capital $\mathcal{K}_B^{\mathrm{BIA}}$ is lower and equal to \$36.55 mn. In the case of the standardized approach, the beta coefficients are respectively equal to 18%, 12%, 15% and 12%. We deduce that:

$$\begin{aligned}\mathcal{K}_A^{\mathrm{TSA}} &= \frac{1}{3} \times (\max(18\% \times 10 + 12\% \times 250 + 15\% \times 10 + 12\% \times 70, 0) + \\ &\quad \max(18\% \times 15 + 12\% \times 230 + 15\% \times 10 + 12\% \times 65, 0) + \\ &\quad \max(-18\% \times 30 + 12\% \times 205 + 15\% \times 12 + 12\% \times 72, 0)) \\ &= \$36.98 \text{ mn}\end{aligned}$$

We also have $\mathcal{K}_B^{\text{TSA}} = \$42.24$ mn. We notice that $\mathcal{K}_A^{\text{BIA}} > \mathcal{K}_A^{\text{TSA}}$ and $\mathcal{K}_B^{\text{BIA}} < \mathcal{K}_B^{\text{TSA}}$. Bank A has a lower capital charge when using TSA instead of BIA, because it is more exposed to low-risk business lines (retail banking and asset management). For Bank B, it is the contrary because its main exposure concerns high-risk business lines (corporate finance). However, if we assume that the gross income of the corporate finance for Bank B at time $t-3$ is equal to -150 instead of $+150$, we obtain $\mathcal{K}_B^{\text{BIA}} = \$46.13$ mn and $\mathcal{K}_B^{\text{TSA}} = \$34.60$ mn. In this case, the TSA approach is favorable, because the gross income at time $t-3$ is negative implying that the capital contribution at time $t-3$ is equal to zero.

Contrary to the basic indicator approach that requires no criteria to be used, banks must satisfy a list of qualifying criteria for the standardized approach. For instance, the board of directors must be actively involved in the oversight of the operational risk management framework and each business line must have sufficient resources to manage operational risk. International active banks must also collect operational losses and use this information for taking appropriate action.

5.2.3 Advanced measurement approaches

Like the internal model-based approach for market risk, the AMA method is defined by certain criteria without refereing to a specific statistical model:

- The capital charge should cover the one-year operational loss at the 99.9% confidence level. It corresponds to the sum of expected loss (EL) and unexpected loss (UL).
- The model must be estimated using a minimum five-year observation period of internal loss data, and capture tail loss events by considering for example external loss data when it is needed. It must also include scenario analysis and factors reflecting internal control systems.
- The risk measurement system must be sufficiently granular to capture the main operational risk factors. By default, the operational risk of the bank must be divided into the 8 business lines and the 7 event types. For each cell of the matrix, the model must estimate the loss distribution and may use correlations to perform the aggregation.
- The allocation of economic capital across business lines must create incentives to improve operational risk management.
- The model can incorporate the risk mitigation impact of insurance, which is limited to 20% of the total operational risk capital charge.

The validation of the AMA model does not only concern the measurement aspects, but also the soundness of the entire operational risk management system. This concerns governance of operational risk, dedicated resources, management structure, risk maps and key risk indicators (KRI), notification and action procedures, emergency and crisis management, business continuity and disaster recovery plans.

In order to better understand the challenges of an internal model, we have reported in Table 5.3 the distribution of annualized loss amounts by business line and event type obtained with the 2008 loss data collection exercise. We first notice an heterogeneity between business lines. For instance, losses were mainly concentrated in the fourth event type (clients, products & business practices) for the corporate finance business line (93.7%) and the seventh event type (execution, delivery & process management) for the payment & settlement business line (76.4%). On average, these two event types represented more than 75% of the total loss amount. In contrast, fifth and sixth event types (damage to physical assets, business disruption and system failures) had a small contribution close to 1%. We also notice that operational losses mainly affected retail banking, followed by corporate finance and trading & sales. One of the issues is that this picture of operational risk is no longer valid after 2008 with the increase of losses in trading & sales, but also in payment & settlement. The nature of operational risk changes over time, which is a big challenge to build an internal model to calculate the required capital.

TABLE 5.3: Distribution of annualized operational losses (in %)

Business line	Event type							All
	1	2	3	4	5	6	7	
Corporate Finance	0.2	0.1	0.6	93.7	0.0	0.0	5.4	28.0
Trading & Sales	11.0	0.3	2.3	29.0	0.2	1.8	55.3	13.6
Retail Banking	6.3	19.4	9.8	40.4	1.1	1.5	21.4	32.0
Commercial Banking	11.4	15.2	3.1	35.5	0.4	1.7	32.6	7.6
Payment & Settlement	2.8	7.1	0.9	7.3	3.2	2.3	76.4	2.6
Agency Services	1.0	3.2	0.7	36.0	18.2	6.0	35.0	2.6
Asset Management	11.1	1.0	2.5	30.8	0.3	1.5	52.8	2.5
Retail Brokerage	18.1	1.4	6.3	59.5	0.1	0.2	14.4	5.1
Unallocated	6.5	2.8	28.4	28.3	6.5	1.3	26.2	6.0
All	6.1	8.0	6.0	52.4	1.4	1.2	24.9	100.0

Source: BCBS (2009d).

5.2.4 Basel III approach

From January 2022, the standardized measurement approach (SMA) will replace the three approaches of the Basel II framework. The SMA is based on three components: the business indicator (BI), the business indicator component (BIC) and the internal loss multiplier (ILM). The business indicator is a proxy of the operational risk:

$$\mathrm{BI} = \mathrm{ILDC} + \mathrm{SC} + \mathrm{FC}$$

where ILDC is the interest, leases and dividends component, SC is the services component and FC is the financial component. The underlying idea is to list the main activities that generate operational risk:

$$\begin{cases} \mathrm{ILDC} = \min\left(\left|\mathrm{INC} - \mathrm{EXP}\right|, 2.25\% \cdot \mathrm{IRE}\right) + \mathrm{DIV} \\ \mathrm{SC} = \max\left(\mathrm{OI}, \mathrm{OE}\right) + \max\left(\mathrm{FI}, \mathrm{FE}\right) \\ \mathrm{FC} = \left|\Pi_{\mathrm{TB}}\right| + \left|\Pi_{\mathrm{BB}}\right| \end{cases}$$

where INC represents the interest income, EXP the interest expense, IRE the interest earning assets, DIV the dividend income, OI the other operating income, OE the other operating expense, FI the fee income, FE the fee expense, Π_{TB} the net P&L of the trading book and Π_{BB} the net P&L of the banking book. All these variables are calculated as the average over the last three years. We can draw a parallel between the business indicator components and the TSA components. For example, ILDC concerns corporate finance, retail banking, commercial banking, SC is related to payment & settlement, agency services, asset management, retail brokerage, while FC mainly corresponds to trading & sales. Once the BI is calculated and expressed in \$ bn, we determine the business indicator component, which is given by:

$$\mathrm{BIC} = 12\% \cdot \min\left(\mathrm{BI}, \$1 \text{ bn}\right) + 15\% \cdot \min\left(\mathrm{BI} - 1, \$30 \text{ bn}\right) + 18\% \cdot \min\left(\mathrm{BI} - 30\right)^{+}$$

The BIC formula recalls the basic indicator approach of Basel II, but it introduces a marginal weight by BI tranches. Finally, the bank has to compute the internal loss multiplier, which is defined as:

$$\text{ILM} = \ln\left(e^1 - 1 + \left(\frac{15 \cdot \bar{L}}{\text{BIC}}\right)^{0.8}\right)$$

where $\bar{L}$ is the average annual operational risk losses over the last 10 years. ILM can be lower or greater than one, depending on the value of $\bar{L}$:

$$\begin{cases} \text{ILM} < 1 \Leftrightarrow \bar{L} < \text{BIC}/15 \\ \text{ILM} = 1 \Leftrightarrow \bar{L} = \text{BIC}/15 \\ \text{ILM} > 1 \Leftrightarrow \bar{L} > \text{BIC}/15 \end{cases}$$

The capital charge for the operational risk is then equal to[5]:

$$\mathcal{K} = \text{ILM} \cdot \text{BIC}$$

Remark 61 *The SMA of the Basel III framework may be viewed as a mix of the three approaches of the Basel II framework: BIA, TSA and AMA. Indeed, SMA is clearly a modified version of BIA by considering a basic indicator based on sources of operational risk. In this case, the business indicator can be related to TSA. Finally, the introduction of the* ILM *coefficient is a way to consider a more sensitive approach based on internal losses, which is the basic component of AMA.*

5.3 Loss distribution approach

Although the Basel Committee does not advocate any particular method for the AMA method in the Basel II framework, the loss distribution approach (LDA) is the recognized standard model for calculating the capital charge. This model is not specific to operational risk because it was developed in the case of the collective risk theory at the beginning of 1900s. However, operational risk presents some characteristics that need to be considered.

5.3.1 Definition

The loss distribution approach is described in Klugman *et al.* (2012) and Frachot *et al.* (2001). We assume that the operational loss L of the bank is divided into a matrix of homogenous losses:

$$L = \sum_{k=1}^{K} S_k \tag{5.1}$$

where S_k is the sum of losses of the k^{th} cell and K is the number of cells in the matrix. For instance, if we consider the Basel II classification, the mapping matrix contains 56 cells corresponding to the 8 business lines and 7 event types. The loss distribution approach is a

[5]However, the computation of the ILM coefficient is subject to some standard requirements. For instance, ILM is set to one for banks with a BIC lower than \$1 bn and supervisors can impose the value of the ILM coefficient for banks that do not meet loss data collection criteria.

method to model the random loss S_k of a particular cell. It assumes that S_k is the random sum of homogeneous individual losses:

$$S_k = \sum_{n=1}^{N_k(t)} X_n^{(k)} \tag{5.2}$$

where $N_k(t)$ is the random number of individual losses for the period $[0, t]$ and $X_n^{(k)}$ is the n^{th} individual loss. For example, if we consider internal fraud in corporate finance, we can write the loss for the next year as follows:

$$S = X_1 + X_2 + \ldots + X_{N(1)}$$

where X_1 is the first observed loss, X_2 is the second observed loss, $X_{N(1)}$ is the last observed loss of the year and $N(1)$ is the number of losses for the next year. We notice that we face two sources of uncertainty:

1. we don't know what will be the magnitude of each loss event(severity risk);
2. we don't know how many losses will occur in the next year (frequency risk).

In order to simplify the notations, we omit the index k and rewrite the random sum as follows:

$$S = \sum_{n=1}^{N(t)} X_n \tag{5.3}$$

The loss distribution approach is based on the following assumptions:

- the number $N(t)$ of losses follows the loss frequency distribution $\mathbf{P}$; the probability that the number of loss events is equal to n is denoted by $p(n)$;
- the sequence of individual losses X_n is independent and identically distributed (*iid*); the corresponding probability distribution $\mathbf{F}$ is called the loss severity distribution;
- the number of events is independent from the amount of loss events.

Once the probability distributions $\mathbf{P}$ and $\mathbf{F}$ are chosen, we can determine the probability distribution of the aggregate loss S, which is denoted by $\mathbf{G}$ and is called the compound distribution.

Example 52 *We assume that the number of losses is distributed as follows:*

n	0	1	2	3
$p(n)$	50%	30%	17%	3%

The loss amount can take the values \$100 *and* \$200 *with probabilities* 70% *and* 30%.

To calculate the probability distribution $\mathbf{G}$ of the compound loss, we first define the probability distribution of X_1, $X_1 + X_2$ and $X_1 + X_2 + X_3$, because the maximum number of losses is equal to 3. If there is only one loss, we have $\Pr\{X_1 = 100\} = 70\%$ and $\Pr\{X_1 = 200\} = 30\%$. In the case of two losses, we obtain $\Pr\{X_1 + X_2 = 200\} = 49\%$, $\Pr\{X_1 + X_2 = 300\} = 42\%$ and $\Pr\{X_1 + X_2 = 400\} = 9\%$. Finally, the sum of three losses takes the values 300, 400, 500 and 600 with probabilities 34.3%, 44.1%, 18.9% and 2.7%

respectively. We notice that these probabilities are in fact conditional to the number of losses. Using Bayes theorem, we obtain:

$$\Pr\{S=s\} = \sum_n \Pr\left\{\sum\nolimits_{i=1}^{n} X_i = s \,\middle|\, N(t)=n\right\} \cdot \Pr\{N(t)=n\}$$

We deduce that:

$$\begin{aligned}\Pr\{S=0\} &= \Pr\{N(t)=0\} \\ &= 50\%\end{aligned}$$

and:

$$\begin{aligned}\Pr\{S=100\} &= \Pr\{X_1=100\} \times \Pr\{N(t)=1\} \\ &= 70\% \times 30\% \\ &= 21\%\end{aligned}$$

The compound loss can take the value 200 in two different ways:

$$\begin{aligned}\Pr\{S=200\} &= \Pr\{X_1=200\} \times \Pr\{N(t)=1\} + \\ &\quad \Pr\{X_1+X_2=200\} \times \Pr\{N(t)=2\} \\ &= 30\% \times 30\% + 49\% \times 17\% \\ &= 17.33\%\end{aligned}$$

For the other values of S, we obtain $\Pr\{S=300\} = 8.169\%$, $\Pr\{S=400\} = 2.853\%$, $\Pr\{S=500\} = 0.567\%$ and $\Pr\{S=600\} = 0.081\%$.

The previous example shows that the cumulative distribution function of S can be written as[6]:

$$\mathbf{G}(s) = \begin{cases} \sum_{n=1}^{\infty} p(n)\,\mathbf{F}^{n\star}(s) & \text{for} \quad s>0 \\ p(0) & \text{for} \quad s=0 \end{cases} \tag{5.4}$$

where $\mathbf{F}^{n\star}$ is the n-fold convolution of $\mathbf{F}$ with itself:

$$\mathbf{F}^{n\star}(s) = \Pr\left\{\sum\nolimits_{i=1}^{n} X_i \leq s\right\} \tag{5.5}$$

In Figure 5.1, we give an example of a continuous compound distribution when the annual number of losses follows the Poisson distribution $\mathcal{P}(50)$ and the individual losses follow the log-normal distribution $\mathcal{LN}(8,5)$. The capital charge, which is also called the capital-at-risk (CaR), corresponds then to the percentile α:

$$\text{CaR}(\alpha) = \mathbf{G}^{-1}(\alpha) \tag{5.6}$$

The regulatory capital is obtained by setting α to 99.9%: $\boldsymbol{\mathcal{K}} = \text{CaR}(99.9\%)$. This capital-at-risk is valid for one cell of the operational risk matrix. Another issue is to calculate the capital-at-risk for the bank as a whole. This requires defining the dependence function between the different compound losses $(S_1, S_2, \ldots, S_K)$. In summary, here are the different steps to implement the loss distribution approach:

- for each cell of the operational risk matrix, we estimate the loss frequency distribution and the loss severity distribution;
- we then calculate the capital-at-risk;
- we define the dependence function between the different cells of the operational risk matrix, and deduce the aggregate capital-at-risk.

[6]When $\mathbf{F}$ is a discrete probability function, it is easy to calculate $\mathbf{F}^{n\star}(s)$ and then deduce $\mathbf{G}(s)$. However, the determination of $\mathbf{G}(s)$ is more difficult in the general case of continuous probability functions. This issue is discussed in Section 5.3.3 on page 327.

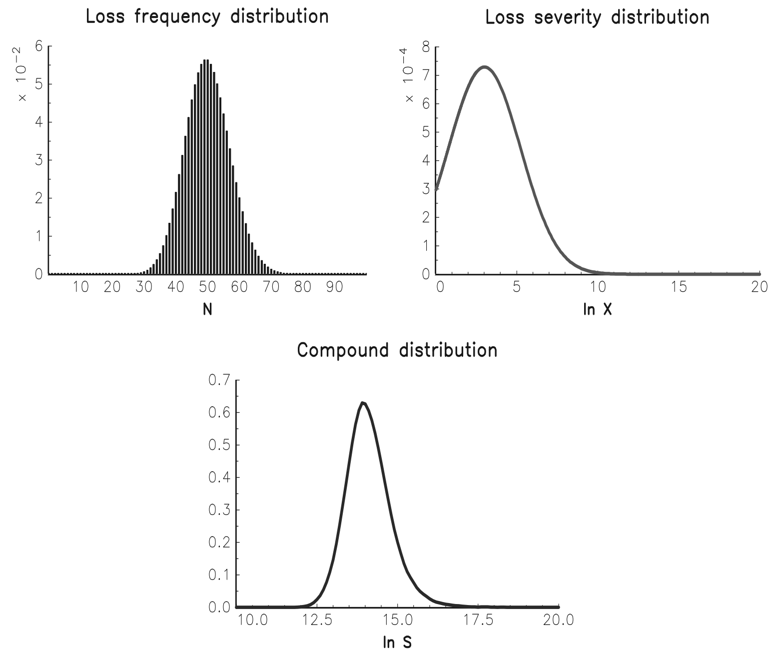

FIGURE 5.1: Compound distribution when $N \sim \mathcal{P}(50)$ and $X \sim \mathcal{LN}(8, 5)$

5.3.2 Parametric estimation

We first consider the estimation of the severity distribution, because we will see that the estimation of the frequency distribution can only be done after this first step.

5.3.2.1 Estimation of the loss severity distribution

We assume that the bank has an internal loss database. We note $\{x_1, \ldots, x_T\}$ the sample collected for a given cell of the operational risk matrix. We consider that the individual losses follow a given parametric distribution $\mathbf{F}$:

$$X \sim \mathbf{F}(x; \theta)$$

where θ is the vector of parameters to estimate.

In order to be a good candidate for modeling the loss severity, the probability distribution $\mathbf{F}$ must satisfy the following properties: the support of $\mathbf{F}$ is the interval $\mathbb{R}_+$, it is sufficiently flexible to accommodate a wide variety of empirical loss data and it can fit large losses. We list here the cumulative distribution functions that are the most used in operational risk models:

- Gamma $X \sim \mathcal{G}(\alpha, \beta)$

$$\mathbf{F}(x; \theta) = \frac{\gamma(\alpha, \beta x)}{\Gamma(\alpha)}$$

where $\alpha > 0$ and $\beta > 0$.

- Log-gamma $X \sim \mathcal{LG}(\alpha, \beta)$

$$\mathbf{F}(x; \theta) = \frac{\gamma(\alpha, \beta \ln x)}{\Gamma(\alpha)}$$

where $\alpha > 0$ and $\beta > 0$.

- Log-logistic $X \sim \mathcal{LL}(\alpha, \beta)$

$$\begin{aligned}\mathbf{F}(x;\theta) &= \frac{1}{1+(x/\alpha)^{-\beta}} \\ &= \frac{x^{\beta}}{\alpha^{\beta}+x^{\beta}}\end{aligned}$$

where $\alpha > 0$ and $\beta > 0$.

- Log-normal $X \sim \mathcal{LN}(\mu, \sigma^2)$

$$\mathbf{F}(x;\theta) = \Phi\left(\frac{\ln x - \mu}{\sigma}\right)$$

where $x > 0$ and $\sigma > 0$.

- Generalized extreme value $X \sim \mathcal{GEV}(\mu, \sigma, \xi)$

$$\mathbf{F}(x;\theta) = \exp\left\{-\left[1+\xi\left(\frac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$$

where $x > \mu - \sigma/\xi$, $\sigma > 0$ and $\xi > 0$.

- Pareto $X \sim \mathcal{P}(\alpha, x_-)$

$$\mathbf{F}(x;\theta) = 1 - \left(\frac{x}{x_-}\right)^{-\alpha}$$

where $x \geq x_-$, $\alpha > 0$ and $x_- > 0$.

The vector of parameters θ can be estimated by the method of maximum likelihood (ML) or the generalized method of moments (GMM). In Chapter 10, we show that the log-likelihood function associated to the sample is:

$$\boldsymbol{\ell}(\theta) = \sum_{i=1}^{T} \ln f(x_i;\theta) \tag{5.7}$$

where $f(x;\theta)$ is the density function. In the case of the GMM, the empirical moments are:

$$\left\{\begin{array}{l} h_{i,1}(\theta) = x_i - \mathbb{E}[X] \\ h_{i,2}(\theta) = (x_i - \mathbb{E}[X])^2 - \operatorname{var}(X) \end{array}\right. \tag{5.8}$$

In Table 5.4, we report the density function $f(x;\theta)$, the mean $\mathbb{E}[X]$ and the variance $\operatorname{var}(X)$ when X follows one of the probability distributions described previously. For instance, if we consider that $X \sim \mathcal{LN}(\mu, \sigma^2)$, it follows that the log-likelihood function is:

$$\boldsymbol{\ell}(\theta) = -\sum_{i=1}^{T} \ln x_i - \frac{T}{2}\ln\sigma^2 - \frac{T}{2}\ln 2\pi - \frac{1}{2}\sum_{i=1}^{T}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2$$

whereas the empirical moments are:

$$\left\{\begin{array}{l} h_{i,1}(\theta) = x_i - e^{\mu+\frac{1}{2}\sigma^2} \\ h_{i,2}(\theta) = \left(x_i - e^{\mu+\frac{1}{2}\sigma^2}\right)^2 - e^{2\mu+\sigma^2}\left(e^{\sigma^2}-1\right) \end{array}\right.$$

TABLE 5.4: Density function, mean and variance of parametric probability distribution

Distribution	$f(x;\theta)$	$\mathbb{E}[X]$	$\operatorname{var}(X)$
$\mathcal{G}(\alpha,\beta)$	$\dfrac{\beta^{\alpha}x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)}$	$\dfrac{\alpha}{\beta}$	$\dfrac{\alpha}{\beta^2}$
$\mathcal{LG}(\alpha,\beta)$	$\dfrac{\beta^{\alpha}(\ln x)^{\alpha-1}}{x^{\beta+1}\Gamma(\alpha)}$	$\left(\dfrac{\beta}{\beta-1}\right)^{\alpha}$ if $\beta>1$	$\left(\dfrac{\beta}{\beta-2}\right)^{\alpha}-\left(\dfrac{\beta}{\beta-1}\right)^{2\alpha}$ if $\beta>2$
$\mathcal{LL}(\alpha,\beta)$	$\dfrac{\beta(x/\alpha)^{\beta-1}}{\alpha\left(1+(x/\alpha)^{\beta}\right)^2}$	$\dfrac{\alpha\pi}{\beta\sin(\pi/\beta)}$ if $\beta>1$	$\alpha^2\left(\dfrac{2\pi}{\beta\sin(2\pi/\beta)}-\dfrac{\pi^2}{\beta^2\sin^2(\pi/\beta)}\right)$ if $\beta>2$
$\mathcal{LN}\left(\mu,\sigma^2\right)$	$\dfrac{1}{x\sigma\sqrt{2\pi}}\exp\left(-\dfrac{1}{2}\left(\dfrac{x-\mu}{\sigma}\right)^2\right)$	$\exp\left(\mu+\dfrac{1}{2}\sigma^2\right)$	$\exp\left(2\mu+\sigma^2\right)\left(\exp\left(\sigma^2\right)-1\right)$
$\mathcal{GEV}(\mu,\sigma,\xi)$	$\dfrac{1}{\sigma}\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{-(1+1/\xi)}$ $\exp\left\{-\left[1+\xi\left(\dfrac{x-\mu}{\sigma}\right)\right]^{-1/\xi}\right\}$	$\mu+\dfrac{\sigma}{\xi}(\Gamma(1-\xi)-1)$ if $\xi<1$	$\dfrac{\sigma^2}{\xi^2}\left(\Gamma(1-2\xi)-\Gamma^2(1-\xi)\right)$ if $\xi<\frac{1}{2}$
$\mathcal{P}(\alpha,x_-)$	$\dfrac{\alpha x_-^{\alpha}}{x^{\alpha+1}}$	$\dfrac{\alpha x_-}{\alpha-1}$ if $\alpha>1$	$\dfrac{\alpha x_-^2}{(\alpha-1)^2(\alpha-2)}$ if $\alpha>2$

In the case of the log-normal distribution, the vector θ is composed of two parameters μ and σ, implying that two moments are sufficient to define the GMM estimator. This is also the case of other probability distributions, except the GEV distribution that requires specification of three empirical moments[7].

Example 53 *We assume that the individual losses take the following values expressed in thousand dollars:* 10.1*,* 12.5*,* 14*,* 25*,* 317.3*,* 353*,* 1 200*,* 1 254*,* 52 000 *and* 251 000*.*

Using the method of maximum likelihood, we find that $\hat{\alpha}_{\mathrm{ML}}$ and $\hat{\beta}_{\mathrm{ML}}$ are equal to 15.70 and 1.22 for the log-gamma distribution and 293 721 and 0.51 for the log-logistic distribution. In the case of the log-normal distribution[8], we obtain $\hat{\mu}_{\mathrm{ML}}=12.89$ and $\hat{\sigma}_{\mathrm{ML}}=3.35$.

The previous analysis assumes that the sample of operational losses for estimating θ represents a comprehensive and homogenous information of the underlying probability distribution **F**. In practice, loss data are plagued by various sources of bias. The first issue lies in the data generating process which underlies the way data have been collected. In almost all cases, loss data have gone through a truncation process by which data are recorded only when their amounts are higher than some thresholds. In practice, banks' internal thresholds are set in order to balance two conflicting wishes: collecting as many data as possible while reducing costs by collecting only significant losses. These thresholds, which are defined by the global risk management policy of the bank, must satisfy some criteria:

[7] We can use the moment of order 3, which corresponds to:

$$\mathbb{E}\left[(X-\mathbb{E}[X])^3\right]=\frac{\sigma^3}{\xi^3}\left(\Gamma(1-3\xi)-3\Gamma(1-2\xi)\Gamma(1-\xi)+2\Gamma^3(1-\xi)\right)$$

[8] If we consider the generalized method of moments, the estimates are $\hat{\mu}_{\mathrm{GMM}}=16.26$ and $\hat{\sigma}_{\mathrm{GMM}}=1.40$.

> "*A bank must have an appropriate de minimis gross loss threshold for internal loss data collection, for example €10 000. The appropriate threshold may vary somewhat between banks, and within a bank across business lines and/or event types. However, particular thresholds should be broadly consistent with those used by peer banks*" (BCBS, 2006, page 153).

The second issue concerns the use of relevant external data, especially when there is reason to believe that the bank is exposed to infrequent, yet potentially severe losses. Typical examples are rogue trading or cyber attacks. If the bank has not yet experienced a large amount of loss due to these events in the past, this does not mean that it will never experience such problems in the future. Therefore, internal loss data must be supplemented by external data from public and/or pooled industry databases. Unfortunately, incorporating external data is rather dangerous and requires careful methodology to avoid the pitfalls regarding data heterogeneity, scaling problems and lack of comparability between too heterogeneous data. Unfortunately, there is no satisfactory solution to deal with these scaling issues, implying that banks use external data by taking into account only reporting biases and a fixed and known threshold[9].

The previous issues imply that operational risk loss data cannot be reduced to the sample of individual losses, but also requires specifying the threshold H_i for each individual loss x_i. The form of operational loss data is then $\{(x_i, H_i), i = 1, \ldots, T\}$, where x_i is the observed value of X knowing that X is larger than the threshold H_i. Reporting thresholds affect severity estimation in the sense that the sample severity distribution (i.e. the severity distribution of reported losses) is different from the '*true*' one (i.e. the severity distribution one would obtain if all the losses were reported). Unfortunately, the true distribution is the most relevant for calculating capital charge. As a consequence, linking the sample distribution to the true one is a necessary task. From a mathematical point of view, the true distribution is the probability distribution of X whereas the sample distribution is the probability distribution of $X \mid X \geq H_i$. We deduce that the sample distribution for a given threshold H is the conditional probability distribution defined as follows:

$$\begin{aligned}
\mathbf{F}^{\star}(x;\theta \mid H) &= \Pr\{X \leq x \mid X \geq H\} \\
&= \frac{\Pr\{X \leq x, X \geq H\}}{\Pr\{X \geq H\}} \\
&= \frac{\Pr\{X \leq x\} - \Pr\{X \leq \min(x, H)\}}{\Pr\{X \geq H\}} \\
&= \mathbb{1}\{x \geq H\} \cdot \frac{\mathbf{F}(x;\theta) - \mathbf{F}(H;\theta)}{1 - \mathbf{F}(H;\theta)} \qquad (5.9)
\end{aligned}$$

It follows that the density function is:

$$f^{\star}(x;\theta \mid H) = \mathbb{1}\{x \geq H\} \cdot \frac{f(x;\theta)}{1 - \mathbf{F}(H;\theta)}$$

To estimate the vector of parameters θ, we continue to use the method of maximum likelihood or the generalized method of moments by considering the correction due to the

[9] See Baud *et al.* (2002, 2003) for more advanced techniques based on unknown and stochastic thresholds.

truncation of data. For the ML estimator, we have then:

$$\begin{aligned}\boldsymbol{\ell}(\theta) &= \sum_{i=1}^{T} \ln f^{\star}(x_i;\theta \mid H_i) \\ &= \sum_{i=1}^{T} \ln f(x_i;\theta) + \sum_{i=1}^{T} \ln \mathbb{1}\{x_i \geq H_i\} - \sum_{i=1}^{T} \ln(1-\mathbf{F}(H_i;\theta))\end{aligned} \tag{5.10}$$

where H_i is the threshold associated to the i^{th} observation. The correction term $-\sum_{i=1}^{T} \ln(1-\mathbf{F}(H_i;\theta))$ shows that maximizing a conventional log-likelihood function which ignores data truncation is totally misleading. We also notice that this term vanishes when H_i is equal to zero[10]. For the GMM estimator, the empirical moments become:

$$\begin{cases} h_{i,1}(\theta) = x_i - \mathbb{E}[X \mid X \geq H_i] \\ h_{i,2}(\theta) = (x_i - \mathbb{E}[X \mid X \geq H_i])^2 - \operatorname{var}(X \mid X \geq H_i) \end{cases} \tag{5.11}$$

There is no reason that the conditional moment $\mathbb{E}[X^m \mid X \geq H_i]$ is equal to the unconditional moment $\mathbb{E}[X^m]$. Therefore, the conventional GMM estimator is biased and this is why we have to apply again the threshold correction.

If we consider again the log-normal distribution, the expression of the log-likelihood function (5.10) is[11]:

$$\begin{aligned}\boldsymbol{\ell}(\theta) &= -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln\sigma^2 - \sum_{i=1}^{T}\ln x_i - \frac{1}{2}\sum_{i=1}^{T}\left(\frac{\ln x_i - \mu}{\sigma}\right)^2 - \\ &\quad \sum_{i=1}^{T}\ln\left(1-\Phi\left(\frac{\ln H_i - \mu}{\sigma}\right)\right)\end{aligned}$$

Let us now calculate the conditional moment $\mu'_m(X) = \mathbb{E}[X^m \mid X \geq H]$. By using the notation $\Phi_c(x) = 1 - \Phi((x-\mu)/\sigma)$, we have:

$$\begin{aligned}\mu'_m(X) =& \frac{1}{\Phi_c(\ln H)}\int_H^{\infty} \frac{x^m}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right) \mathrm{d}x \\ =& \frac{1}{\Phi_c(\ln H)}\int_{\ln H}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2 + my\right) \mathrm{d}y \\ =& \frac{\exp\left(m\mu + m^2\sigma^2/2\right)}{\Phi_c(\ln H)}\int_{\ln H}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{y-(\mu+m\sigma^2)}{\sigma}\right)^2\right) \mathrm{d}y \\ =& \frac{\Phi_c\left(\ln H - m\sigma^2\right)}{\Phi_c(\ln H)} \exp\left(m\mu + m^2\sigma^2/2\right)\end{aligned}$$

We deduce that:

$$\mathbb{E}[X \mid X \geq H] = a(\theta, H) = \frac{1-\Phi\left(\frac{\ln H - \mu - \sigma^2}{\sigma}\right)}{1-\Phi\left(\frac{\ln H - \mu}{\sigma}\right)} e^{\mu + \frac{1}{2}\sigma^2}$$

[10] Indeed, we have $\mathbf{F}(0;\theta) = 0$ and $\ln(1-\mathbf{F}(0;\theta)) = 0$.

[11] By construction, the observed value x_i is larger than the threshold H_i, meaning that $\ln \mathbb{1}\{x_i \geq H_i\}$ is equal to 0.

and:

$$\mathbb{E}\left[X^2 \mid X \geq H\right] = b(\theta, H) = \frac{1 - \Phi\left(\frac{\ln H - \mu - 2\sigma^2}{\sigma}\right)}{1 - \Phi\left(\frac{\ln H - \mu}{\sigma}\right)} e^{2\mu + 2\sigma^2}$$

We finally obtain:

$$\left\{ \begin{array}{l} h_{i,1}(\theta) = x_i - a(\theta, H_i) \\ h_{i,2}(\theta) = x_i^2 - 2x_i a(\theta, H_i) + 2a^2(\theta, H_i) - b(\theta, H_i) \end{array} \right.$$

In order to illustrate the impact of the truncation, we report in Figure 5.2 the cumulative distribution function and the probability density function of $X \mid X > H$ when X follows the log-normal distribution $\mathcal{LN}(8, 5)$. The threshold H is set at \$10 000, meaning that the bank collects operational losses when the amount is larger than this threshold. In the bottom panels of the figure, we indicate the mean and the variance with respect to the threshold H. We notice that data truncation increases the magnitude of the mean and the variance. For instance, when H is set at \$10 000, the conditional mean and variance are multiplied by a factor equal to 3.25 with respect to the unconditional mean and variance.

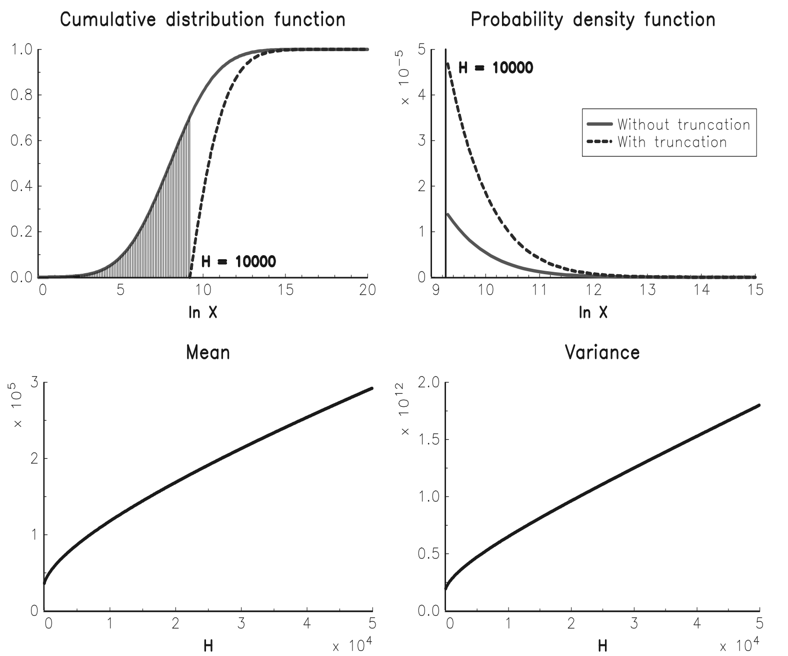

FIGURE 5.2: Impact of the threshold H on the severity distribution

Example 54 *We consider Example 53 and assume that the losses have been collected using a unique threshold that is equal to* \$5 000.

By using the truncation correction, the ML estimates become $\hat{\mu}_{\mathrm{ML}} = 8.00$ and $\hat{\sigma}_{\mathrm{ML}} = 5.71$ for the log-normal model. In Figure 5.3, we compare the log-normal cumulative distribution function without and with the truncation correction. We notice that the results are very different.

The previous example shows that estimating the parameters of the probability distribution is not sufficient to define the severity distribution. Indeed, ML and GMM give

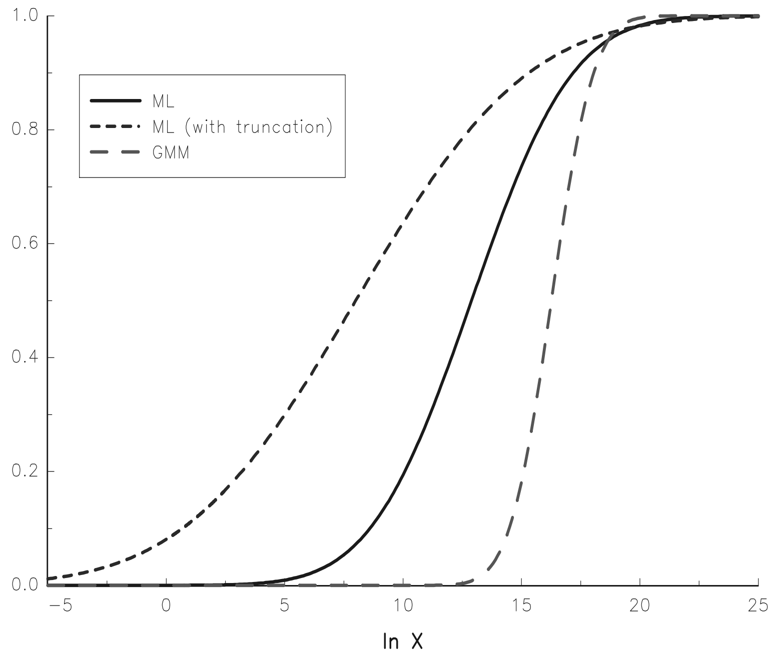

FIGURE 5.3: Comparison of the estimated severity distributions

two different log-normal probability distributions. The issue is to decide which is the best parametrization. In a similar way, the choice between the several probability families (log-normal, log-gamma, GEV, Pareto, etc.) is an open question. This is why fitting the severity distribution does not reduce to estimate the parameters of a given probability distribution. It must be completed by a second step that consists in selecting the best estimated probability distribution. However, traditional goodness-of-fit tests (Kolmogorov-Smirnov, Anderson-Darling, etc.) are not useful, because they concern the entire probability distribution. In operational risk, extreme events are more relevant. This explains why QQ plots and order statistics are generally used to assess the fitting of the upper tail. A QQ plot represents the quantiles of the empirical distribution against those of the theoretical model. If the statistical model describes perfectly the data, we obtain the diagonal line $y = x$. In Figure 5.4, we show an example of QQ plot. We notice that the theoretical quantiles obtained from the statistical model are in line with those calculated with the empirical data when the quantile is lower than 80%. Otherwise, the theoretical quantiles are above the empirical quantiles, meaning that extreme events are underestimated by the statistical model. We deduce that the body of the distribution is well estimated, but not the upper tail of the distribution. However, medium losses are less important than high losses in operational risk.

5.3.2.2 Estimation of the loss frequency distribution

In order to model the frequency distribution, we have to specify the counting process $N(t)$, which defines the number of losses occurring during the time period $[0, t]$. The number of losses for the time period $[t_1, t_2]$ is then equal to:

$$N(t_1; t_2) = N(t_2) - N(t_1)$$

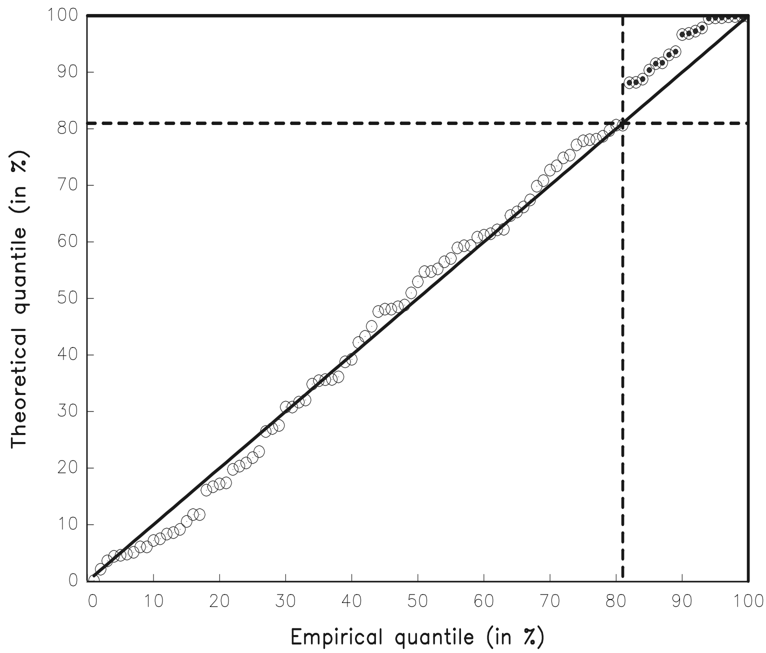

FIGURE 5.4: An example of QQ plot where extreme events are underestimated

We generally made the following statements about the stochastic process $N(t)$:

- the distribution of the number of losses $N(t; t+h)$ for each $h > 0$ is independent of t; moreover, $N(t; t+h)$ is stationary and depends only on the time interval h;
- the random variables $N(t_1; t_2)$ and $N(t_3; t_4)$ are independent if the time intervals $[t_1, t_2]$ and $[t_3, t_4]$ are disjoint;
- no more than one loss may occur at time t.

These simple assumptions define a Poisson process, which satisfies the following properties:

1. there exists a scalar $\lambda > 0$ such that the distribution of $N(t)$ has a Poisson distribution with parameter λt;
2. the duration between two successive losses is *iid* and follows the exponential distribution $\mathcal{E}(\lambda)$.

Let $p(n)$ be the probability to have n losses. We deduce that:

$$\begin{aligned} p(n) &= \Pr\{N(t) = n\} \\ &= \frac{e^{-\lambda t} \cdot (\lambda t)^n}{n!} \end{aligned} \tag{5.12}$$

Without loss of generality, we can fix $t = 1$ because it corresponds to the required one-year time period for calculating the capital charge. In this case, $N(1)$ is simply a Poisson distribution with parameter λ. This probability distribution has a useful property for time aggregation. Indeed, the sum of two independent Poisson variables N_1 and N_2 with parameters λ_1 and λ_2 is also a Poisson variable with parameter $\lambda_1 + \lambda_2$. This property is a direct result of the definition of the Poisson process. In particular, we have:

$$\sum_{k=1}^{K} N\left(\frac{k-1}{K}; \frac{k}{K}\right) = N(1)$$

where $N((k-1)/K; k/K) \sim \mathcal{P}(\lambda/K)$. This means that we can estimate the frequency distribution at a quarterly or monthly period and convert it to an annual period by simply multiplying the quarterly or monthly intensity parameter by 4 or 12.

The estimation of the annual intensity λ can be done using the method of maximum likelihood. In this case, $\hat{\lambda}$ is the mean of the annual number of losses:

$$\hat{\lambda} = \frac{1}{n_y} \sum_{y=1}^{n_y} N_y \tag{5.13}$$

where N_y is the number of losses occurring at year y and n_y is the number of observations. One of the key features of the Poisson distribution is that the variance equals the mean:

$$\lambda = \mathbb{E}[N(1)] = \operatorname{var}(N(1)) \tag{5.14}$$

We can use this property to estimate λ by the method of moments. If we consider the first moment, we obtain the ML estimator, whereas we have with the second moment:

$$\hat{\lambda} = \frac{1}{n_y} \sum_{y=1}^{n_y} \left(N_y - \bar{N}\right)^2$$

where $\bar{N}$ is the average number of losses.

Example 55 *We assume that the annual number of losses from 2006 to 2015 is the following:* 57*,* 62*,* 45*,* 24*,* 82*,* 36*,* 98*,* 75*,* 76 *and* 45*.*

The mean is equal to 60 whereas the variance is equal to 474.40. In Figure 5.5, we show the probability mass function of the Poisson distribution with parameter 60. We notice that the parameter λ is not enough large to reproduce the variance and the range of the sample. However, using the moment estimator based on the variance is completely unrealistic.

When the variance exceeds the mean, we use the negative binomial distribution $\mathcal{NB}(r, p)$, which is defined as follows:

$$\begin{aligned} p(n) &= \binom{r+n-1}{n} (1-p)^r p^n \\ &= \frac{\Gamma(r+n)}{n!\,\Gamma(r)} (1-p)^r p^n \end{aligned}$$

where $r > 0$ and $p \in [0,1]$. The negative binomial distribution can be viewed as the probability distribution of the number of successes in a sequence of *iid* Bernoulli random variables $\mathcal{B}(p)$ until we get r failures. The negative binomial distribution is then a generalization of the geometric distribution. Concerning the first two moments, we have:

$$\mathbb{E}[\mathcal{NB}(r,p)] = \frac{p \cdot r}{1-p}$$

and:

$$\operatorname{var}(\mathcal{NB}(r,p)) = \frac{p \cdot r}{(1-p)^2}$$

We verify that:

$$\begin{aligned} \operatorname{var}(\mathcal{NB}(r,p)) &= \frac{1}{1-p} \cdot \mathbb{E}[\mathcal{NB}(r,p)] \\ &> \mathbb{E}[\mathcal{NB}(r,p)] \end{aligned}$$

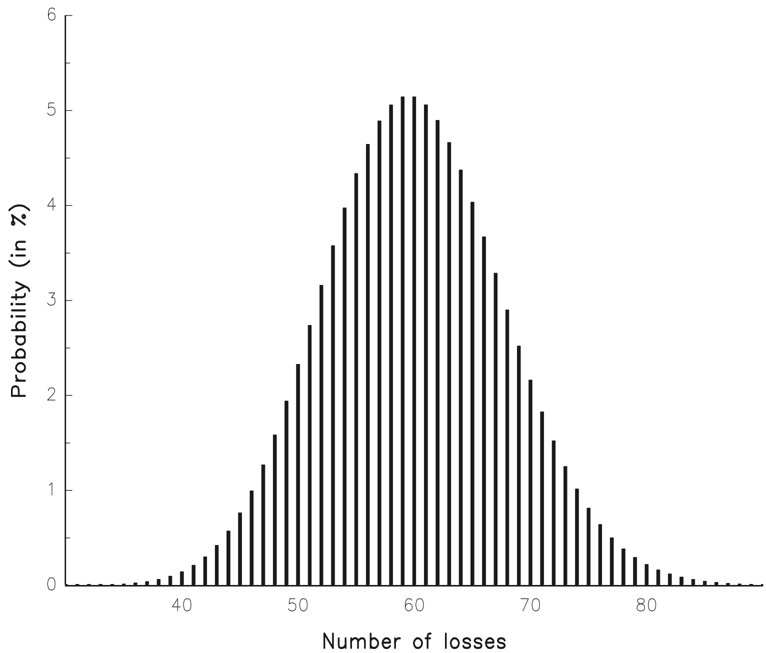

FIGURE 5.5: PMF of the Poisson distribution $\mathcal{P}(60)$

Remark 62 *The negative binomial distribution corresponds to a Poisson process where the intensity parameter is random and follows a gamma distribution*[12]*:*

$$\mathcal{NB}(r,p) \sim \mathcal{P}(\Lambda) \quad \text{and} \quad \Lambda \sim \mathcal{G}(\alpha,\beta)$$

where $\alpha = r$ *and* $\beta = (1-p)/p$.

We consider again Example 55 and assume that the number of losses is described by the negative binomial distribution. Using the method of moments, we obtain the following estimates:

$$\hat{r} = \frac{m^2}{v-m} = \frac{60^2}{474.40-60} = 8.6873$$

and

$$\hat{p} = \frac{v-m}{v} = \frac{474.40-60}{474.40} = 0.8735$$

where m is the mean and v is the variance of the sample. Using these estimates as the starting values of the numerical optimization procedure, the ML estimates are $\hat{r} = 7.7788$ and $\hat{p} = 0.8852$. We report the corresponding probability mass function $p(n)$ in Figure 5.6. We notice that this distribution better describes the sample that the Poisson distribution, because it has a larger support. In fact, we show in Figure 5.7 the probability density function of λ for the two estimated counting processes. For the Poisson distribution, λ is constant and equal to 60, whereas λ has a gamma distribution $\mathcal{G}(7.7788, 0.1296)$ in the case of the negative binomial distribution. The variance of the gamma distribution explains the larger variance of the negative binomial distribution with respect to the Poisson distribution, while we notice that the two distributions have the same mean.

[12] See Exercise 5.4.6 on page 346.

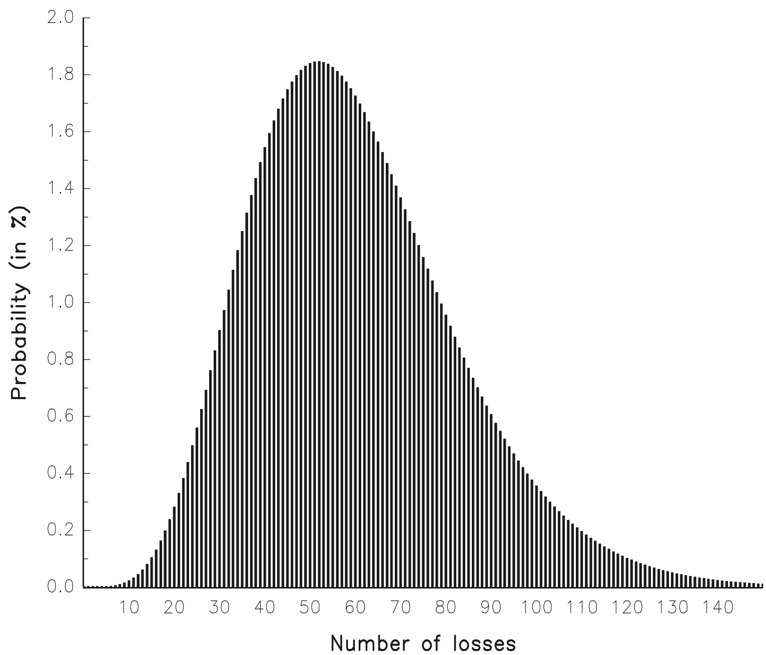

FIGURE 5.6: PMF of the negative binomial distribution

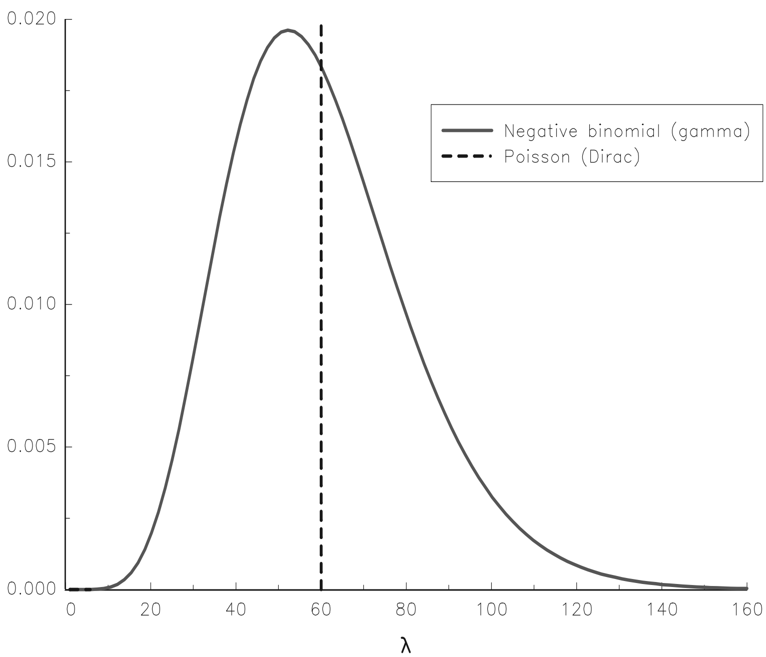

FIGURE 5.7: Probability density function of the parameter λ

As in the case of the severity distribution, data truncation and reporting bias have an impact of the frequency distribution (Frachot *et al.*, 2006). For instance, if one bank's reporting threshold H is set at a high level, then the average number of reported losses will be low. It does not imply that the bank is allowed to have a lower capital charge than another bank that uses a lower threshold and is otherwise identical to the first one. It simply means that the average number of losses must be corrected for reporting bias as well. It appears that the calibration of the frequency distribution comes as a second step (after having calibrated the severity distribution) because the aforementioned correction needs an estimate of the exceedance probability $\Pr\{X > H\}$ for its calculation. This is rather straightforward: the difference (more precisely the ratio) between the number of reported events and the '*true*' number of events (which would be obtained if all the losses were reported, *i.e.* with a zero-threshold) corresponds exactly to the probability of one loss being higher than the threshold. This probability is a direct by-product of the severity distribution.

Let $N_H(t)$ be the number of events that are larger than the threshold H. By definition, $N_H(t)$ is the counting process of exceedance events:

$$N_H(t) = \sum_{i=1}^{N(t)} \mathbb{1}\{X_i > H\}$$

It follows that:

$$\begin{aligned}
\mathbb{E}[N_H(t)] &= \mathbb{E}\left[\sum_{i=1}^{N(t)} \mathbb{1}\{X_i > H\}\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n} \mathbb{1}\{X_i > H\} \middle| N(t) = n\right] \\
&= \mathbb{E}[N(t)] \cdot \mathbb{E}[\mathbb{1}\{X_i > H\}]
\end{aligned}$$

because the random variables $X_1, \ldots, X_n$ are *iid* and independent from the random number of events $N(t)$. We deduce that:

$$\begin{aligned}
\mathbb{E}[N_H(t)] &= \mathbb{E}[N(t)] \cdot \Pr\{X_i > H\} \\
&= \mathbb{E}[N(t)] \cdot (1 - \mathbf{F}(H;\theta))
\end{aligned} \tag{5.15}$$

This latter equation provides information about the transformation of the counting process $N(t)$ into the exceedance process. However, it only concerns the mean and not the distribution itself. One interesting feature of data truncation is when the distribution of the threshold exceedance process belongs to the same distribution class of the counting process. It is the case of the Poisson distribution:

$$\mathbf{P}_H(\lambda) = \mathbf{P}(\lambda_H)$$

Using Equation (5.15), it follows that the Poisson parameter λ_H of the exceedance process is simply the product of the Poisson parameter λ by the exceedance probability $\Pr\{X > H\}$:

$$\lambda_H = \lambda \cdot (1 - \mathbf{F}(H;\theta))$$

We deduce that the estimator $\hat{\lambda}$ has the following expression:

$$\hat{\lambda} = \frac{\hat{\lambda}_H}{1 - \mathbf{F}\left(H;\hat{\theta}\right)}$$

where $\hat{\lambda}_H$ is the average number of losses that are collected above the threshold H and $\mathbf{F}\left(x;\hat{\theta}\right)$ is the parametric estimate of the severity distribution.

Example 56 *We consider that the bank has collected the loss data from 2006 to 2015 with a threshold of* \$20 000. *For a given event type, the calibrated severity distribution corresponds to a log-normal distribution with parameters* $\hat{\mu} = 7.3$ *and* $\hat{\sigma} = 2.1$*, whereas the annual number of losses is the following:* 23, 13, 50, 12, 25, 36, 48, 27, 18 *and* 35.

Using the Poisson distribution, we obtain $\hat{\lambda}_H = 28.70$. The probability that the loss exceeds the threshold H is equal to:

$$\Pr\{X > 20\,000\} = 1 - \Phi\left(\frac{\ln(20\,000) - 7.3}{2.1}\right) = 10.75\%$$

This means that only 10.75% of losses can be observed when we apply a threshold of \$20 000. We then deduce that the estimate of the Poisson parameter is equal to:

$$\hat{\lambda} = \frac{28.70}{10.75\%} = 266.90$$

On average, there are in fact about 270 loss events per year.

We could discuss whether the previous result remains valid in the case of the negative binomial distribution. If it is the case, then we have:

$$\mathbf{P}_H(r, p) = \mathbf{P}(r_H, p_H)$$

Using Equation (5.15), we deduce that:

$$\frac{p_H \cdot r_H}{1 - p_H} = \frac{p \cdot r}{1 - p} \cdot (1 - \mathbf{F}(H; \theta))$$

If we assume that r_H is equal to r, we obtain:

$$p_H = \frac{p \cdot (1 - \mathbf{F}(H; \theta))}{1 - p \cdot \mathbf{F}(H; \theta)}$$

We verify the following inequality $p \leq p_H \leq 1$. However, this solution is not completely satisfactory.

5.3.3 Calculating the capital charge

Once the frequency and severity distributions are calibrated, the computation of the capital charge is straightforward. For that, we can use the Monte Carlo method or different analytical methods. The Monte Carlo method is much more used, because it is more flexible and gives better results in the case of low frequency/high severity events. Analytical approaches, which are very popular in insurance, can be used for high frequency/low severity events. One remaining challenge, however, is aggregating the capital charge of the different cells of the mapping matrix. By construction, the loss distribution approach assumes that aggregate losses are independent. Nevertheless, regulation are forcing banks to take into account positive correlation between risk events. The solution is then to consider copula functions.

5.3.3.1 Monte Carlo approach

We reiterate that the one-year compound loss of a given cell is defined as follows:

$$S = \sum_{i=1}^{N(1)} X_i$$

where $X_i \sim \mathbf{F}$ and $N(1) \sim \mathbf{P}$. The capital-at-risk is then the 99% quantile of the compound loss distribution. To estimate the capital charge by Monte Carlo, we first simulate the annual number of losses from the frequency distribution and then simulate individual losses in order to calculate the compound loss. Finally, the quantile is estimated by order statistics. The algorithm is described below.

Algorithm 1 Compute the capital-at-risk for an operational risk cell

```
Initialize the number of simulations n_S
for j = 1 : n_S do
    Simulate an annual number n of losses from the frequency distribution P
    S_j ← 0
    for i = 1 : n do
        Simulate a loss X_i from the severity distribution F
        S_j = S_j + X_i
    end for
end for
Calculate the order statistics S_{1:n_S}, ..., S_{n_S:n_S}
Deduce the capital-at-risk CaR = S_{αn_S:n_S} with α = 99.9%
return CaR
```

Let us illustrate this algorithm when $N(1) \sim \mathcal{P}(4)$ and $X_i \sim \mathcal{LN}(8, 4)$. Using a linear congruential method, the simulated values of $N(1)$ are 3, 4, 1, 2, 3, etc. while the simulated values of X_i are 3388.6, 259.8, 13328.3, 39.7, 1220.8, 1486.4, 15197.1, 3205.3, 5070.4, 84704.1, 64.9, 1237.5, 187073.6, 4757.8, 50.3, 2805.7, etc. For the first simulation, we have three losses and we obtain:

$$S_1 = 3388.6 + 259.8 + 13328.3 = \$16\,976.7$$

For the second simulation, the number of losses is equal to four and the compound loss is equal to:

$$S_2 = 39.7 + 1220.8 + 1486.4 + 15197.1 = \$17\,944.0$$

For the third simulation, we obtain $S_3 = \$3\,205.3$, and so on. Using n_S simulations, the value of the capital charge is estimated with the 99.9% empirical quantile based on order statistics. For instance, Figure 5.8 shows the histogram of 2 000 simulated values of the capital-at-risk estimated with one million simulations. The true value is equal to \$3.24 mn. However, we notice that the variance of the estimator is large. Indeed, the range of the MC estimator is between \$3.10 mn and \$3.40 mn in our experiments with one million simulation runs.

The estimation of the capital-at-risk with a high accuracy is therefore difficult. The convergence of the Monte Carlo algorithm is low and the estimated quantile can be very far from the true quantile especially when the severity loss distribution is heavy tailed and the confidence level α is high. That's why it is important to control the accuracy of $\mathbf{G}^{-1}(\alpha)$. This can be done by verifying that the estimated moments are close to the theoretical ones. For the first two central moments, we have:

$$\mathbb{E}[S] = \mathbb{E}[N(1)] \cdot \mathbb{E}[X_i]$$

and:

$$\operatorname{var}(S) = \mathbb{E}[N(1)] \cdot \operatorname{var}(X_i) + \operatorname{var}(N(1)) \cdot \mathbb{E}^2[X_i]$$

To illustrate the convergence problem, we consider the example of the compound Poisson distribution where $N(1) \sim P(10)$ and $X_i \sim \mathcal{LN}(5, \sigma^2)$. We compute the aggregate loss

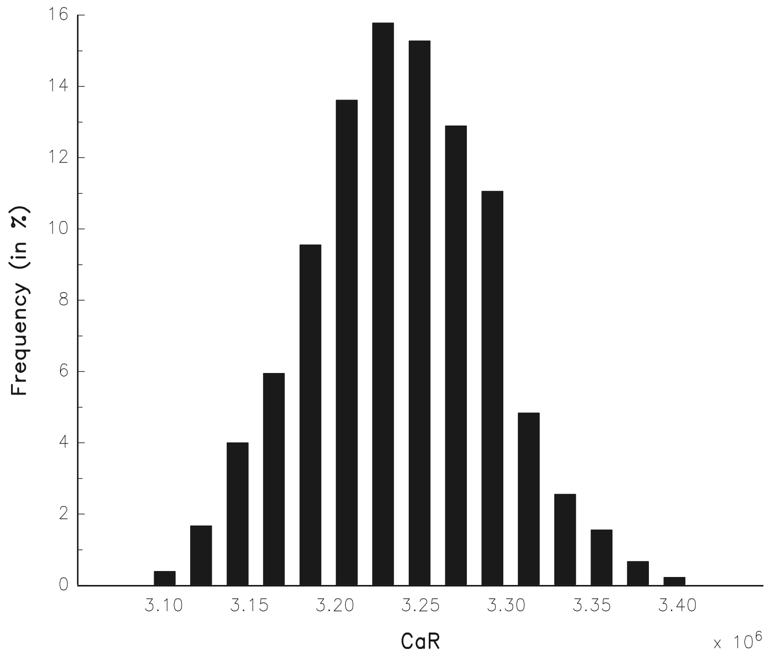

FIGURE 5.8: Histogram of the MC estimator $\widehat{\text{CaR}}$

distribution by the Monte Carlo method for different number n_S of simulations and different runs. To measure the accuracy, we calculate the ratio between the MC standard deviation $\hat{\sigma}_{n_S}(S)$ and the true value $\sigma(S)$:

$$R(n_s) = \frac{\hat{\sigma}_{n_S}(S)}{\sigma(S)}$$

We notice that the convergence is much more erratic when σ takes a high value (Figure 5.10) than when σ is low (Figure 5.9). When σ takes the value 1, the convergence of the Monte Carlo method is verified with 100 000 simulations. When σ takes the value 2.5, 100 million simulations are not sufficient to estimate the second moment, and then the capital-at-risk. Indeed, the occurrence probability of extreme events is generally underestimated. Sometimes, a severe loss is simulated implying a jump in the empirical standard deviation (see Figure 5.10). This is why we need a large number of simulations in order to be confident when estimating the 99.9% capital-at-risk with high severity distributions.

Remark 63 *With the Monte Carlo approach, we can easily integrate mitigation factors such as insurance coverage. An insurance contract is generally defined by a deductive*[13] A *and the maximum amount* B *of a loss, which is covered by the insurer. The effective loss* $\tilde{X}_i$ *suffered by the bank is then the difference between the loss of the event and the amount paid by the insurer:*

$$\tilde{X}_i = X_i - \max(\min(X_i, B) - A, 0)$$

The relationship between X_i *and* $\tilde{X}_i$ *is shown in Figure 5.11. In this case, the annual loss of the bank becomes:*

$$S = \sum_{i=1}^{N(1)} \tilde{X}_i$$

[13]It corresponds to the loss amount the bank has to cover by itself.

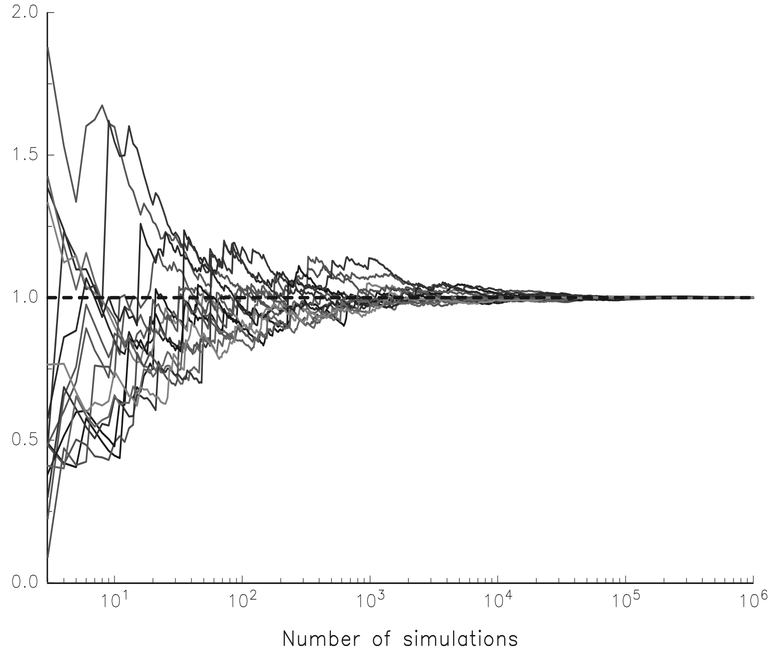

FIGURE 5.9: Convergence of the accuracy ratio $R(n_s)$ when $\sigma = 1$

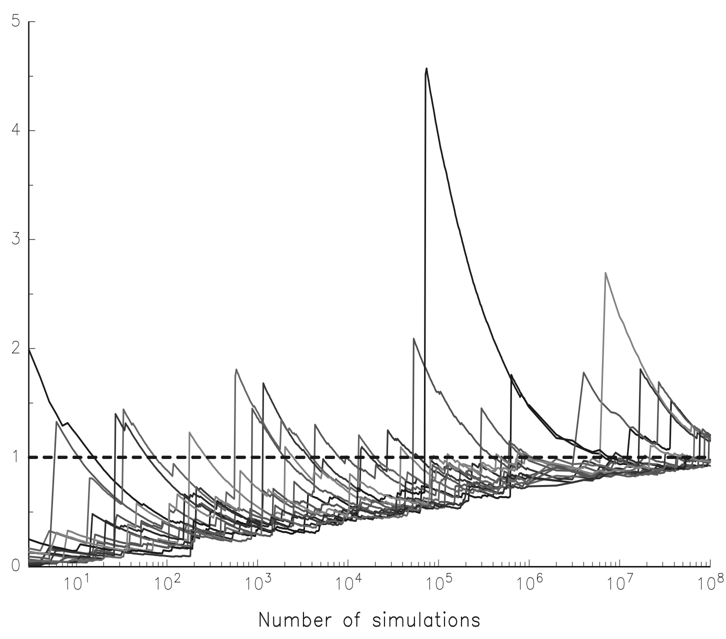

FIGURE 5.10: Convergence of the accuracy ratio $R(n_s)$ when $\sigma = 2.5$

Taking into account an insurance contract is therefore equivalent to replace X_i by $\tilde{X}_i$ in the Monte Carlo simulations.

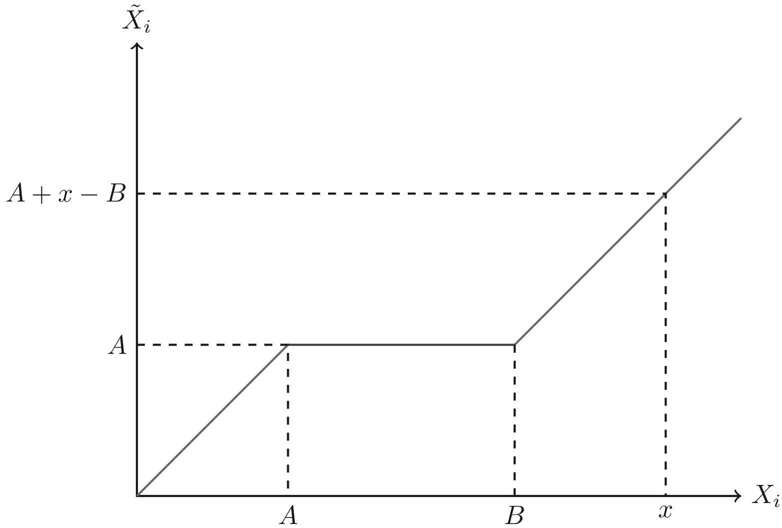

FIGURE 5.11: Impact of the insurance contract on the operational risk loss

5.3.3.2 Analytical approaches

There are three analytical (or semi-analytical) methods to compute the aggregate loss distribution: the solution based on characteristic functions, Panjer recursion and the single loss approximation.

Method of characteristic functions Formally, the characteristic function of the random variable X is defined by:

$$\varphi_X(t) = \mathbb{E}\left[e^{itX}\right]$$

If X has a continuous probability distribution $\mathbf{F}$, we obtain:

$$\varphi_X(t) = \int_0^\infty e^{itx}\,\mathrm{d}\mathbf{F}(x)$$

We notice that the characteristic function of the sum of n independent random variables is the product of their characteristic functions:

$$\begin{aligned}
\varphi_{X_1+\ldots+X_n}(t) &= \mathbb{E}\left[e^{it(X_1+X_2+\cdots+X_n)}\right] \\
&= \prod_{i=1}^n \mathbb{E}\left[e^{itX_i}\right] \\
&= \prod_{i=1}^n \varphi_{X_i}(t)
\end{aligned}$$

It comes that the characteristic function of the compound distribution $\mathbf{G}$ is given by:

$$\varphi_S(t) = \sum_{n=0}^\infty p(n)\left(\varphi_X(t)\right)^n = \varphi_{N(1)}\left(\varphi_X(t)\right)$$

where $\varphi_{N(1)}(t)$ is the probability generating function of $N(1)$. For example, if $N(1) \sim \mathcal{P}(\lambda)$, we have:

$$\varphi_{N(1)}(t) = e^{\lambda(t-1)}$$

and:

$$\varphi_S(t) = e^{\lambda(\varphi_X(t)-1)}$$

We finally deduce that S has the probability density function given by the Laplace transform of $\varphi_S(t)$:

$$g(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \varphi_S(t)\, \mathrm{d}t$$

Using this expression, we can easily compute the cumulative distribution function and its inverse with the fast fourier transform.

Panjer recursive approach Panjer (1981) introduces recursive approaches to compute high-order convolutions. He showed that if the probability mass function of the counting process $N(t)$ satisfies:

$$p(n) = \left(a + \frac{b}{n}\right) p(n-1)$$

where a and b are two scalars, then the following recursion holds:

$$g(x) = p(1)\, f(x) + \int_0^x \left(a + b\frac{y}{x}\right) f(y)\, g(x-y)\, \mathrm{d}y$$

where $x > 0$. For discrete severity distributions satisfying $f_n = \Pr\{X_i = n\delta\}$ where δ is the monetary unit (e.g. \$10 000), the Panjer recursion becomes:

$$\begin{aligned} g_n &= \Pr\{S = n\delta\} \\ &= \frac{1}{1 - af_0} \sum_{j=1}^{n} \left(a + \frac{bj}{n}\right) f_j g_{n-j} \end{aligned}$$

where:

$$\begin{aligned} g_0 &= \sum_{n=0}^{\infty} p(n)\,(f_0)^n \\ &= \begin{cases} p(0)\, e^{bf_0} & \text{if } a = 0 \\ p(0)\,(1 - af_0)^{-1-b/a} & \text{otherwise} \end{cases} \end{aligned}$$

The capital-at-risk is then equal to:

$$\mathrm{CaR}(\alpha) = n^\star \delta$$

where:

$$n^\star = \inf \left\{ n : \sum_{j=0}^{n} g_j \geq \alpha \right\}$$

Like the method of characteristic functions, the Panjer recursion is very popular among academics, but produces significant numerical errors in practice when applied to operational risk losses. The issue is the support of the compound distribution, whose range can be from zero to several billions.

Example 57 *We consider the compound Poisson distribution with log-normal losses and different sets of parameters:*

(a) $\lambda = 5$, $\mu = 5$, $\sigma = 1.0$*;*

(b) $\lambda = 5$, $\mu = 5$, $\sigma = 1.5$*;*

(c) $\lambda = 5$, $\mu = 5$, $\sigma = 2.0$*;*

(d) $\lambda = 50$, $\mu = 5$, $\sigma = 2.0$.

In order to implement the Panjer recursion, we have to perform a discretization of the severity distribution. Using the central difference approximations, we have:

$$\begin{aligned} f_n &= \Pr\left\{n\delta - \frac{\delta}{2} \leq X_i \leq n\delta + \frac{\delta}{2}\right\} \\ &= \mathbf{F}\left(n\delta + \frac{\delta}{2}\right) - \mathbf{F}\left(n\delta - \frac{\delta}{2}\right) \end{aligned}$$

To initialize the algorithm, we use the convention $f_0 = \mathbf{F}(\delta/2)$. In Figure 5.12, we compare the cumulative distribution function of the aggregate loss obtained with the Panjer recursion and Monte Carlo simulations[14]. We deduce the capital-at-risk for different values of α in Table 5.5. In our case, the Panjer algorithm gives a good approximation, because the support of the distribution is '*bounded* '. When the aggregate loss can take very large values, we need a lot of iterations to achieve the convergence[15]. Moreover, we may have underflow in computations because $g_0 \approx 0$.

TABLE 5.5: Comparison of the capital-at-risk calculated with Panjer recursion and Monte Carlo simulations

α	Panjer recursion				Monte Carlo simulations			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
90%	2400	4500	11000	91000	2350	4908	11648	93677
95%	2900	6500	19000	120000	2896	6913	19063	123569
99%	4300	13500	52000	231000	4274	13711	51908	233567
99.5%	4900	18000	77000	308000	4958	17844	77754	310172
99.9%	6800	32500	182000	604000	6773	32574	185950	604756

Single loss approximation If the severity belongs to the family of subexponential distributions, then Böcker and Klüppelberg (2005) and Böcker and Sprittulla (2006) show that the percentile of the compound distribution can be approximated by the following expression:

$$\mathbf{G}^{-1}(\alpha) \approx (\mathbb{E}[N(1)] - 1) \cdot \mathbb{E}[X_i] + \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{\mathbb{E}[N(1)]}\right) \tag{5.16}$$

It follows that the capital-at-risk is the sum of the expected loss and the unexpected loss defined as follows:

$$\begin{aligned} \mathrm{EL} &= \mathbb{E}[N(1)] \cdot \mathbb{E}[X_i] \\ \mathrm{UL}(\alpha) &= \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{N(1)}\right) - \mathbb{E}[X_i] \end{aligned}$$

[14]We use one million simulations.

[15]In this case, it is not obvious that the Panjer recursion is faster than Monte Carlo simulations.

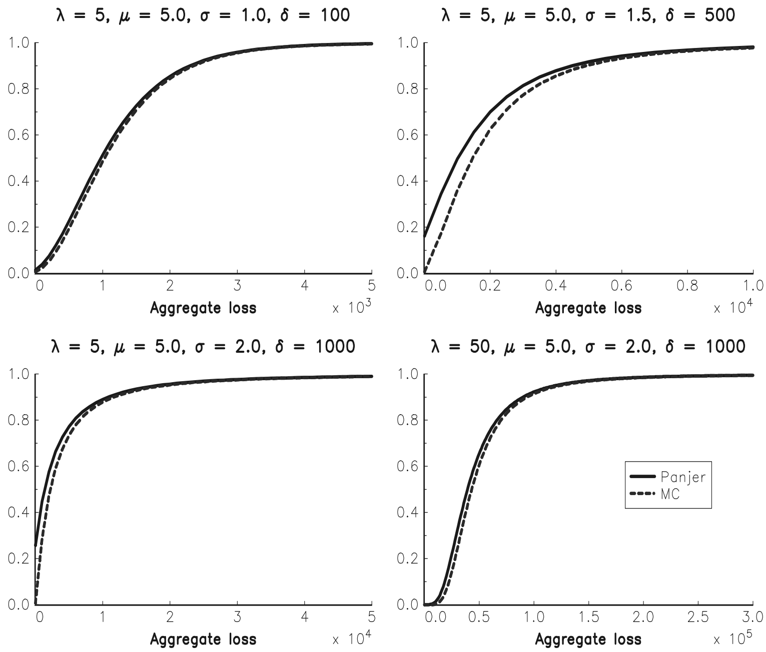

FIGURE 5.12: Comparison between the Panjer and MC compound distributions

To understand Formula (5.16), we recall that subexponential distributions are a special case of heavy-tailed distributions and satisfy the following property:

$$\lim_{x \to \infty} \frac{\Pr\{X_1 + \cdots + X_n > x\}}{\Pr\{\max(X_1, \ldots, X_n) > x\}} = 1$$

This means that large values of the aggregate loss are dominated by the maximum loss of one event. If we decompose the capital-at-risk as a sum of risk contributions, we obtain:

$$\mathbf{G}^{-1}(\alpha) = \sum_{i=1}^{\mathbb{E}[N(1)]} \mathcal{RC}_i$$

where:

$$\mathcal{RC}_i = \mathbb{E}[X_i] \qquad \text{for } i \neq i^\star$$

and:

$$\mathcal{RC}_{i^\star} = \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{N(1)}\right)$$

In this model, the capital-at-risk is mainly explained by the single largest loss $i^\star$. If we neglect the small losses, the capital-at-risk at the confidence level α_{CaR} is related to the quantile $\alpha_{\mathrm{Severity}}$ of the loss severity:

$$\alpha_{\mathrm{Severity}} = 1 - \frac{1-\alpha_{\mathrm{CaR}}}{N(1)}$$

This relationship[16] is shown in Figure 5.13 and explains why this framework is called the single loss approximation (SLA). For instance, if the annual number of losses is equal to 100 on average, computing the capital-at-risk with a 99.9% confidence level is equivalent to estimate the quantile 99.999% of the loss severity.

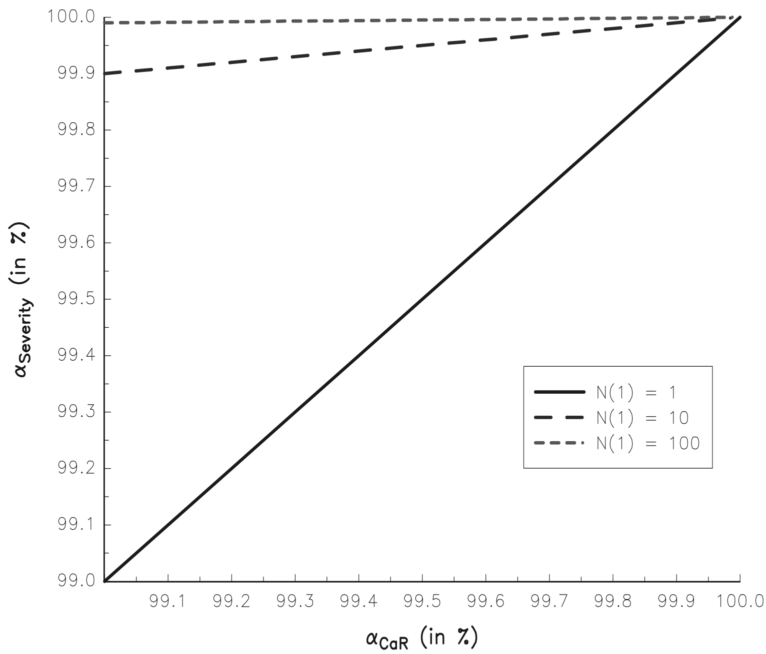

FIGURE 5.13: Relationship between α_{CaR} and α_{Severity}

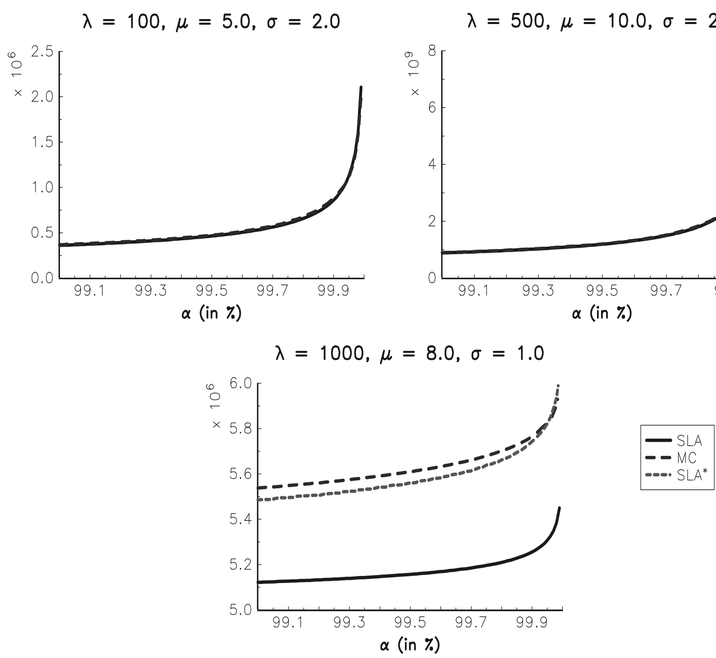

FIGURE 5.14: Numerical illustration of the single loss approximation

The most popular subexponential distributions used in operational risk modeling are the log-gamma, log-logistic, log-normal and Pareto probability distributions (BCBS, 2014f). For instance, if $N(1) \sim \mathcal{P}(\lambda)$ and $X_i \sim \mathcal{LN}(\mu, \sigma^2)$, we obtain:

$$\mathrm{EL} = \lambda \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

and:

$$\mathrm{UL}(\alpha) = \exp\left(\mu + \sigma\Phi^{-1}\left(1 - \frac{1-\alpha}{\lambda}\right)\right) - \exp\left(\mu + \frac{1}{2}\sigma^2\right)$$

In Figure 5.14, we report the results of some experiments for different values of parameters. In the top panels, we assume that $\lambda = 100$, $\mu = 5.0$ and $\sigma = 2.0$ (left panel), and $\lambda = 500$, $\mu = 10.0$ and $\sigma = 2.5$ (right panel). These two examples correspond to medium severity/low frequency and high severity/low frequency events. In these cases, we obtain a good approximation. In the bottom panel, the parameters are $\lambda = 1000$, $\mu = 8.0$ and $\sigma = 1.0$. The approximation does not work very well, because we have a low severity/high frequency events and the risk can then not be explained by an extreme single loss. The underestimation of the capital-at-risk is due to the underestimation of the number of losses. In fact, with low severity/high frequency events, the risk is not to face a large single loss, but to have a high number of losses in the year. This is why it is better to approximate the capital-at-risk with the following formula:

$$\mathbf{G}^{-1}(\alpha) \approx \left(\mathbf{P}^{-1}(\alpha) - 1\right)\mathbb{E}[X_i] + \mathbf{F}^{-1}\left(1 - \frac{1-\alpha}{\mathbb{E}[N(1)]}\right)$$

where $\mathbf{P}$ is the cumulative distribution function of the counting process $N(1)$. In Figure 5.14, we have also reported this approximation SLA* for the third example. We verify that it gives better results for high frequency events than the classic approximation.

5.3.3.3 Aggregation issues

We recall that the loss at the bank level is equal to:

$$L = \sum_{k=1}^{K} S_k$$

where S_k is the aggregate loss of the k^{th} cell of the mapping matrix. For instance, if the matrix is composed of the eight business lines (BL) and seven even types (ET) of the Basel II classification, we have $L = \sum_{k \in \mathcal{K}} S_k$ where $\mathcal{K} = \{(\mathrm{BL}_{k_1}, \mathrm{ET}_{k_2}), k_1 = 1, \ldots, 8; k_2 = 1, \ldots, 7\}$. Let $\mathrm{CaR}_{k_1,k_2}(\alpha)$ be the capital charge calculated for the business line k_1 and the event type k_2. We have:

$$\mathrm{CaR}_{k_1,k_2}(\alpha) = \mathbf{G}_{k_1,k_2}^{-1}(\alpha)$$

One solution to calculate the capital charge at the bank level is to sum up all the capital charges:

$$\begin{aligned}\mathrm{CaR}(\alpha) &= \sum_{k_1=1}^{8}\sum_{k_2=1}^{7} \mathrm{CaR}_{k_1,k_2}(\alpha) \\ &= \sum_{k_1=1}^{8}\sum_{k_2=1}^{7} \mathbf{G}_{k_1,k_2}^{-1}(\alpha)\end{aligned}$$

[16] In Chapter 12, we will see that such transformation is common in extreme value theory.

From a theoretical point of view, this is equivalent to assume that all the aggregate losses S_k are perfectly correlated. This approach is highly conservative and ignores diversification effects between business lines and event types.

Let us consider the two-dimensional case:

$$\begin{aligned} L &= S_1 + S_2 \\ &= \sum_{i=1}^{N_1} X_i + \sum_{j=1}^{N_2} Y_j \end{aligned}$$

In order to take into account the dependence between the two aggregate losses S_1 and S_2, we can assume that frequencies N_1 and N_2 are correlated or severities X_i and Y_j are correlated. Thus, the aggregate loss correlation $\rho(S_1, S_2)$ depends on two key parameters:

- the frequency correlation $\rho(N_1, N_2)$;
- the severity correlation $\rho(X_i, Y_j)$.

For example, we should observe that historically, the number of external fraud events is high (respectively low) when the number of internal fraud events is also high (respectively low). Severity correlation is more difficult to justify. In effect, a basic feature of the LDA model requires assuming that individual losses are jointly independent. Therefore it is conceptually difficult to assume simultaneously severity independence within each class of risk and severity correlation between two classes. By assuming that $\rho(X_i, Y_j) = 0$, Frachot *et al. (*2004) find an upper bound of the aggregate loss correlation. We have:

$$\begin{aligned} \operatorname{cov}(S_1, S_2) &= \mathbb{E}[S_1 S_2] - \mathbb{E}[S_1] \cdot \mathbb{E}[S_2] \\ &= \mathbb{E}\left[\sum_{i=1}^{N_1} X_i \cdot \sum_{j=1}^{N_2} Y_j\right] - \mathbb{E}\left[\sum_{i=1}^{N_1} X_i\right] \cdot \mathbb{E}\left[\sum_{j=1}^{N_2} Y_j\right] \\ &= \mathbb{E}[N_1 N_2] \cdot \mathbb{E}[X_i] \cdot \mathbb{E}[Y_j] - \mathbb{E}[N_1] \cdot \mathbb{E}[X_i] \cdot \mathbb{E}[N_2] \cdot \mathbb{E}[Y_j] \\ &= (\mathbb{E}[N_1 N_2] - \mathbb{E}[N_1] \cdot \mathbb{E}[N_2]) \cdot \mathbb{E}[X_i] \cdot \mathbb{E}[Y_j] \end{aligned}$$

and:

$$\rho(S_1, S_2) = \frac{(\mathbb{E}[N_1 N_2] - \mathbb{E}[N_1] \cdot \mathbb{E}[N_2]) \cdot \mathbb{E}[X_i] \cdot \mathbb{E}[Y_j]}{\sqrt{\operatorname{var}(S_1) \cdot \operatorname{var}(S_2)}}$$

If we assume that the counting processes N_1 and N_2 are Poisson processes with parameters λ_1 and λ_2, we obtain:

$$\rho(S_1, S_2) = \rho(N_1, N_2) \cdot \eta(X_i) \cdot \eta(Y_j)$$

where:

$$\begin{aligned} \eta(X) &= \frac{\mathbb{E}[X]}{\sqrt{\mathbb{E}[X^2]}} \\ &= \frac{1}{\sqrt{1 + \operatorname{CV}^2(X)}} \leq 1 \end{aligned}$$

Here $\operatorname{CV}(X) = \sigma(X) / \mathbb{E}[X]$ denotes the coefficient of variation of the random variable X. As a result, aggregate loss correlation is always lower than frequency correlation:

$$0 \leq \rho(S_1, S_2) \leq \rho(N_1, N_2) \leq 1$$

We deduce that an upper bound of the aggregate loss correlation is equal to:

$$\rho^+ = \eta(X_i) \cdot \eta(Y_j)$$

For high severity events, severity independence likely dominates frequency correlation and we obtain $\rho^+ \simeq 0$ because $\eta(X_i) \simeq 0$.

Let us consider the example of log-normal severity distributions. We have:

$$\rho^+ = \exp\left(-\frac{1}{2}\sigma_X^2 - \frac{1}{2}\sigma_Y^2\right)$$

We notice that this function is decreasing with respect to σ_X and σ_Y. Figure 5.15 shows the relationship between σ_X, σ_Y and ρ^+. We verify that ρ^+ is small when σ_X and σ_Y take large values. For instance, if $\sigma_X = \sigma_Y = 2$, the aggregate loss correlation is lower than 2%.

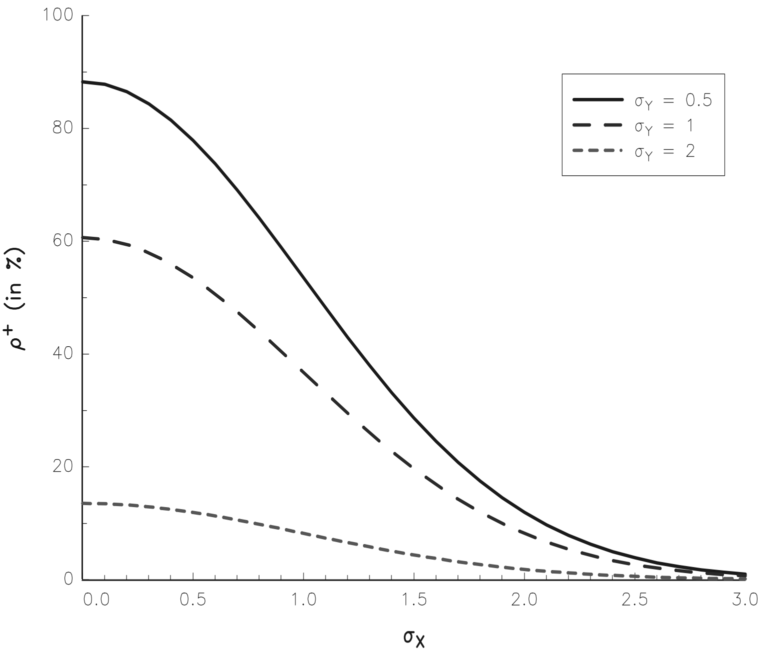

FIGURE 5.15: Upper bound ρ^+ of the aggregate loss correlation

There are two ways to take into account correlations for computing the capital charge of the bank. The first approach is to consider the normal approximation:

$$\mathrm{CaR}(\alpha) = \sum_k \mathrm{EL}_k + \sqrt{\sum_{k,k'} \rho_{k,k'} \cdot (\mathrm{CaR}_k(\alpha) - \mathrm{EL}_k) \cdot (\mathrm{CaR}_{k'}(\alpha) - \mathrm{EL}_{k'})}$$

where $\rho_{k,k'}$ is the correlation between the cells k and k' of the mapping matrix. The second approach consists in introducing the dependence between the aggregate losses using a copula function $\mathbf{C}$. The joint distribution of $(S_1, \ldots, S_K)$ has the following form:

$$\Pr\{S_1 \leq s_1, \ldots, S_K \leq s_K\} = \mathbf{C}(\mathbf{G}_1(s_1), \ldots, \mathbf{G}_K(s_K))$$

where $\mathbf{G}_k$ is the cumulative distribution function of the k^{th} aggregate loss S_k. In this case, the quantile of the random variable $L = \sum_{k=1}^K S_k$ is estimated using Monte Carlo simulations. The difficulty comes from the fact that the distributions $\mathbf{G}_k$ have no analytical expression. The solution is then to use the method of empirical distributions, which is presented on page 806.

5.3.4 Incorporating scenario analysis

The concept of scenario analysis should deserve further clarification. Roughly speaking, when we refer to scenario analysis, we want to express the idea that banks' experts and experienced managers have some reliable intuitions on the riskiness of their business and that these intuitions are not entirely reflected in the bank's historical internal data. As a first requirement, we expect that experts should have the opportunity to give their approval to capital charge results. In a second step, one can imagine that experts' intuitions are directly plugged into severity and frequency estimations. Experts' intuition can be captured through scenario building. More precisely, a scenario is given by a potential loss amount and the corresponding probability of occurrence. As an example, an expert may assert that a loss of one million dollars or higher is expected to occur once every (say) 5 years. This is a valuable information in many cases, either when loss data are rare and do not allow for statistically sound results or when historical loss data are not sufficiently forward-looking. In this last case, scenario analysis allows to incorporate external loss data.

In what follows, we show how scenarios can be translated into restrictions on the parameters of frequency and severity distributions. Once these restrictions have been identified, a calibration strategy can be designed where parameters are calibrated by maximizing some standard criterion subject to these constraints. As a result, parameter estimators can be seen as a mixture of the internal data-based estimator and the scenario-based implied estimator.

5.3.4.1 Probability distribution of a given scenario

We assume that the number of losses $N(t)$ is a Poisson process with intensity λ. Let τ_n be the arrival time of the n^{th} loss:

$$\tau_n = \inf\{t \geq 0 : N(t) = n\}$$

We know that the durations $T_n = \tau_n - \tau_{n-1}$ between two consecutive losses are *iid* exponential random variables with parameter λ. We recall that the losses X_n are also *iid* with distribution $\mathbf{F}$. We note now $T_n(x)$ the duration between two losses exceeding x. It is obvious that the durations are *iid*. It suffice now to characterize $T_1(x)$. By using the fact that a finite sum of exponential times is an Erlang distribution, we have:

$$\begin{aligned}
\Pr\{T_1(x) > t\} &= \sum_{n\geq 1} \Pr\{\tau_n > t; X_1 < x, \ldots, X_{n-1} < x; X_n \geq x\} \\
&= \sum_{n\geq 1} \Pr\{\tau_n > t\} \cdot \mathbf{F}(x)^{n-1} \cdot (1 - \mathbf{F}(x)) \\
&= \sum_{n\geq 1} \mathbf{F}(x)^{n-1} \cdot (1 - \mathbf{F}(x)) \cdot \left(\sum_{k=0}^{n-1} e^{-\lambda t} \frac{(\lambda t)^k}{k!}\right) \\
&= (1 - \mathbf{F}(x)) \cdot \sum_{k=0}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \left(\sum_{n=k}^{\infty} \mathbf{F}(x)^n\right) \\
&= e^{-\lambda t} \sum_{k=0}^{\infty} \frac{(\lambda t)^k}{k!} \mathbf{F}(x)^k \\
&= e^{-\lambda(1-\mathbf{F}(x))t}
\end{aligned}$$

We deduce that $T_n(x)$ follows an exponential distribution with parameter $\lambda(x) = \lambda(1 - \mathbf{F}(x))$. The average duration between two losses exceeding x is also the mean of $T_n(x)$:

$$\mathbb{E}[T_n(x)] = \frac{1}{\lambda(1 - \mathbf{F}(x))}$$

Example 58 *We assume that the annual number of losses follows a Poisson distribution where* $\lambda = 5$ *and the severity of losses are log-normal* $\mathcal{LN}(9,4)$.

In Figure 5.16, we simulate the corresponding Poisson process $N(t)$ and also the events whose loss is larger than \$20 000 and \$50 000. We then show the exponential distribution[17] of $T_n(x)$.

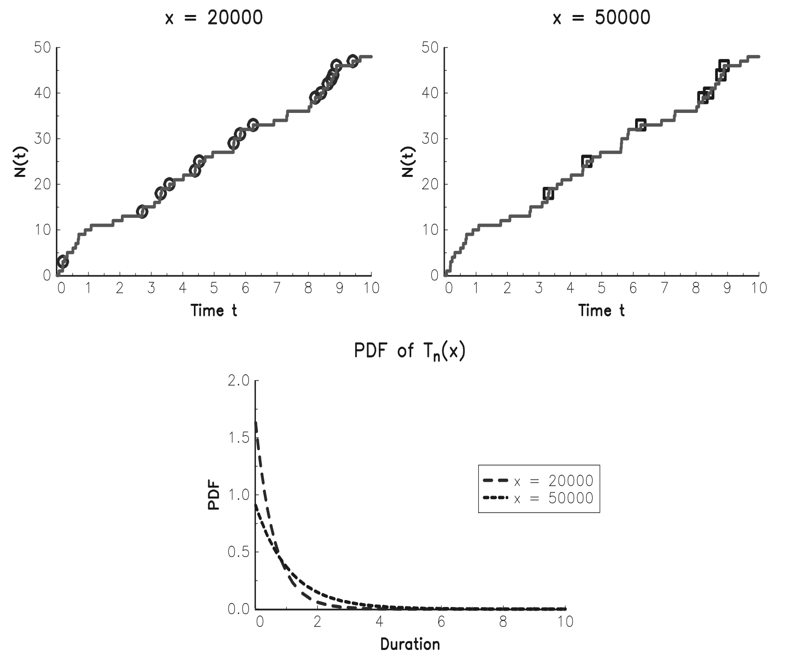

FIGURE 5.16: Simulation of the Poisson process $N(t)$ and peak over threshold events

5.3.4.2 Calibration of a set of scenarios

Let us consider a scenario defined as "a loss of x or higher occurs once every d years". By assuming a compound Poisson distribution with a parametric severity distribution $\mathbf{F}(x;\theta)$, λ is the average number of losses per year, $\lambda(x) = \lambda(1 - \mathbf{F}(x;\theta))$ is the average number of losses higher than x and $1/\lambda(x)$ is the average duration between two losses exceeding x. As a result, for a given scenario (x, d), parameters (λ, θ) must satisfy:

$$d = \frac{1}{\lambda(1 - \mathbf{F}(x;\theta))}$$

Suppose that we face different scenarios $\{(x_s, d_s), s = 1, \ldots, n_S\}$. We may estimate the implied parameters underlying the expert judgements using the quadratic criterion:

$$\left(\hat{\lambda}, \hat{\theta}\right) = \arg\min \sum_{s=1}^{n_S} w_s \cdot \left(d_s - \frac{1}{\lambda(1 - \mathbf{F}(x_s;\theta))}\right)^2$$

[17]For the parameter $\lambda(x)$, we have:

$$\lambda\left(2 \times 10^4\right) = 5 \times \left(1 - \Phi\left(\frac{\ln\left(2 \times 10^4\right) - 9}{2}\right)\right) = 1.629$$

and $\lambda\left(5 \times 10^4\right) = 0.907$.

where w_s is the weight of the s^{th} scenario. The previous approach belongs to the method of moments. As a result, we can show that the optimal weights w_s correspond to the inverse of the variance of d_s:

$$\begin{aligned} w_s &= \frac{1}{\operatorname{var}(d_s)} \\ &= \lambda\left(1 - \mathbf{F}(x_s;\theta)\right) \end{aligned}$$

To solve the previous optimization program, we proceed by iterations. Let $\left(\hat{\lambda}_m, \hat{\theta}_m\right)$ be the solution of the minimization program:

$$\left(\hat{\lambda}_m, \hat{\theta}_m\right) = \arg\min \sum_{j=1}^{p} \hat{\lambda}_{m-1} \cdot \left(1 - \mathbf{F}\left(x_s; \hat{\theta}_{m-1}\right)\right) \cdot \left(d_s - \frac{1}{\lambda\left(1 - \mathbf{F}(x_s;\theta)\right)}\right)^2$$

Under some conditions, the estimator $\left(\hat{\lambda}_m, \hat{\theta}_m\right)$ converge to the optimal solution. We also notice that we can simplify the optimization program by using the following approximation:

$$w_s = \frac{1}{\operatorname{var}(d_s)} = \frac{1}{\mathbb{E}[d_s]} \simeq \frac{1}{d_s}$$

Example 59 *We assume that the severity distribution is log-normal and consider the following set of expert's scenarios:*

x_s *(in \$ mn)*	1	2.5	5	7.5	10	20
d_s *(in years)*	1/4	1	3	6	10	40

If $w_s = 1$, we obtain $\hat{\lambda} = 43.400$, $\hat{\mu} = 11.389$ and $\hat{\sigma} = 1.668$ (#1). Using the approximation $w_s \simeq 1/d_s$, the estimates become $\hat{\lambda} = 154.988$, $\hat{\mu} = 10.141$ and $\hat{\sigma} = 1.855$ (#2). Finally, the optimal estimates are $\hat{\lambda} = 148.756$, $\hat{\mu} = 10.181$ and $\hat{\sigma} = 1.849$ (#3). In the table below, we report the estimated values of the duration. We notice that they are close to the expert's scenarios.

x_s (in \$ mn)	1	2.5	5	7.5	10	20
#1	0.316	1.022	2.964	5.941	10.054	39.997
#2	0.271	0.968	2.939	5.973	10.149	39.943
#3	0.272	0.970	2.941	5.974	10.149	39.944

Remark 64 *We can combine internal loss data, expert's scenarios and external loss data*[18] *by maximizing the penalized likelihood:*

$$\begin{aligned} \hat{\theta} = \arg\max \quad & \varpi_{\text{internal}} \cdot \boldsymbol{\ell}(\theta) - \varpi_{\text{expert}} \cdot \sum_{s=1}^{n_S} w_s \left(d_s - \frac{1}{\lambda\left(1 - \mathbf{F}(x_s;\theta)\right)}\right)^2 - \\ & \varpi_{\text{external}} \cdot \sum_{s=1}^{n_S^\star} w_s^\star \left(d_s^\star - \frac{1}{\lambda\left(1 - \mathbf{F}(x_s^\star;\theta)\right)}\right)^2 \end{aligned}$$

where ϖ_{internal}*,* ϖ_{expert} *and* ϖ_{external} *are the weights reflecting the confidence placed on internal loss data, expert's scenarios and external loss data.*

[18] In this case, each external loss is treated as an expert's scenario.

5.3.5 Stability issue of the LDA model

One of the big issues of AMA (and LDA) models is their stability. It is obvious that the occurrence of a large loss changes dramatically the estimated capital-at-risk as explained by Ames *et al.* (2015):

> "*Operational risk is fundamentally different from all other risks taken on by a bank. It is embedded in every activity and product of an institution, and in contrast to the conventional financial risks (e.g. market, credit) is harder to measure and model, and not straight forwardly eliminated through simple adjustments like selling off a position. While it varies considerably, operational risk tends to represent about 10-30% of the total risk pie, and has grown rapidly since the 2008-09 crisis. It tends to be more fat-tailed than other risks, and the data are poorer. As a result, models are fragile – small changes in the data have dramatic impacts on modeled output – and thus required operational risk capital is unstable*".

In this context, the Basel Committee has decided to review the different measurement approaches to calculate the operational risk capital. In Basel III, advanced measurement approaches have been dropped. This decision marks a serious setback for operational risk modeling. The LDA model continues to be used by Basel II jurisdictions, and will continue to be used by large international banks, because it is the only way to assess an economic capital using internal loss data. Moreover, internal losses continue to be collected by banks in order to implement the SMA of Basel III. Finally, the LDA model will certainly become the standard model for satisfying Pillar 2 requirements. However, solutions for stabilizing the LDA model can only be partial and even hazardous or counter-intuitive, because it ignores the nature of operational risk.

5.4 Exercises

5.4.1 Estimation of the loss severity distribution

We consider a sample of n individual losses $\{x_1, \ldots, x_n\}$. We assume that they can be described by different probability distributions:

(i) X follows a log-normal distribution $\mathcal{LN}(\mu, \sigma^2)$.

(ii) X follows a Pareto distribution $\mathcal{P}(\alpha, x_-)$ defined by:

$$\Pr\{X \leq x\} = 1 - \left(\frac{x}{x_-}\right)^{-\alpha}$$

where $x \geq x_-$ and $\alpha > 0$.

(iii) X follows a gamma distribution $\mathcal{G}(\alpha, \beta)$ defined by:

$$\Pr\{X \leq x\} = \int_0^x \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \, \mathrm{d}t$$

where $x \geq 0$, $\alpha > 0$ and $\beta > 0$.

(iv) The natural logarithm of the loss X follows a gamma distribution: $\ln X \sim \mathcal{G}(\alpha; \beta)$.

1. We consider the case *(i)*.

   (a) Show that the probability density function is:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{\ln x - \mu}{\sigma}\right)^2\right)$$

   (b) Calculate the first two moments of X. Deduce the orthogonal conditions of the generalized method of moments.

   (c) Find the maximum likelihood estimators $\hat{\mu}$ and $\hat{\sigma}$.

2. We consider the case *(ii)*.

   (a) Calculate the first two moments of X. Deduce the GMM conditions for estimating the parameter α.

   (b) Find the maximum likelihood estimator $\hat{\alpha}$.

3. We consider the case *(iii)*. Write the log-likelihood function associated to the sample of individual losses $\{x_1, \ldots, x_n\}$. Deduce the first-order conditions of the maximum likelihood estimators $\hat{\alpha}$ and $\hat{\beta}$.

4. We consider the case *(iv)*. Show that the probability density function of X is:

$$f(x) = \frac{\beta^\alpha (\ln x)^{\alpha-1}}{\Gamma(\alpha)\, x^{\beta+1}}$$

   What is the support of this probability density function? Write the log-likelihood function associated to the sample of individual losses $\{x_1, \ldots, x_n\}$.

5. We now assume that the losses $\{x_1, \ldots, x_n\}$ have been collected beyond a threshold H meaning that $X \geq H$.

   (a) What does the generalized method of moments become in the case *(i)*?

   (b) Calculate the maximum likelihood estimator $\hat{\alpha}$ in the case *(ii)*.

   (c) Write the log-likelihood function in the case *(iii)*.

5.4.2 Estimation of the loss frequency distribution

We consider a dataset of individual losses $\{x_1, \ldots, x_n\}$ corresponding to a sample of T annual loss numbers $\{N_{Y_1}, \ldots, N_{Y_T}\}$. This implies that:

$$\sum_{t=1}^{T} N_{Y_t} = n$$

If we measure the number of losses per quarter $\{N_{Q_1}, \ldots, N_{Q_{4T}}\}$, the previous equation becomes:

$$\sum_{t=1}^{4T} N_{Q_t} = n$$

1. We assume that the annual number of losses follows a Poisson distribution $\mathcal{P}(\lambda_Y)$. Calculate the maximum likelihood estimator $\hat{\lambda}_Y$ associated to the sample $\{N_{Y_1}, \ldots, N_{Y_T}\}$.

2. We assume that the quarterly number of losses follows a Poisson distribution $\mathcal{P}(\lambda_Q)$. Calculate the maximum likelihood estimator $\hat{\lambda}_Q$ associated to the sample $\{N_{Q_1}, \ldots, N_{Q_{4T}}\}$.

3. What is the impact of considering a quarterly or annual basis on the computation of the capital charge?

4. What does this result become if we consider a method of moments based on the first moment?

5. Same question if we consider a method of moments based on the second moment.

5.4.3 Using the method of moments in operational risk models

1. Let $N(t)$ be the number of losses for the time interval $[0, t]$. We note $\{N_1, \ldots, N_T\}$ a sample of $N(t)$ and we assume that $N(t)$ follows a Poisson distribution $\mathcal{P}(\lambda)$. We recall that:
$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$
   (a) Calculate the first moment $\mathbb{E}[N(t)]$.
   (b) Show the following result:
$$\mathbb{E}\left[\prod_{i=0}^{m} (N(t) - i)\right] = \lambda^{m+1}$$
   Then deduce the variance of $N(t)$.
   (c) Propose two estimators based on the method of moments.

2. Let S be the random sum:
$$S = \sum_{i=0}^{N(t)} X_i$$
   where $X_i \sim \mathcal{LN}(\mu, \sigma^2)$, $X_i \perp X_j$ and $N(t) \sim \mathcal{P}(\lambda)$.
   (a) Calculate the mathematical expectation $\mathbb{E}[S]$.
   (b) We recall that:
$$\left(\sum_{i=1}^{n} x_i\right)^2 = \sum_{i=1}^{n} x_i^2 + \sum_{i \neq j} x_i x_j$$
   Show that:
$$\operatorname{var}(S) = \lambda \exp\left(2\mu + 2\sigma^2\right)$$
   (c) How can we estimate μ and σ if we have already calibrated λ?

3. We assume that the annual number of losses follows a Poisson distribution $\mathcal{P}(\lambda)$. We also assume that the individual losses are independent and follow a Pareto distribution $\mathcal{P}(\alpha, x_-)$ defined by:
$$\Pr\{X \leq x\} = 1 - \left(\frac{x}{x_-}\right)^{-\alpha}$$
   where $x \geq x_-$ and $\alpha > 1$.

(a) Show that the duration between two consecutive losses that are larger than ℓ is an exponential distribution with parameter $\lambda x_{-}^{\alpha} \ell^{-\alpha}$.

(b) How can we use this result to calibrate experts' scenarios?

5.4.4 Calculation of the Basel II required capital

We consider the simplified balance sheet of a bank, which is described below.

1. In the Excel file, we provide the price evolution of stocks A and B. The trading portfolio consists of 10 000 shares A and 25 000 shares B. Calculate the daily historical VaR of this portfolio by assuming that the current stock prices are equal to \$105.5 and \$353. Deduce the capital charge for market risk assuming that the VaR has not fundamentally changed during the last 3 months[19].

2. We consider that the credit portfolio of the bank can be summarized by 4 meta-credits whose characteristics are the following:

	Sales	EAD	PD	LGD	M
Bank		\$80 mn	1%	75%	1.0
Corporate	\$500 mn	\$200 mn	5%	60%	2.0
SME	\$30 mn	\$50 mn	2%	40%	4.5
Mortgage		\$50 mn	9%	45%	
Retail		\$100 mn	4%	85%	

   Calculate the IRB capital charge for the credit risk.

3. We assume that the bank is exposed to a single operational risk. The severity distribution is a log-normal probability distribution $\mathcal{LN}(8,4)$, whereas the frequency distribution is the following discrete probability distribution:

$$\begin{aligned} \Pr\{N=5\} &= 60\% \\ \Pr\{N=10\} &= 40\% \end{aligned}$$

   Calculate the AMA capital charge for the operational risk.

4. Deduce the capital charge of the bank and the capital ratio knowing that the capital of the bank is equal to \$70 mn.

5.4.5 Parametric estimation of the loss severity distribution

1. We assume that the severity losses are log-logistic distributed $X_i \sim \mathcal{LL}(\alpha,\beta)$ where:

$$\mathbf{F}(x;\alpha,\beta) = \frac{(x/\alpha)^{\beta}}{1+(x/\alpha)^{\beta}}$$

   (a) Find the density function.

   (b) Deduce the log-likelihood function of the sample $\{x_1,\ldots,x_n\}$.

   (c) Show that the ML estimators satisfy the following first-order conditions:

$$\left\{ \begin{array}{l} \sum_{i=1}^{n} \mathbf{F}\left(x_i;\hat{\alpha},\hat{\beta}\right) = n/2 \\ \sum_{i=1}^{n} \left(2\mathbf{F}\left(x_i;\hat{\alpha},\hat{\beta}\right)-1\right)\ln x_i = n/\hat{\beta} \end{array} \right.$$

[19] The multiplication coefficient ξ is set equal to 0.5.

(d) The sample of loss data is 2 918, 740, 3 985, 2 827, 2 839, 6 897, 7 665, 3 766, 3 107 and 3 304. Verify that $\hat{\alpha} = 3\,430.050$ and $\hat{\beta} = 3.315$ are the ML estimates.

(e) What does the log-likelihood function of the sample $\{x_1, \ldots, x_n\}$ become if we assume that the losses were collected beyond a threshold H?

5.4.6 Mixed Poisson process

1. We consider the mixed poisson process where $N(t) \sim \mathcal{P}(\Lambda)$ and Λ is a random variable. Show that:

$$\operatorname{var}(N(t)) = \mathbb{E}[N(t)] + \operatorname{var}(\Lambda)$$

2. Deduce that $\operatorname{var}(N(t)) \geq \mathbb{E}[N(t)]$. Determine the probability distribution Λ such that the equality holds. Let $\varphi(n)$ be the following ratio:

$$\varphi(n) = \frac{(n+1) \cdot p(n+1)}{p(n)}$$

Show that $\varphi(n)$ is constant.

3. We assume that $\Lambda \sim \mathcal{G}(\alpha, \beta)$.

(a) Calculate $\mathbb{E}[N(t)]$ and $\operatorname{var}(N(t))$.

(b) Show that $N(t)$ has a negative binomial distribution $\mathcal{NB}(r, p)$. Calculate the parameters r and p with respect to α and β.

(c) Show that $\varphi(n)$ is an affine function.

4. We assume that $\Lambda \sim \mathcal{E}(\lambda)$.

(a) Calculate $\mathbb{E}[N(t)]$ and $\operatorname{var}(N(t))$.

(b) Show that $N(t)$ has a geometric distribution $\mathcal{G}(p)$. Determine the parameter p.