Quantification of Operational Risks using a Scenario Based Approach

Université Paris Dauphine

Jeanne FASSIER
10/10/2014
ACKNOWLEDGEMENTS

I would like to take this opportunity to thank all the people I met during my internship, especially those who have contributed in some way to this thesis.

Particularly, my warm thanks go to Jean-Baptiste PETIT who supervised and mentored my work, for his attention, and for all the time he awarded to me, his advice was really useful and appreciated. I am very happy if my modest contributions have helped him get his head above water sometimes.

I also thank all the members of the Operational Risk Team of AXA Group Risk Management – Céline SAMAIN, Mai SAYACHAK, and Olivier CHANSIN - who welcomed me within the team and were always ready to help me. In particular, a warm thank to Céline for trusting me and considering me as an actual member of the team; to Mai for enabling me to exercise on the popularization of some pompous mathematical concepts; to Olivier for challenging my results and teaching me how to use a book binder at the last minute. In addition, I would like to extent a warm thank to Dora ELAMRI and Fabien CHERANCE for their help during my researches.

Then, I would like to thank all the members of the GRM for their patience and the time they spent to explain their work to me. I have definitely “gotten lucky” by learning in this friendly atmosphere gathering very nice, dynamic and smart people.

Finally, I would like to thank Dylan POSSAMAI who gave me all his attention and whose really valued and appreciated advice helped me realizing that thesis.
ABSTRACT

Operational risk is the risk of losses resulting from inadequate or failed internal processes, people and systems, or from external events. Solvency II has defined regulatory requirements for the computation of the operational risk economic capital. Given the heterogeneity and scarcity of the data, modelling operational risk within an insurance company is a challenging task. In this context, an alternative is to use the Scenario Based Approach (SBA): this method focuses on scenario analyses that are fed with expert opinion. The quantification of those risk scenarios on a standalone basis leads to the study of their dependence in order to aggregate those operational risks and finally determine the annual aggregate loss distribution to compute the Value-at-Risk at 99.5%. The objective of this thesis is to scrutinize the steps of the quantification process in a SBA internal model in order to discourse about the robustness of the model. The study starts at the standalone level with the quantification of operational risks using a frequency-severity approach and then gradually reaches a more general point of view with the aggregated losses: it intends to take a step backwards in order to gain a better overall understanding of the model. With the objective of injecting a new perspective into a SBA model, alternatives that may enable to circumvent some of the hurdles encountered during the review are presented throughout the development.

Keywords:
Operational risk, risk management, SCR, Solvency II, internal model, frequency-severity, Scenario Based Approach, SBA, calibration, nearest correlation matrix, aggregation, copula, Monte Carlo, Value-at-Risk, sensitivity analysis, scaling
Le risque opérationnel est le risque de pertes dues à une défaillance des procédures internes, des personnes, des systèmes internes ou à des risques externes. La directive Solvabilité II inclut le risque opérationnel dans le calcul du besoin en fonds propres. Ce large spectre conduit à des bases de données hétérogènes et peu abondantes. Une des alternatives au manque de données a été introduite dans les réglementations bâloises sous le nom de *Scenario Based Approach* (SBA) : cette méthode repose sur l’analyse de scénarios de risques, alimentés par les dires d’experts. La quantification de chacun de ces scénarios de risques individuels mène à l’étude de leurs dépendances afin de déterminer une méthodologie d’agrégation des risques opérationnels, pour finalement établir la distribution des pertes opérationnelles annuelles et calculer la Value-at-Risk à 99.5%. L’objectif de ce mémoire est d’examiner les différentes étapes du processus de quantification d’un modèle interne SBA afin de discuter de la robustesse du modèle. L’étude débute par la quantification individuelle des scénarios par une approche fréquence-sévérité pour se poursuivre progressivement dans le cadre plus général de la vision agrégée des risques : le but est de prendre un certain recul sur la méthode afin d’avoir une compréhension plus globale du modèle. Toujours dans le but d’apporter une nouvelle perspective à un modèle SBA, ce mémoire s’attache à présenter, tout au long de la discussion, des alternatives permettant de contourner d’éventuelles difficultés rencontrées lors du passage en revue du modèle.

*Mots-clés :*

Risque opérationnel, gestion des risques, SCR, Solvabilité II, modèle interne, Scenario Based Approach, SBA, fréquence-sévérité, calibrage, matrice de corrélation, agrégation, copule, Monte Carlo, Value-at-Risk, tests de sensibilité, scaling
EXECUTIVE SUMMARY

Lately, huge operational losses have hit the headlines and raised awareness on the need for banking and insurance companies to be able to manage those risks. In the spirit of Basel II, the European directive Solvency II has defined regulatory requirements for the computation of an operational risk economic capital. The topic of quantification of operational risk has gained increasing attention in both academic research and in risk management teams. However, this discipline is relatively new for the insurance industry, and the scope is wide and constantly evolving. Consequently, the loss databases are not sufficiently dense to put in place a classical internal model, based on the historical losses. Given the heterogeneity and scarcity of the data, modelling operational risk within an insurance company is a challenging task. In this context, an alternative to the lack of available data is to use the Scenario Based Approach (SBA) that was introduced in Basel II regulation: this method focuses on scenario analyses that are fed with expert opinion. The objective of this thesis is to scrutinize the steps of the quantification process in a SBA model in order to discourse about the robustness of the model. The study starts at the standalone level with the quantification of operational risks using a frequency-severity approach and then gradually reaches a more general point of view with the aggregated losses: it intends to take a step backwards in order to gain a better overall understanding of the model. With the objective of injecting a new perspective into a SBA model, alternatives that may enable to circumvent some of the hurdles encountered during the review are presented throughout the development.

This thesis starts by introducing the concept of operational risk to the reader and illustrates it with some well-known examples. Then, chapter 2 presents how a risk management team deals with operational risk. Insurance companies have different methodologies for quantifying their operational risks and computing the associated economic capital, aligned with Solvency II requirements. We will see that the firm can use a Standard Formula or build an internal model. In this case, we will see that insurance companies have recourse to techniques initiated by the banking industry. An overview of the industry practices is presented in chapter 3. Chapter 4 focuses on the methodology of the SBA that will be of interest in this thesis. It consists of identifying, assessing and measuring the most critical Operational risks. The approach, which relies on expert judgment, is therefore forward-looking and adequately reflects the risk profile of the company. The implementation of a strong operational risk framework rests upon figures that help the management evaluate the exposure of the company. Because the accuracy of the quantification process is a major concern, comprehensiveness of the model will be discussed throughout this thesis.

Chapter 5 introduces the standalone modelling of Operational risks by a frequency-severity approach. The specificity of this modelling comes from the fact that the distributions are fitted on assessed points. As classical parameter estimation techniques are not consistent with such approach, the calibration exercise is very specific to the SBA model: it consists of finding the solution of a system of equations that gives closed formulae for the parameters. Nevertheless, the solution of a system composed of more than two equations with only two variables is not ensured. The removal of one of the calibration points is not desirable which motivates the introduction of a weighting factor between the calibration points. Instead of being fixed to a default value, the weighting factor can be considered as a third variable in the calibration exercise: the one that minimizes the model
discrepancies. With numerical methods, finding the optimal weighting factor is very feasible. However, the outcomes highly depend on the choice of the error function to be minimized. This uncertainty indicates a lack of robustness of this method and suggests being cautious when choosing one of the proposed minimization methods.

The insurance company has to evaluate the amount of capital it should hold at aggregated level. Of course, the simple sum of all the risks is not conceivable. Indeed, that all operational risks occur simultaneously is intuitively not likely to happen. For instance, it is expected that an internal fraud occurring in Japan would not be correlated with a fire incident in Spain. Therefore, insurance companies have to take into account the dependence that exists between operational risks and introduce this into their internal model. Chapter 6 gives an overview on dependence measures that are used in operational risk. It explains why the usual linear correlation coefficient is not adapted to our problem. This section introduces non-parametric methods that will be of interest in the context of the improving the robustness of the model. Moreover, supervisors want insurance companies to justify their underlying assumptions, in order to ensure the full adequacy of their internal model. Thus, this chapter describes how the underlying model assumptions of independence can be back-tested in a SBA model, and discusses the hypothesis of independence between frequency and severity necessary to use an aggregate loss model. Indeed, the annual loss of a given event type is driven by two sources of randomness: frequency and severity. Therefore the correlation between two risks could be the result of correlation between frequencies, severities, or both. Usually, measurements of frequency correlations are used to derive models for correlations at the aggregate loss level, while assuming independence of loss severities. For the sake of simplicity, the methodology for modelling dependence between operational risks presented hereafter is based on the introduction of dependence at the aggregate loss level.

The dependence structures are summarized into a correlation matrix. However, the expert-judgement matrix does not always have the required properties whereas usual aggregation techniques require the input matrix to be semi-definite positive in order to perform the Cholesky algorithm. Chapter 7 suggests a method that slightly modifies the input matrix in order to build a correlation matrix that is very close to the initial one. Then, this chapter presents some usual aggregation techniques. The first one is the variance-covariance method. The second method is based on the theory of copulas. The variance-covariance method should be dismissed straightaway because operational risks are not modelled with the Gaussian distribution. Copula methods should definitely be preferred. The choice of the copula will be driven by required properties for the dependence structure, evaluated by the expert. The aggregation presented in this chapter is inspired by the copula theory: the ranks of the vectors correlated with the desired copula (Gaussian or Student) are reproduced in the standalone risks so that they have the same rank correlations. Both copulas take a correlation matrix into parameter but the Student copula also has a degree of liberty that cannot be determined via expert judgment. Because of the uncertainty caused by the calibration of the Student copula in a SBA internal model, we conclude that the Gaussian assumption seems more reasonable.

The formula for the calculation of the Operational Risk economic capital raises a few questions that are discussed in chapter 8. It explains the notion of expected loss: interpreting this as the median seems more prudent because median remains null most of the time and is smaller than the mean. In
fact, when it comes to summarizing the core of the loss distribution with only one figure, the median should be preferred as it is less sensitive to outliers and therefore more robust than the mean. As expected, the robustness of the computation of the capital charge is highly dependent on the convergence of the Monte Carlo algorithm. Accuracy of the results and computational speed have to counterbalance each other in order to reach an acceptable compromise. The Monte Carlo study also highlights the capital reduction after proceeding to the aggregation which illustrates the diversification effects. The use of the Value-at-Risk to measure the risk borne by the insurer is also discussed in this chapter. As this topic has been widely treated in the literature we will give a brief overview on this topic in order to ensure that the reader is aware of the limits of the usage of Value-at-Risk. Although its limits have been largely pointed out, among which its absence of subadditivity, other papers have insisted on the robustness of the Value-at-Risk compared to the robustness of recently proposed alternatives. However, the existence of a threshold under which the median is null certainly exists for extreme quantiles as well and could seriously put in doubt the consistency of the use of the VaR.

Last chapter aims to take a step backwards in order to gain a better overall understanding of the model. In this respect, this chapter aims to test the robustness of the model by conducting sensitivity analyses. We will see that there exist some combinations of parameters for which the model is highly sensitive. In addition to the identification of those inputs, the study led to notice the importance of the design of the scenarios. Indeed, it is pretty intuitive that the accuracy of the design of the scenario itself highly contributes to the model robustness in a Scenario Based Approach. This chapter suggests some workaround to be performed in those specific cases in order to reduce the model uncertainty. Still with the objective of injecting a new perspective into a SBA model, this chapter suggests an alternative to the use of available data when feeding the scenario risk assessment. It consists of developing a scaling method that can help the expert benchmark its measurements by comparing them to actual losses that occurred in another entity of the Group. The method is based on the analysis that the size of the log-losses is highly correlated with the geographical region and size of the company where the loss occurred. The regression model aims to capture the characteristics of the occurred losses that can be explained by some exposure factors in order to adjust the losses to another entity. The results of the validation exercise that consists of comparing scaled losses to losses that were genuinely suffered by the entity will be provided at the end of the chapter, in order to verify the soundness of the scaling model.

After all, the existence of some flaws in the models is inherent to their nature. Indeed, a model is never perfect, or totally representative of the reality. In Operational risk, the major stake is to succeed in finding a compromise between the necessary pragmatism of the approach and the level of abstraction of the mathematical model. However, lack of agreement between theoretical mathematical models and experimental measurements often leads to important advances as better theories are developed. This is the reason why insurance companies have to be aware of the imperfections of their model so that they can make effort in order to perpetually try to find ameliorations and alternatives that may improve their model.
NOTE DE SYNTHESE

Ces dernières années, de nombreuses pertes opérationnelles spectaculaires ont fait les gros titres, mettant en évidence le besoin certain pour les compagnies bancaires et assurantielles d’apprendre à gérer ce risque. La directive européenne Solvabilité II, inspirée de la règlementation bancaire, inclut à son tour le risque opérationnel dans le calcul du besoin en fonds propres. Les techniques de quantification du risque opérationnel se sont alors rapidement développées à la fois dans le monde académique et dans les équipes de gestion des risques. Puisque cette discipline est encore récente dans le monde de l’assurance, alors même que son champ d’application est particulièrement large et en perpétuelle évolution, les données dont disposent les assureurs sont encore trop rares et hétérogènes pour mesurer le risque opérationnel dans le cadre d’un modèle interne basé sur ces pertes. Ainsi, une des alternatives au manque de données a été introduite dans les réglementations bâloises sous le nom de Scenário Based Approach (SBA) : cette méthode repose sur l’analyse de scénarios de risques, alimentés par les dires d’experts. L’objectif de ce mémoire est d’examiner les différentes étapes du processus de quantification d’un modèle SBA afin de discuter de la robustesse du modèle. L’étude débute par la quantification individuelle des scénarios par une approche fréquence-sévérité pour se poursuivre ensuite dans le cadre plus général de la vision agrégée des risques : le but est de prendre un certain recul sur la méthode afin d’avoir une compréhension plus globale du modèle. Toujours dans le but d’apporter une nouvelle perspective à un modèle SBA, ce mémoire s’attache à présenter, tout au long de la discussion, des alternatives permettant de contourner d’éventuelles difficultés rencontrées lors du passage en revue du modèle.

Ce mémoire commence par introduire la notion de risque opérationnel au lecteur, et illustre son propos en présentant de célèbres pertes opérationnelles. Le chapitre suivant présente la manière dont ce risque est géré au sein d’une équipe de Risk Management. Les compagnies d’assurances utilisent des méthodologies différentes pour quantifier le risque opérationnel et calculer le capital économique qui lui est associé. Nous verrons qu’elles peuvent utiliser une Formule Standard, proposée par la directive Solvabilité II, mais aussi construire leur propre modèle interne. Dans ce cas, nous verrons que la plupart des compagnies d’assurances ont recours à des techniques précédemment développées par l’industrie bancaire. Le chapitre 3 donnera une vue d’ensemble des pratiques de marché. Ensuite, le chapitre 4 s’attachera à détailler la méthodologie de l’approche SBA, qui nous intéresse dans ce mémoire. Elle consiste à identifier, évaluer et quantifier les plus gros risques opérationnels auxquels la compagnie est exposée. Cette approche, qui repose largement sur les dires d’experts, est donc tournée vers l’avenir afin de refléter au mieux le profil de risque de l’assureur. La qualité d’un tel dispositif repose sur les chiffres fournis aux équipes de direction afin d’évaluer l’exposition de l’entreprise. L’exactitude de ces données quantitatives étant un enjeu stratégique majeur, la cohérence du modèle sera discutée tout au long de cette étude.

L’approche fréquence-sévérité utilisée pour modéliser les risques opérationnels individuellement est introduite dans le chapitre 5. La spécificité de ce modèle tient du fait que les distributions sont calibrées à partir de points mesurés par l’expert. Comme les méthodes classiques d’estimations des paramètres ne sont pas envisageables dans ce cas, l’exercice de calibrage est très spécifique à l’approche SBA : il s’agit de résoudre des systèmes d’équations afin de déterminer une formule fermée qui donne les paramètres des distributions en fonction des points de calibrage. Toutefois, la
résolution d’un système de plus de deux équations à deux inconnues n’est pas toujours assurée. Il n’est pas pour autant préférable de supprimer ne serait-ce qu’un de ces points de calibrage, puisque cela induirait une perte d’information. Cela amène donc à l’introduction d’un facteur de pondération, entre plusieurs des points de calibrage. Plutôt que de voir attribuer une valeur par défaut, ce facteur de pondération peut être considéré comme une troisième variable dans cet exercice de calibrage : il s’agit alors de trouver le facteur minimisant l’erreur de modèle. La résolution numérique du problème rend l’implémentation de cette méthode très réalisable. Toutefois, nous verrons que les résultats dépendent fortement du choix de la fonction à minimiser. Cette incertitude indique un manque de robustesse de cette méthode et suggère ainsi de rester prudent quant au choix et à l’implémentation de ce problème d’optimisation.

L’assureur doit alors évaluer le niveau de capital qu’il doit détenir au niveau agrégé. Bien sûr, additionner tous les risques est inconcevable. En effet, la survenance simultanée de tous les risques opérationnels est intuitivement peu probable. Par exemple, on s’attend à ce qu’un cas de fraude interne survenu au Japon soit totalement indépendant de la déclaration d’un incendie au sein d’une entité en Espagne. Ainsi, les compagnies doivent prendre en compte les dépendances existant entre les risques opérationnels afin d’introduire cette notion dans leur modèle interne. Le chapitre 6 présente une vue d’ensemble des mesures de dépendance utilisées en risque opérationnel. Il explique pourquoi le coefficient de corrélation linéaire n’est pas adapté à notre problème et introduit des méthodes non-paramétriques qui nous intéresseront dans un contexte d’éventuelle amélioration de la robustesse du modèle. De plus, les autorités exigent des compagnies de justifier leurs hypothèses afin de s’assurer de la justesse de leur modèle. Ainsi, ce chapitre décrit comment les hypothèses d’indépendance peuvent être vérifiées dans un modèle SBA, alors même que l’indépendance entre fréquence et sévérité est nécessaire à la mise en place du modèle collectif. En effet, il y a deux sources d’aléa dans la perte annuelle d’un scénario : la fréquence et la sévérité. Ainsi, la corrélation entre les risques peut être le fruit d’une de ces sources, ou bien de deux en même temps. Classiquement, certains modèles évaluent la corrélation entre les variables de fréquence, en supposant l’indépendance entre les variables de sévérité. Afin de simplifier le problème, la méthodologie présentée dans cette étude introduit la notion de dépendance au niveau des pertes agrégées.

Les structures de dépendance sont alors résumées par une matrice de corrélation. Cependant, une matrice construite à dires d’experts ne respecte pas toujours les propriétés requises pour une matrice de corrélation alors que les techniques classiques d’agrégation prennent en paramètre une matrice semi-définie positive afin de procéder à l’algorithme de Cholesky. Le chapitre 7 propose des méthodes permettant de modifier légèrement les coefficients de la matrice pour la rendre semi-définie positive. Les principales techniques d’agrégation sont présentées à la suite. La première est la méthode de variance-covariance. La seconde est basée sur la théorie des copules. Nous verrons que la méthode de variance-covariance est éliminée d’emblée parce que les risques opérationnels ne sont pas des vecteurs gaussiens. Les méthodes inspirées des copules sont donc à privilégier. Le choix de la copule sera alimenté par des propriétés voulues pour la structure de dépendance, choisis par l’expert. L’agrégation présentée dans ce chapitre s’inspire donc de la théorie des copules : les rangs des vecteurs corrélés avec la copule choisie (Normale ou Student) sont reproduits pour les risques individuels afin qu’ils aient les mêmes corrélations de rang. Ces deux copules ont pour paramètre une matrice de corrélation mais la copule de Student dépend aussi du nombre de degrés de liberté, qui ne peut pas être évalué par l’expert, faute d’interprétation pratique de ce paramètre. A cause de
l’incertitude liée au calibrage de la copule de Student dans le cadre d’une modèle SBA, on préfèrera alors utiliser la copule gaussienne.

La formule qui donne le capital économique relatif au risque opérationnel soulève certaines questions qui sont discutées dans le chapitre 8. Il explique notamment la notion d’*expected loss* : l’interpréter comme la médiane de la distribution de perte semble plus prudent parce que la médiane reste nulle dans la plupart des cas et reste inférieure à la moyenne. En général, nous verrons que la médiane devrait être préférée à la moyenne pour résumer la distribution en un seul chiffre, puisqu’elle est moins sensible aux valeurs aberrantes et donc plus robuste que la moyenne. De plus, la robustesse du calcul de la charge en capital est très dépendante de la convergence des simulations de Monte Carlo. Précision des résultats et rapidité d’exécution doivent se compenser afin de parvenir à un juste compromis. L’étude sur Monte Carlo permet aussi de souligner la réduction en capital due au processus d’agrégation et donc d’illustrer les effets de la diversification. Finalement, ce chapitre discute de l’utilisation de la Value-at-Risk pour mesurer le risque porté par la compagnie. Ce mémoire donne une brève présentation du sujet afin de s’assurer que le lecteur connait les limites de l’utilisation de cette mesure de risque. Bien que ces limites aient été largement montrées du doigt, parmi lesquelles l’absence de sous-additivité, certaines recherches ont insisté sur la robustesse de la Value-at-Risk comparée à celle des alternatives récemment proposées. En revanche, on montrera l’existence d’un seuil en dessous duquel la médiane est nulle. Cette propriété étant aussi vraie pour les quantiles extrêmes, on a là un argument mettant sérieusement en doute le choix de la Value-at-Risk pour mesurer le risque.

Le dernier chapitre a pour objectif de prendre un certain recul sur ce qui a été présenté afin de mieux comprendre les subtilités du modèle. Pour cela, des tests de sensibilité sont conduits dans ce chapitre, afin de tester la robustesse du modèle. Nous verrons qu’il existe des combinaisons de paramètres pour lesquelles le modèle est très sensible. En plus d’identifier ces paramètres, l’étude mène à souligner l’importance de la construction du scénario. En effet, il semble assez naturel que la validité d’un modèle SBA soit très liée à la justesse des scénarios qui le composent. Ce chapitre suggère alors certaines manipulations qui peuvent mener, dans certains cas, à réduire l’incertitude du modèle. Toujours dans le but d’apporter une nouvelle perspective à un modèle SBA, ce chapitre propose une alternative à l’utilisation de données disponibles pour nourrir l’évaluation des scénarios. Cela consiste à développer une méthode de *scaling* pour aider l’expert à comparer les résultats de son évaluation à des pertes réelles, survenues dans d’autres entités du Groupe. La méthode se base sur le constat que les montants des pertes sont très corrélés avec la région géographique et la taille de l’entité dans laquelle elle a eu lieu. Le modèle de régression capte des caractéristiques des pertes observées en fonction de variables explicatives bien choisies afin de transposer ces pertes à d’autres entités. Les résultats de l’exercice de validation consistant à comparer les pertes ajustées à celles que l’entité a réellement subies seront présentés à la fin du chapitre afin de vérifier la justesse du modèle proposé.

Finalement, l’existence de défauts dans le modèle est naturelle. En effet, un modèle n’est jamais parfait et ne reflète pas complètement la réalité. En risque opérationnel, l’enjeu majeur est de trouver un compromis entre le caractère nécessairement pragmatique de l’approche et le niveau d’abstraction du modèle mathématique. Par ailleurs, c’est justement après la découverte d’incohérences entre la théorie mathématique et les observations empiriques que l’on observe les plus importantes avancées dans la recherche. C’est la raison pour laquelle les compagnies
d’assurances se doivent de connaître les imperfections de leurs modèles, afin de continuer à faire les efforts nécessaires pour perpétuellement chercher des alternatives permettant d’améliorer leur modèle.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>1</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>2</td>
</tr>
<tr>
<td>RESUME</td>
<td>3</td>
</tr>
<tr>
<td>EXECUTIVE SUMMARY</td>
<td>4</td>
</tr>
<tr>
<td>NOTE DE SYNTHESE</td>
<td>7</td>
</tr>
<tr>
<td>TABLE OF CONTENTS</td>
<td>11</td>
</tr>
<tr>
<td>CHAPTER 1</td>
<td>14</td>
</tr>
<tr>
<td>INTRODUCTION</td>
<td>14</td>
</tr>
<tr>
<td>1 What is Operational Risk?</td>
<td>14</td>
</tr>
<tr>
<td>1.1 Definition</td>
<td>14</td>
</tr>
<tr>
<td>1.2 Well known examples</td>
<td>14</td>
</tr>
<tr>
<td>2 Basel Committee</td>
<td>15</td>
</tr>
<tr>
<td>3 Solvency II</td>
<td>16</td>
</tr>
<tr>
<td>3.1 A three pillars approach</td>
<td>16</td>
</tr>
<tr>
<td>3.2 The introduction of operational risk measurement in insurance</td>
<td>17</td>
</tr>
<tr>
<td>4 Objective of this thesis</td>
<td>18</td>
</tr>
<tr>
<td>CHAPTER 2</td>
<td>20</td>
</tr>
<tr>
<td>OPERATIONAL RISK FRAMEWORK</td>
<td>20</td>
</tr>
<tr>
<td>1 AXA Group</td>
<td>20</td>
</tr>
<tr>
<td>2 Group Risk Management</td>
<td>20</td>
</tr>
<tr>
<td>3 Group Operational Risk Team</td>
<td>21</td>
</tr>
<tr>
<td>4 Operational Risk Framework</td>
<td>21</td>
</tr>
<tr>
<td>4.1 AXA’s Definition of Operational Risk</td>
<td>21</td>
</tr>
<tr>
<td>4.2 Group risk grid</td>
<td>22</td>
</tr>
<tr>
<td>4.3 Boundaries with other risks</td>
<td>22</td>
</tr>
<tr>
<td>CHAPTER 3</td>
<td>24</td>
</tr>
<tr>
<td>INDUSTRY PRACTICE</td>
<td>24</td>
</tr>
<tr>
<td>1 Operational Risk Data</td>
<td>24</td>
</tr>
<tr>
<td>1.1 Basel II recommendations</td>
<td>24</td>
</tr>
<tr>
<td>1.2 The four data elements</td>
<td>25</td>
</tr>
<tr>
<td>2 Standard formula</td>
<td>27</td>
</tr>
<tr>
<td>3 Internal model</td>
<td>29</td>
</tr>
<tr>
<td>3.1 Loss Distribution Approach</td>
<td>29</td>
</tr>
<tr>
<td>3.2 Bayesian Approach</td>
<td>30</td>
</tr>
<tr>
<td>3.3 SBA</td>
<td>31</td>
</tr>
<tr>
<td>CHAPTER 4</td>
<td>32</td>
</tr>
<tr>
<td>THE SCENARIO BASED APPROACH</td>
<td>32</td>
</tr>
<tr>
<td>1 Operational risk SCR</td>
<td>32</td>
</tr>
<tr>
<td>2 Standards</td>
<td>33</td>
</tr>
<tr>
<td>2.1 Single methodology</td>
<td>33</td>
</tr>
<tr>
<td>2.2 Top-down vs. Bottom-up</td>
<td>34</td>
</tr>
<tr>
<td>3 Guidelines and Methodology</td>
<td>35</td>
</tr>
<tr>
<td>3.1 Identification: Risk assessment</td>
<td>35</td>
</tr>
<tr>
<td>3.2 Identification: loss data collection</td>
<td>37</td>
</tr>
</tbody>
</table>
3.3 Quantification 38
3.4 Monitoring & Reporting 40
4 Conclusion 41

CHAPTER 5 42

MODELLING STANDALONE SCENARIOS 42
1 Theory: From the risk assessment to the standalone quantification 42
2 Calibration 43
2.1 Frequency distribution 43
2.2 Severity distribution 45
3 Weighting factor 49
3.1 Discrepancies in the model 49
3.2 Minimization of the model error 51
4 Conclusion 55

CHAPTER 6 56

THE CORRELATION BETWEEN RISKS 56
1 Measure of dependence 56
1.1 Linear correlation 57
1.2 Rank correlation coefficients 58
1.3 Coefficient of tail dependence 63
2 Independence between frequency and severity 63
2.1 Pearson’s chi-squared test 64
2.2 Significance tests 65
2.3 Independence tests in practice 67
3 Back testing the iid hypothesis 70
3.1 Non-parametric tests 70
3.2 Back testing the iid hypothesis in practice 71
4 Conclusion 72

CHAPTER 7 73

AGGREGATION 73
1 Correlation matrix 73
1.1 Cholesky decomposition and positive definiteness 74
1.2 Nearest correlation matrix methods 77
2 Aggregation methods 82
2.1 Variance-covariance method 82
2.2 Gaussian copula 83
2.3 Alternatives to the Gaussian copula 87
3 Conclusion 89

CHAPTER 8 90

COMPUTATION OF THE CAPITAL CHARGE 90
1 Risk measure 90
1.1 Value-at-Risk 91
1.2 Coherent risk measure 91
1.3 Limitations of the VaR 92
2 Expected loss 92
2.1 Frequency threshold for the null median 93
2.2 Median is more prudent than the mean 95
3 Monte Carlo: Study of the convergence of the VaR 96
3.1 Standalone scenarios 96
CHAPTER 1

INTRODUCTION

1 What is Operational Risk?

1.1 Definition

According to Basel II regulations, Operational risk is "the risk of a change in value caused by the fact that actual losses, incurred for inadequate or failed internal processes, people and systems, or from external events (including legal risk), differ from the expected losses". This definition is similar to the one adopted by the European Union Solvency II Directive that applies to insurance companies.

Thus, they are risks that do not only exist in the banking or insurance industry. It can be, for instance, a ‘fat finger’ trade execution error, a case of internal fraud from a tied agent or an IT disruption that prevents the staff from working for a few hours. But it can also be a worldwide pandemic that prevent the staff from going to work. Depending on the virulence of the virus, significant employee absenteeism is expected due to illness, caring for sick family members and employees avoiding work for fear of becoming sick. Absenteeism would impact productivity and could indirectly trigger lapses due to deterioration of the quality of customer service. The abnormal rate of absenteeism for the workforce for several weeks could have an economic impact similar to recession. Thus, operational risk can be very expensive for an international company like AXA.

1.2 Well known examples

1.2.1 Barings (1995)

In 1995 Britain’s oldest merchant bank of two hundred years came to a dramatic and fatal halt. This bank was Barings. The bankruptcy was a result of the actions of its head derivative trader based in Singapore, Nick Leeson. The economic impact of the earthquake in Kobe turned the markets against him. Leeson requested extra funds to continue his trading activities and tried to extricate himself from the financial mess. He hid the losses in an obscure account that was discovered when his bosses carried out a spot audit after being alerted by those requests. The losses amounted to GBP800m, almost the entire assets of the bank at that time. Barings finally crashed and was bought for GDP1 by the banking and insurance company ING. Leeson pleaded guilty to fraud and wrote “Rogue Trader”, a book in which he condemned the practices that allowed him to gamble with such large amounts of money unchecked.1

1.2.2 Crédit Lyonnais (1996)

A major fire destroyed the headquarters of Crédit Lyonnais in Paris in 1996. It began in the main trading room and spread to burn almost the entire building, along with the safe room, crucial bank

1 http://news.bbc.co.uk/2/hi/business/375259.stm (visited 10 Sept 2014)
archives and computer data. Credit Lyonnais in Paris was forced to activate its disaster recovery plan and move to a back-up site. The insurers evaluated the damages to an amount of 1.6 billion of francs.²

1.2.3 Société Générale (2008)
Jerome Kerviel, 37, was convicted of forgery and breach of trust after an illegal trading scheme amounting to €50 billion resulted in a loss of €4.9 billion at his bank, Société Générale, in 2008. He argued that the bank had known what he was doing but turned a blind eye. This case illustrates the need for companies with risky business models to design effective internal controls.

1.2.4 AXA Rosenberg (2011)
In 2010, AXA Rosenberg sent a letter to its clients and told them that it had made a “coding error” affecting returns in several of its portfolios. One year later, the SEC (Securities and Exchange Commission) charged AXA Rosenberg with securities fraud for “concealing a significant error in the computer code of the quantitative investment model that they use to manage client assets. (…) AXA Rosenberg has agreed to settle the SEC’s charges by paying $217 million to harmed clients plus a $25 million penalty, and hiring an independent consultant with expertise in quantitative investment techniques who will review disclosures and enhance the role of compliance personnel.”³

1.2.5 BNP Paribas (2014)
This summer, BNP pleaded guilty and agreed to pay almost $9 billion to resolve accusations it violated US sanctions against Sudan, Cuba and Iran—despite warnings by some within the firm about the legality and morality of the transactions. The bank had put in place a sophisticated scheme to disguise its role in illicit transactions of billions of dollars.

2 Basel Committee
The regulation regarding operational risk considerably evolved in the two last decades at the instigation of Basel Committee on Banking Supervision. A brief summary of the Basel Committee⁴ is presented below:

- 1988: “Basel Capital Accord” defines the Cooke ratio (Basel I)
- 1993: CAD is implemented in Europe
- 1999 Consultative paper on the McDonough ratio
- 2004 Publication of Basel II regulations: “Revised Capital Framework”
- 2006 Implementation of Basel II regulations
- 2010 Publication of Basel III regulations: “International framework for liquidity risk measurement, standards and measurement”

² http://www.lesechos.fr/22/01/1997/LesEchos/17319-084-ECH_incendie-du-credit-lyonnais--un-sinistre-de-1-6-milliard-de-francs.htm (visited on 10 Sept 2014)
⁴ http://www.bis.org/bcbs/history.htm (visited on 9 Sept 2014)
Quantification of Operational Risks using a Scenario Based Approach

- 2013: Beginning of the implementation of Basel III (expected to be fully implemented by the end of 2017)

The Cooke ratio defined in the First Basel Accord of 1988 calculates the amount of capital a bank should have as a percentage of its total risk-adjusted assets. The calculation is used to determine a minimum capital adequacy standard that must be maintained by banks in case of unexpected loss. The minimum requirement is that the Cooke ratio is greater than or equal to 8%. 5

The McDonough ratio was developed during the Basel II conference and was named after William McDonough, head of Basel Committee. Improvements were made to update the Cooke ratio because the development of new financial instruments was creating problems for determining the risk carried by banks. The major difference comes in the calculation of a bank’s risk-weighted assets. The Cooke ratio did not give different weights to borrowers of differing quality. This was changed under the McDonough ratio, which used the probability of default and the expected loss from default to assign varying weights to assets in the denominator. 6

Basel regulations were designed to improve the way capital requirements reflect underlying risks, among which operational risks. In 2001, a consultative document on Operational Risk by Basel Committee on Banking Supervision 7 states that: “The Committee wants to enhance operational risk assessment efforts by encouraging the industry to develop methodologies and collect data related to managing operational risk. Consequently, [Basel Committee] (...) encourages the industry to further develop techniques for measuring, monitoring and mitigating operational risk.”

3 Solvency II

In the spirit of Basel II, the European Commission wishes to improve the measurement and the management of risks in the insurance industry. Solvency II Directive aims to reinforce the solvency requirements so that the commitments to the insured are guaranteed.

3.1 A three pillars approach

In November 2009, the Directive 2009/138/EC, also called Solvency II Directive, was published in the Official Journal of the European Union and is to enter in force in 2016. This European insurance legislation aims to harmonize the European insurance market and enhance consumer protection. Solvency II also reflects new risk management practices to define, quantify and manage risk. The regulation framework is structured along a three pillars approach:

- Pillar I: Quantitative Capital Requirements

It defines quantitative rules for the computation of its own funds: the MCR 8 and SCR 9 are two levels of own funds that could trigger the intervention of the regulator in order to protect the clients from

---

5 http://www.investopedia.com/terms/c/cookeratio.asp (visited on 9 Sept 2014)
6 http://www.investopedia.com/terms/m/mcdonoughratio.asp (visited 9 Sept 2014)
7 http://www.bis.org/publ/bcbsca07.pdf (visited 9 Sept 2014)
8 MCR=Minimum Capital Requirement
the possible default of insurance companies. The calculations are based on the Value-at-Risk with a 99.5% confidence level over a one-year horizon.

- Pillar II: Qualitative Supervisory Review

It sets out the settlement of a supervisory system. It also sets out requirements for the governance and risk management of insurers by defining key functions such as internal audit, compliance, internal control, risk management. It leads to a reinforcement of internal controls and enhance the power of the regulator. It introduces the ORSA\textsuperscript{10} report that constitutes an auto-evaluation of the transversal risk management.

- Pillar III: Market Discipline

The third and last pillar of Solvency II focuses on the disclosure and the transparency to the different stakeholders of the insurance company. Moreover, it enables harmonizing the reporting processes to the market and to the authorities for European insurance companies.

3.2 The introduction of operational risk measurement in insurance

One of the new topics of this regulation is the introduction of the Operational risk in Pillar I with the calculation of the capital requirements as well as in Pillar II with the set-up of its management. In article 101.4, it can be read that "The Solvency Capital Requirement shall cover at least the following risks: (a) non-life underwriting risk, (b) life underwriting risk, (c) health underwriting risk, (d) market risk, (e) credit risk, (f) operational risk. Operational risk […] shall include legal risks, and exclude risks arising from strategic decisions, as well as reputation risks."

In France, the ACPR\textsuperscript{11} is assigned with the mission of validating and controlling the Solvency II framework developed by insurance companies. The original deadline for the implementation of this European insurance legislation was 2012. It was postponed to January, 1\textsuperscript{st} 2014. To complete the Solvency II Directive, Omnibus II was adopted on the March, 11\textsuperscript{th} 2014. Following this vote, the implementation stage began. It is expected to be concluded at the end of 2015 in preparation of the final entry into force of Solvency II, whose expected date is January, 1\textsuperscript{st} 2016.

Insurance companies have two options to quantify their operational risks: either a standard approach or an internal model. Standard approach is a simplified approach where the Operational risk SCR is calculated as a percentage of their earned premiums and technical provisions. In an internal model,
risks correspond to the actual situation of the firm. It is more than just quantifying its operational risks: insurance companies also have to use their internal model as a steering tool for risk management.

![Figure 1: Repartition of the main risks in the SCR](image)

In EIOPA report on the fifth Quantitative Impact Study for Solvency II, it is stated that the weight of Operational Risks in the SCR of companies that have developed and internal model is around 5%. According to Dan CHELLY, Director at Optimind Winter, it is roughly smaller than for banks (around 8% of the total own funds)\(^\text{13}\).

### 4 Objective of this thesis

This thesis aims to present how insurance companies can quantify their operational risks using an internal approach in order to cope with the specificity of operational risks. As it will be presented further, those quantification techniques have been widely inspired by researches conducted in the banking industry, with the development of Advanced Measurement Approaches introduced by Basel Committee.

Given the heterogeneity and insufficiency of data, modelling operational risks remains a challenging task that requires specific modelling techniques: one of them is called the Scenario Based Approach and will be presented throughout this thesis. This approach is essentially based on scenario analysis: it is a forward-looking approach. Firstly, it consists of identifying all the operational risks borne by the company in order to build a comprehensive risk profile. Afterwards, the frequency and the severity of those scenarios are measured using a blending of expert judgement, internal and external data, when available. The calibration of the distributions is done by quantile-matching that give closed

---

\(^{12}\) EIOPA Report on QIS5 for Solvency II, March 2011 (page 32)

\(^{13}\) http://www.ffsa.fr/webffsa/risques.nsf/html/Risques_89_0006.htm#n5 (9 Sept 2014)
formulae for the distribution parameters. Finally, the aggregate loss is computed virtually using numerical methods.

Quantification is performed so that companies can measure the level of economic capital they have to hold. Therefore standalone quantification of operational risk naturally brings to the question of aggregation. Diversification effects are then crucial, when common sense suggests that operational risks should be, at least partially, decorrelated. Frachot et al. (2004) stated: “that all severe operational risk losses occur simultaneously and systematically in the same year is rather dubious and is hardly supported by empirical evidence”\(^{14}\). Thus, the dependence of operational risks has to be evaluated so that aggregation techniques can be put in place. Because distributions are not elliptical, this thesis details some possible alternatives to linear correlation that could be considered and how they are used in the aggregation process.

As of today, the largest European insurance companies are discussing with their local regulators in order to have their internal models validated. As it is essential for those companies to be able to justify all the assumptions of their internal model to their regulators, this report puts specific attention in ensuring the model consistency with regards of available information. However, the operational framework detailed in this thesis endeavours to develop the internal loss database in order to ensure the adequacy of the expert opinion by back-testing scenarios with available internal data. Nevertheless, the actuary must be aware that discrepancies are inherent to the model. Thus, when designing a model for quantifying risk, an important component to take into account in the sensitivity of the model to changes in the initial assumptions.

The objective of this thesis is to scrutinize the steps of the quantification process in a SBA model in order to discourse about the robustness of the model. The study starts at the standalone level with the quantification of operational risks using a frequency-severity approach and then gradually reaches a more general point of view with the aggregated losses: it intends to take a step backwards in order to gain a better overall understanding of the model. With the objective of injecting a new perspective into a SBA model, alternatives that may enable to circumvent some of the hurdles encountered during the review are presented throughout the development.

---

\(^{14}\) FRACHOT A., RONCALLI T., SALOMON E.. The Correlation Problem in Operational Risk, Groupe de Recherche Opérationnelle Group Risk Management, Crédit Agricole SA, France, 2004
CHAPTER 2
OPERATIONAL RISK FRAMEWORK

This section aims to describe the structure of AXA Group Operation Risk Management team, in which I had the opportunity to pursue my internship. It presents the risk profile to be studied in this thesis. It is largely inspired by the internal communication information available on the intranet.

1 AXA Group

AXA chose to gather its central functions within an economic interest grouping (GIE). In France, a GIE is a grouping which enables its members to share some of their activities in order to develop, improve and increase the results of these activities while they are keeping their own individuality at the same time. Within this organization, lays the Group Risk Management (GRM).

2 Group Risk Management

Risk management in an insurance company is about selecting risks as companies select their clients. The role of risk management teams is to make sure that the top management endorse the risks they hold in their company, understand the consequences of an adverse development of these risks, and have actionable plans if needed. On the intranet, it can be read that “Our role is to create a secure risk framework encouraging effectively managed underwriting, guaranteeing protection for the company over the long term”.

Group Risk Management’s missions and main activities are:

- Creating a secure risk framework

By clarifying decision-making in order to ensure better risk selection. For this, GRM has defined a series of standards, such as the Product Approval Process (PAP), covering all new products before their market release, or Risk Appetite, defining the limits within which the Group would like to operate.

- Protecting the company over the long term

By testing AXA’s capacity to overcome all types of major crises. For this, GRM teams use a series of stress scenarios to assess AXA’s capacity to withstand rare and extreme conditions with multiple impacts.

- Implementing the Solvency II project

By driving the process for approval with the regulator and implementing AXA’s internal model for computing the Economic Capital to guide decision-making.
• Developing the risk culture

Building awareness among operational staff on the importance of good risk management enables to enhance the risk culture. For this, GRM works to develop a strong risk culture through training and communication.

The Group Operational Risk Team belongs to the GRM.

3 Group Operational Risk Team

The objective of this team is to ensure that a consistent framework is in place across the Group to systematically identify, measure, mitigate, report and monitor the most significant Operational risks that the company may face. The team covers all insurance operating entities, banking activities, asset managers, pension funds, service companies and holding activities: it coordinates a network of local operational risk teams, which conduct processes of local ownership regarding operational risk.

It consists mainly of the identification of the most significant Operational risks, collection of the actual losses in a Group database, quantification of the frequency and severity associated to each main Operational risk, inclusion of an Operational risk measure of these risks into the Group Solvency Capital Requirement (SCR), compliance with regulatory requirements (Solvency II) which require quantification of these risks. It ensures the resilience of entities to face catastrophic or extreme risk scenarios and the monitoring of the operational risk profile consistent with the Group risk appetite/tolerance framework.

4 Operational Risk Framework

4.1 AXA’s Definition of Operational Risk

Based on Basel II and Solvency II definition, AXA defines Operational risk as the “risk of loss arising from inadequate or failed internal processes, or from personnel and systems or from external events”. It also detailed that:

- Failure or inadequacy may result from both internal and external causes
- Legal impacts are included
- Reputation risks and risks arising from strategic decisions are excluded.

In order to best define Operational Risk, AXA highlights the main differences between Operational Risk and “classic risks” that are insurance and financial risks.

For instance, think about the common interpretation of taking risks. Usually, it is said that more risk means more expected returns. In operational risk it is not the case: more risks only means more losses, with no upside.
One of the specificities of Operational Risk lies in the fact that it is a recent discipline, which entails that it cannot rely on a long experience. Contrary to “classic risks” for which data have been collected and analysed for years leading to complete, consistent and homogeneous historical records, operational risk consistency is challenging as collecting operational loss data is an issue because of the limited industry experience.

As we will discuss more in details after presenting AXA’s cartography of risks, the scope of operational risk is also particular: it is wider than for well delimited insurance and financial risk portfolio. Thus, special attention needs to be put in the delimitation of boundaries between operational risks and other types of risks in order to avoid double-counting.

4.2 Group risk grid

Segmentation of such heterogeneous set of risks as operational risk constitutes a crucial assessment process. A common approach starts by using the operational risk categories from Basel II framework, that is:

- IF: Internal Fraud
- EF: External Fraud
- EPWS: Employment Practices and Workplace Safety
- CPBP: Clients, Products and Business Practices
- DPA: Damage to Physical Assets
- BDSF: Business Disruption and System Failure
- EDPM: Execution, Delivery and Process Management

Within AXA, Operational risks are grouped into those 7 risk categories. But the actual cartography used for quantification purpose is thinner: it consists of 100 event-types. The segmentation comes from the need of explaining the different categories to the experts in a more granular way.

4.3 Boundaries with other risks

Operational risk encompasses any type of risks which are triggered by an operational failure or an external event. The definition of Operational Risk might therefore lead to include risk events that have operational failures, but might be already included in the Solvency Capital Requirement related to other risks (Financial/Market or Insurance risks).

- Operational Risk vs. Financial risks

Operational risk failures that cause a market, credit loss are reported as Operational risk, whereas Operational risk failure outside of the company that causes market risk losses are not included into Operational risk, but under the Market SCR.

- Operational risk vs. Insurance risks
  - Reserves risks
A reserve risk arises from adverse development of claims payments resulting from events occurred in previous accident years. Operational risk failures that cause a misestimating of the reserves are on principle reported as Operational risk.

- Claims fraud

When modelled, insurance risks include a portion of claims fraud (detected and undetected) which is embedded in the P&C loss ratio and reserve development. Operational risk however focuses on low frequency/high severity claims fraud events. By exception, a certain amount of double counting may exist reflecting a conservative approach.

- Operational risk and strategic risks

Strategic risk is the risk that a negative impact (current or prospective) on earnings or capital, material at Group level, arises from a lack of responsiveness to industry changes or adverse business decisions

- Operational risk and legal risks

Regulatory risk is the risk relating to the evolving legal and regulatory environment in which AXA operates. It is considered separately from compliance or litigation risks related to compliance with current regulation and laws, which are monitored by the Compliance and Legal functions, captured through the operational risk assessment framework or reflected in the reserves for asserted claims or commenced investigations. In consistency with Solvency II definition on Operational risk and AXA risk grid, Operational risk includes legal and regulatory loss arising from an Operational failure. Legal impacts include litigation, arbitration and tribunals. They typically include fines, penalties, punitive damages, private settlements, external lawyers’ fees, court fees, other litigation expenses, etc.

- Operational risks vs Reputational risks

Reputation risk is the risk that an event will negatively influence stakeholder’s perceptions. In consistency with Solvency II definition on Operational risk, reputational risks are excluded. However reputation impact caused by an operational risk event is qualitatively captured through the operational risk assessment exercise and shared with Group communication.

At the very beginning, Basel Accord suggested that banks should take into account direct financial losses as well as indirect losses triggered by an operational risk event. Later on, Basel Committee only kept direct financial losses in order to ease and increase the robustness of operational risk understanding by avoiding any speculation on any hypothetical indirect loss. However, measuring operational risks remains a delicate problem whereas their quantification has become mandatory.
CHAPTER 3
INDUSTRY PRACTICE

1 Operational Risk Data

In spite of the risk classification and the improvement of the loss data collection, the profile of operational risks remains atypical. As shown in figure 1, there is heterogeneity within risk categories and across risk categories. Moreover, the quantity of available data highly depends on the risk categories, certain categories being almost empty.

![Illustrative boxplot of internal loss data for the 7 event risk categories]

Operational risk modelling is a challenging task given the heterogeneity and insufficiency of data for all the risk types that fall into its scope. Different sources of information are available, namely:

- Internal data
- External data
- Expert opinion

These different sources do not give information about the same regions of an operational risk loss distribution. It is expected that internal data should give information about the attritional events, followed by external data and expert opinion to assess the distribution tails.

1.1 Basel II recommendations

In Basel Committee’s Supervisory Guidelines for the AMA paper, it is said that: “An advanced measurement approach for calculating the operational risk capital charge of a bank requires the use of four data elements which are:

- ILD: internal loss data.
According to Basel II, “supervisors expect internal loss data to be used in the operational risk measurement system to assist in the estimation of loss frequencies, to inform the severity distribution to the extent possible and to serve as an input into scenario analysis.” Hence, Internal Loss Data can be used directly by fitting a statistical distribution to the data or indirectly as a reference source to inform the risk assessment.

- ED: external data.

It contributes to the sparse amount of low frequency-high impact data that helps to estimate the tail of an operational risk distribution. External data provides valuable information about the losses experienced by peers. It is useful for an insurance undertaking to quantify its exposure to risk events that have not been experienced internally.

- SA: scenario analysis.

Scenario Analysis focuses on the examination of rare, significant, yet plausible future events. By involving experts in the risk assessment process, it participates to the commitment of managers in the evaluation of operational risks; scenario analysis provides an important link between the measurement and management of operational risk.

- BEICFs: business environment and internal control factors.

“BEICFs are operational risk management indicators that provide forward-looking assessments of business risk factors as well as a bank’s internal control environment. BEIFCs are commonly used as an indirect input into the quantification framework and as an ex-post adjustment to model output. Ex-post adjustments serve as an important link between the risk management and the risk measurement processes.”

Those four sources of information must be taken into account when building an internal model to be validated by the regulator.

1.2 The four data elements

This section describes the characteristics of these sources of information.

1.2.1 Internal data

It seems highly logical that any company that would be confronted to any claim would record it in an internal database with an appropriate classification. Nevertheless, it is not always the case and companies’ internal databases are not that thick, especially when it comes to operational losses.

The main issue about the internal loss database is that reconciling to the category is possible, though scenarios are broken down into risk events of higher granularity. Given the multiplicity of the risks in operational risk scope, it is unlikely to obtain a full database. The heterogeneous nature of
Operational risk within a category makes it more difficult to use an approach based on historical data (LDA\textsuperscript{15} method), in particular for rare and extreme events.

Moreover, internal controls can introduce a bias in the historical data that becomes hard to overcome. For instance, it is more likely that a major loss detected would go back up to the risk management and then be recorded in the loss data collection. Furthermore, internal controls could be more efficient for larger amounts at stake. Conversely, we can assume that many minor losses would remain undetected or if detected, the information would not always be passed on because the impact would be considered to be insignificant. Consequently, major losses could tend to be over-weighted in the historical data, whereas minor losses could be under-weighted: databases are biased to the extreme levels. Consequently, banks and insurance companies tend to have recourse to external databases.

### 1.2.2 External data

Insurers can rely on two sources of information:

- Public data (e.g. newspapers articles)
- Consortiums: some companies pool their operational risk databases. For instance, ORX (Operational Risk eXchange association) Global Loss Database contains almost 400,000 operational risk loss events over €20,000 in value to a total value of €213 billion. As of today, ORX is own and controlled by its 67 member financial service companies from 20 countries, on an equal basis. ORX is currently working with a group of leading insurance companies to establish the ORX Global Insurance Service.

Nevertheless, finding a robust source of external data is not easy because companies do not feel like circulating operational losses that could have a significant reputational impact. Moreover, confidentiality agreements would hinder the disclosure of accurate information. Furthermore, it is a challenge to find a classification that would be suitable for everyone and produce a coherent database.

The context should also be taken into account: internal controls are not the same in all insurance companies. Hence, all databases are not comparable. Operational losses are also sensitive to the regulatory, political and economic environment: major changes should be taken into account in order to build a coherent external database.

In addition, the historical data bias is still present in external databases. For instance, let’s assume that there are 10,000 losses that cost 1,000,000 euros each, and 10 losses of 1 billion each. It is very likely that all the major losses would have been reported. It is also likely that only a tenth of the 1 million euros losses would have been reported. In the end, the ratio would be 100/1 instead of 1000/1, which would lead to wrong results for the loss distribution calibration.

\textsuperscript{15} LDA = Loss Distribution Approach
1.2.3 **Expert judgment for the Scenario Analysis and the BEIFCs**

In order to overcome the difficulty of the scarce data, an alternative used for the quantification assessment is the expert judgment.

Though, relying on expert opinion for the assessment of the risks borne by the company implies the introduction of an expert judgment bias, conscious or not. We can distinguish different types of bias in the traditional dissonances in expert opinion:

- **Overconfidence**: Experts could overestimate the precision of their own knowledge and over-rely on limited evidence.
- **Optimism and wishful thinking**: There is a danger of experts having unrealistically bright forecasts. In particular, managers tend to be too optimistic about the effectiveness of their reactions to extremely rare and adverse events.
- **Sample size neglect**: People often assume a sample is embedded with all the key properties of the population, without taking into account the size of the sample.
- **Group polarization**: This refers to the tendency of groups to adopt more extreme positions than would individual members.
- **Motivation bias**: Incentive conflicts that arise when participants have a vested personal interest in making either themselves or the results of scenario analysis programme more advantageous.
- **Availability or memory**: There is a people’s propensity to overestimate the likelihood of incidents they had close or recent contact with. Also, we pay excessive attention to events accompanied by high drama but overlook events that happen in a routine fashion.
- **Framing**: Inconsistent choices or predictions for the same problem when phrased slightly differently.
- **Representativeness**: Misinterpretation of links between two hypothetical events leading to insufficient attention to individual probabilities of occurrence.
- **Anchoring**: Central values often work as anchors. Experts start with them as reference and the adjustment is often too small. For instance, the expert is fully influenced by the scenarios pre-existing design and does not bring sophistication from his own knowledge.

2 **Standard formula**

As of today, many firms are still using the standard formula for calculating their operational risk capital charge. Indeed, many firms are using the standard formula for all risks. Most of the companies using internal models tend to be large groups. According to Milliman\(^\text{16}\), many of the companies using internal model still currently use the standard-formula approach for the operational risk component. The detailed standard formula presented hereafter was issued by CEIOPS in QIS5 Technical Specifications, issued in July 2010. It is detailed below:

The inputs for the Operational Risk module are the following:

Quantification of Operational Risks using a Scenario Based Approach

- $TP_{life}$ = Life insurance obligations. For the purpose of this calculation, technical provisions should not include the risk margin, should be without deduction of recoverable from reinsurance contracts and special purpose vehicles.
- $TP_{life-\text{ul}}$ = Life insurance obligations for life insurance obligations where the investment risk is borne by the policyholders. For the purpose of this calculation, technical provisions should not include the risk margin, should be without deduction of recoverable from reinsurance contracts and special purpose vehicles.
- $TP_{nl}$ = Total non-life insurance obligations excluding obligations under non-life contracts which are similar to life obligations, including annuities. For the purpose of this calculation, technical provisions should not include the risk margin and should be without deduction of recoverable from reinsurance contracts and special purpose vehicles.
- $pEarn_{life}$ = Earned premium during the 12 months prior to the previous 12 months for life insurance obligations, without deducting premium ceded to reinsurance.
- $pEarn_{life-\text{ul}}$ = Earned premium during the 12 months prior to the previous 12 months for life insurance obligations where the investment risk is borne by the policyholders, without deducting premium ceded to reinsurance.
- $pEarn_{nl}$ = Earned premium during the 12 months prior to the previous 12 months for non-life insurance obligations, without deducting premium ceded to reinsurance.
- $Earn_{life}$ = Earned premium during the previous 12 months for life insurance obligations, without deducting premium ceded to reinsurance.
- $Earn_{life-\text{ul}}$ = Earned premium during the previous 12 months for life insurance obligations where the investment risk is borne by the policyholders, without deducting premium ceded to reinsurance.
- $Earn_{nl}$ = Earned premium during the previous 12 months for non-life insurance obligations, without deducting premiums ceded to reinsurance.
- $Exp_{ul}$ = Amount of annual expenses incurred during the previous 12 months in respect life insurance where the investment risk is borne by the policyholders.
- $BSCR$ = Basic SCR

All the aforementioned inputs should be available for the last economic period and the previous one, on order to calculate their last annual variations.

The capital charge for operational risk is calculated as follows:

$$SCR_{op} = \min\{0.30 \times BSCR; Op_{\text{nl}}\} + 0.25 \times Exp_{ul}$$

where $Op$ = Basic operational risk charge for all business other than life insurance where the investment risk is borne by the policyholders.

$$Op = \max\{Op_{\text{premiums}}; Op_{\text{provisions}}\}$$

where $Op_{\text{premiums}} = 0.04 \times (Earn_{life} - Earn_{life-\text{ul}}) + 0.03 \times Earn_{nl}$

$$+ \max\{0; 0.04 \times (Earn_{life-\text{ul}} - 1.1 \times pEarn_{life} - (Earn_{life-\text{ul}} - 1.1 \times pEarn_{life-\text{ul}}))\}$$

$$+ \max\{0; 0.03 \times (Earn_{nl} - 1.1 \times pEarn_{nl})\}$$
\[ OP_{\text{provisions}} = 0.0045 \times \max\{0; TP_{\text{life}} - TP_{\text{life-ut}}\} + 0.03 \times \max\{0; TP_{\text{nl}}\} \]

While the Solvency II standard formula applies a factor-based approach to calculating the operational risk, it remains unclear if this approach really captures the risks specific to the insurance industry, and the varying risk profiles of different insurance companies. An actuary from Standard Life UK said that: "The standard formula approach to measurement of operational risk does not reflect the operational processes and controls that a business has in place. In particular, as the charge for unit-linked business depends only on your expense levels, companies that have strong controls and low risk exposure would get the same capital charge as companies with similar expenses levels, but with weak controls and higher risk exposure."\(^{17}\)

To address those shortcomings, some insurance companies have decided to implement an internal model for their Operational risks.

### 3 Internal model

The choice between an internal model and a standard formula is not straightforward. Many companies are avoiding the internal model given its high cost of implementation. According to a study on Solvency II conducted by Optimind/Opinionway\(^{18}\) in 2013, 49% of the French insurance companies use the standard formula for computing the SCR. In addition, approximately a quarter of those companies are using a partial internal model. Only 12% of them state that they are using an internal model. According to a study published by Milliman\(^{19}\), where companies are using an internal model approach to calculating the Operational Risk SCR, many are using methodologies similar to those employed by banks under Basel II.

When building an internal model to be validated by the regulator, insurance companies refer to what has been done in the banking industry. Indeed, there are several manners to include each of the four data elements either directly or indirectly in the risk evaluation. The most sophisticated and complex option to calculate regulatory capital for operational risk under Basel II is the Advanced Measurement Approach (AMA). This is the only risk-sensitive approach for operational risk allowed and described in Basel II. Its most significant benefit is that it improves risk management processes thanks to the sharp risk analysis. Also, those approaches can allow for a reduction of the capital charge.

#### 3.1 Loss Distribution Approach

The most popular AMA methodology is by far the Loss Distribution Approach (LDA). It is an application of actuarial methods that consists in fitting statistical distributions for the frequency and

---


the severity with available historical data. The risk profile of the company is split into risk cells where the columns represent the Event Type (EL\textsuperscript{20}) and the lines are the Business Lines (BL). For each risk cell (i.e. for one type of event in one business line of activities), this modular procedure calibrates a loss distribution with the available historical data.

<table>
<thead>
<tr>
<th>BL1</th>
<th>EL1</th>
<th>EL2</th>
<th>EL3</th>
<th>EL4</th>
<th>EL5</th>
<th>EL6</th>
<th>EL7</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Modular view of operational risks split into 7 Event Lines and 8 Business Lines

Under the Loss Distribution Approach, the company estimates the probability distribution functions of the single event impact and the event frequency for the next year using its internal data, and computes the probability distribution function of the cumulative operational loss. The correlation of distributions between different cells has to be modelled: it requires to estimate the dependency between univariate distributions to produce the joint distribution.

This methodology utilises actual and verified data, which does not leave room to subjectivity. On the other hand, it consists in using historical data to model the future: it is a backward-looking approach. Moreover, as operational officers cannot escalate all losses, there is a reporting threshold in the loss data collection: only losses above the said threshold are collected. This bias should be taken into consideration when calibrating the loss and frequency distributions by introducing adjustments in the curve fitting. Another drawback of this method is the large amount of data needed for this modelling purpose. As it has been discussed, operational risk is still a recent discipline and the databases are not thick enough yet. Thus, this method should be retained for modelling high frequency and low severity risks, for which the number of observations would be satisfying for modelling purpose.

### 3.2 Bayesian Approach

Bayesian inference is a statistical technique well suited for combining expert opinions and historical data: expert judgment is incorporated into the model via specifying distributions for the model parameters, called the prior distributions. It consists in building a Bayesian network which is a probabilistic causal graph derived from Bayes’ theorem:

\[
P(A|B) = \frac{P(B|A)P(A)}{P(B)}.
\]

In particular, each node in the graph represents uncertain variables of interest, while the arrows between the nodes are the causal links between the corresponding random variables. The structure of the Bayesian network is defined by experts. Then, the parameters can be determined with expert judgment or with statistical data. In the end, the different Bayesian networks are grouped to create a global network: the aggregate loss is the sum of all losses.

\textsuperscript{20} EL: Event Line
The main advantage of this approach is the link of causality between variables. There is no need to estimate correlation between risks: by construction, when launching a simulation that goes through every random variable, the risks are aggregated with the estimated conditional distributions. The correlations are a result of the design of the network. However, because of the variety of operational risks, the implementation of such a Bayesian network for each individual scenario takes a long time.

This approach is mainly based on scenario analysis: it consists in identifying all the operational risks borne by the company and quantifying their frequency and severity. The frequency module is broken down into several frequency criteria and the financial impact is decomposed into several impact types. Experts are asked to assess each of the components of the scenario so that he provides an estimation of the average frequency as well as the average and extreme impacts. This method is mainly motioned through expert judgment but available internal and external loss data is also utilized to feed the risk measurement. Usually, quantile-matching methods are used to fit the assessment points with frequency and severity distribution. The aggregate loss is then computed with numerical approximations such as Monte Carlo simulations, taking into account diversification effects. As it better represent its risk profile, this forward-looking approach was retained by AXA for modelling its operational and will be presented more in details in the next section.

3.3 SBA

This approach is mainly based on scenario analysis: it consists in identifying all the operational risks borne by the company and quantifying their frequency and severity. The frequency module is broken down into several frequency criteria and the financial impact is decomposed into several impact types. Experts are asked to assess each of the components of the scenario so that he provides an estimation of the average frequency as well as the average and extreme impacts. This method is mainly motioned through expert judgment but available internal and external loss data is also utilized to feed the risk measurement. Usually, quantile-matching methods are used to fit the assessment points with frequency and severity distribution. The aggregate loss is then computed with numerical approximations such as Monte Carlo simulations, taking into account diversification effects. As it better represent its risk profile, this forward-looking approach was retained by AXA for modelling its operational and will be presented more in details in the next section.
CHAPTER 4
THE SCENARIO BASED APPROACH

AXA has developed an Operational risk internal model to reflect its risk profile, taking into account local and Group specificities, addressing shortcomings in Solvency II standard formula and allowing for a better evolution of the model over time. It is a forward-looking and Scenario-Based Approach (SBA). It consists of identifying, assessing and measuring the most critical Operational risks of each entity complemented by a set of transversal Group scenarios. The scenarios are forward-looking and rely on expert judgment back-tested with any appropriate documentation, internal or external previous experience. These scenarios intend to stress the severity of some operational failures, which may have never happened in the past, to test the resilience of the firm to face extreme and potentially catastrophic types of situations. Other complementary elements of the framework such as internal losses, external losses, business environment and internal control factors and key risk indicators contribute to strengthen the accuracy of the expert opinion approach in selecting and assessing the most critical risk situations.

The main goal of using an internal model as opposed to the standard formula is to better reflect the company’s risk profile in the SCR. This is particularly relevant for Operational risks, as the Standard formula for Operational risk is purely factor-based, with no risk factors related to any Operational risk criteria. Internal model is considered from several aspects:

- Taking into account local specificities: AXA is a global company, and caters to a wide range of insurance markets with a variety of products offered and targeting certain demographics and with differing risk exposures
- Addressing shortcomings of the Standard formula: Operational risk Standard formula is not risk-sensitive whereas internal models adequately reflect the accuracy of the Operational risk profile of the company
- Allowing for better evolution of the model over time: As the Group experience increases, business expands to new markets and product innovations create different risks to consider, the flexibility of an internal model allows the specificities of these developments to be reflected.

1 Operational risk SCR

Operational Risk follows Group SCR framework, guidelines and timeframe. Operational Risk component of the Economic Capital is based on the main local Operational Risks assessed and reported by each operating entity. Local validation by a dedicated governance body is ensured by local Risk Management, before submission of these Operational Risks to Group Risk Management.

SCR is meant to be the internal model to measure Solvency Capital Requirement (SCR) in Solvency II. It is based on a one year horizon Value-at-Risk with a 99.5% confidence interval.
Operational risk quantification process aims to assess at a certain confidence level the maximal financial loss that can be caused by operational risk failures over a one year period:

- On a standalone basis for each risks.
- On an aggregated basis: at business lines, entity and Group level.

The calculation of the Operation Risk SCR figure is based on a Monte Carlo approach and performed as follows:

\[
SCR = V\text{a}R_{99.5\%} - EL
\]

where \( V\text{a}R_{99.5\%} \) is the 99.5% percentile of the \( n \) simulations of the aggregate loss

\( EL \) is the expected loss of the \( n \) simulations

2 Standards

2.1 Single methodology

The implementation of the Operational risk framework is not limited to insurance activities. It concerns all entities, including insurance companies, banking activities, asset managers and internal service providers. All the entities within the Group follow a single methodology, a single typology and use a single tool to identify, measure, quantify and mitigate their Operational risks. GRM organises a series of trainings, coordinates the entities and conducts a benchmarking that is shared with all the entities. This methodology enables to gain the buy-in of the risk managers and strongly supports the development of a risk culture in line with the Solvency II spirit.
One of the major points in Solvency II is described in Pillar II and points out the need for insurance companies to identify the risk they bear. In order to optimize the management of their risks, companies shall put efforts on the identification of their risks before working on their quantification. Aligned with these requirements, a Scenario Based Approach for Operational Risks enables to build a risk profile for each entity of the Group.

A study conducted at Group level showed that on average, the 20 biggest risks of an entity contribute to most of its capital charge. Nevertheless, a minimum number of reported is required to ensure a comprehensive risk profile. Moreover, this approach contributes to the risk culture and awareness as well as it reflect the business and promote discussions between the main stakeholders: operations, transversal function, management and risk management. As the entities’ maturity improves, the number Operational risk scenarios reported increases. The number has stabilized which entails that entities are getting mature enough to well identify their Operational risks. In this context, increasing the number of top scenarios is no longer a priority rather than the validation and use test of this risk profile to set up mitigation actions.

Scenario Based approach is complemented by additional elements of the framework such as Internal losses, External losses, Business Environment and Internal control factors and Key risk indicators.

2.2 Top-down vs. Bottom-up

Scenario analysis could typically be focused either on group driven ‘top-down’ measurement models, or business driven ‘bottom-up’ management models. The methodology presented in this thesis is a convergence of those two approaches.

- Top-down

A set of transversal “Group scenarios” are identified and assessed following a forward-looking and expert-opinion approach: they are the major risks identified at corporate level. Group scenarios are similar risk situations impacting at least two different entities with a similar cause and potentially requiring activation of crisis management. The quantification of these scenarios is consolidated at Group level with the input of each relevant entity which follows Group instructions for the quantification of these risks. These scenarios are then aggregated to estimate the capital allocation needed to cover Operational risks based on advanced models aligned with Solvency II principles and consistent with the advanced Operational risk models developed in the banking industry.

- Bottom-up

Each entity reports on an annual basis its most critical operational risks identified from local assessments to Group Risk Management. These are called “local scenarios”. These risks are quantified through local workshops involving relevant experts, transversal or support functions and business lines managers and following a single Group methodology.
3 Guidelines and Methodology

GRM has defined Group minimum requirements on Operational Risk Management organisation and deliverables. The objective is to ensure that a framework is in place to:

- Identify the most significant Operational risks, detect the sources of risks and collect the actual losses in a Group database
- Quantify the frequency and severity associated to each main Operational risk in compliance with regulatory requirements (Solvency II)
- Mitigate by putting in place concrete action plans in order to ensure the resilience of entities to face catastrophic or extreme risk scenarios
- Report & monitor to ensure that the risk profile is consistent with the risk appetite framework.

The guidelines and methodology are presented in this section, which is highly inspired by a presentation of AXA’s Operational Risk Management by the head of Operational Risk Team.

3.1 Identification: Risk assessment

The objective of a risk identification process is to understand the scope of the risks that the entire organization and its strategy are exposed to. The qualitative risk assessment process should be practical, sustainable and easy to understand. It must proceed in a structured and disciplined fashion in order to produce a comprehensive list of risks. Its purpose is to assess how big the risks are, both individually and collectively, in order to focus management’s attention on the most important threats, and to lay the groundwork for risk response.

The data used in modelling are based on quantitative risk assessment’s outputs. Out of the operational risks reported, the top operational risks must be detailed using a scenario template. The document gives the assessed levels for frequency and severity and gives a full qualitative description, background, and context of the scenario. It also describes the criteria chosen for frequency and severity quantification in line with the Technical documentation.

Operational risk modelling is based on a frequency/severity approach. The annual loss is equal to the sum of random number of losses. This section details how the frequency and severity are being assessed.

3.1.1 Frequency measurement

Frequency describes the number of events that occur given a period of time. In our case, it is a one-year period. The frequency of a risk is measured considering three different elements as contributing to a potential loss. These are:

- a triggering factor: it is the cause or sources that make the risk occur
- a resource failure: it represents the vulnerability of the company to a specific type of failure (given the failure considered as triggering the risk)
- a loss generating event: the events producing a financial loss
These three elements contributing to a potential loss can occur separately or be combined.

Once these factors have been identified and their likelihood has been measured, the frequency is computed the following way:

\[
Frequency = \left\{ \begin{array}{l}
\text{Likelihood of occurrence of the triggering factor} \\
\times \text{Likelihood of having a resource failure} \\
\times \text{Likelihood of occurrence of the loss generation event}
\end{array} \right.
\]

### 3.1.2 Impacts measurement

Severity assessment is performed in order to give the amount of an incurred loss. Given the scarcity of data and the diversity of the risks, a global approach that can be replicated throughout the Group is needed. A straightforward approach is to consider parametric families of distributions, and calibrate them given assessed fitting points. Moments, mode and quantiles could be considered. The approach for the self-risk assessment provides a good illustration of the advantages and difficulties brought by the structure of parameterization. Let’s first describe it so we may after bring light into the aspects that need our attention.

The diversity of the risks considered has to be controlled by a limited number of parameters giving the shape of the loss distribution. When conceiving the internal model, it was decided that three impact levels would be given to assess the severity distribution. With most of distribution laws used in industry being determined by two parameters, the choice of giving an extra level was motivated by the fact that expert opinion would need more than one value to measure risk in the tails.

The expert is asked to assess:

- **Typical impact:** it is the mode of the loss distribution that is to say the most frequent value the impact will take.
- **Serious impact:** it is the $\beta$ quantile of the severity distribution (i.e. the worst amount of financial loss expected once every $\frac{1}{1-\beta}$ incident).
- **Extreme impact:** it is the $\gamma$ quantile of the severity distribution (i.e. the worst amount of financial loss expected once every $\frac{1}{1-\gamma}$ incident).

These three levels are embedded in the risk assessment template and are parameterized in the Operational risk tool. In the template, the expert details all the possible impacts that could be triggered if the scenario occurs. For each of them, he estimates the level of loss that could happen in the typical, serious and extreme case. The impact of the scenario is split into several blocks in order to give a more accurate estimation of the level of the impact.

Training and education guidance ensure a common and shared understanding of what typical, serious and extreme impact levels are representing and thus ensuring expert estimates are consistently considered. Indeed, it is mandatory to ensure a good understanding of the calibration points for an expert based approach.
3.1.3 Documentation of the top scenarios

It is acknowledged that the scenario design process is, to a certain extent, expert opinion driven. Therefore a data quality process for operational risk is more about ‘the quality of the evidence’ used to support the assumptions and data used to feed the risk scenarios. AXA methodology reduces to the minimum the level of uncertainty of the assumptions taken to model risks. The documentation of the top scenarios must include the description of the context in which the risk occurs and define the nature of the causes, the type of failures and the impacts that the event would have. To incorporate the four data elements into the internal model, most criteria used to calculate frequency and severity are based on existing information that can be sourced such as scorecards, management data, internal or external statistics and databases.

Appropriate level of documentation ensures first that the coverage and comprehensiveness of the identification of potential risks incorporates all business activities and demonstrates that no significant risks have been missed. Then, it contributes to the accuracy of the criteria used in order to define the frequency and severity of the scenarios.

3.2 Identification: loss data collection

The firm insists on the fact that the collection of past errors is not a punitive process but as a way to prevent losses from occurring again, mitigate their impacts and spread risk culture. Management exemplarity is required and the blame should focus on the concealment of losses. Loss data collection is one of the components to meet Solvency II use test. It covers 5 objectives:

- Risk culture: increase Operational risk knowledge and risk awareness across the business;
- Prevention: Escalation of major issues and prevention actions across the Group, ensure actions are taken in front of major losses in order to mitigate their impact;
- Draw lessons from past errors: learn from Operational losses that occurred in order to prevent them recurring in the entity or across the Group;
- Back test: complement the development of the modelling approach, in particular by a better understanding of correlations and probabilities;
- Regulation: supervisors expect companies to systematically collect Operational risk data in an internal database. It is part of the regulatory expectations for internal model approval.

The reporting process was launched in 2008, with losses prior to 2008 being reported as well. As of today, around 40 000 losses have been reported in the database by all the entities.

The most mature databases are in categories:

- Suitability, Disclosure, Fiduciary and Advisory Activities: it is the largest category in terms of amount of losses because of some losses that had a very significant impact;
- Breach of Regulation and Legislation: second largest category in terms of amount of losses because life entities are very exposed to regulatory breaches;
• Transaction Capture, Execution and Maintenance: third in terms of amount, and first category in terms of number of reported losses because it includes a large number of high frequency risks (e.g. trade execution errors).

3.3 Quantification

Risk measurement is an important component to rank more precisely the risk identified during the identification phase in a consistent manner and to provide rationale for capital allocation. It allows senior management to consider systematically the risk of extreme but plausible events. This helps them proactively consider how to react quickly and decisively should the worst happen. It is essential to review the results of scenario analysis to ensure that consistent and defensible estimates and outcomes are delivered.

3.3.1 Stand-alone quantification

The most critical risks of each entity identified through the risk identification phase described above are quantified with comparable manner and metrics around across the Group. The common practice is to compute the annual loss distribution with Monte Carlo simulations. The frequency is modelled independently from the severity.

Their combination gives the annual loss with the following formula:

\[ X = \sum_{i=1}^{N} C_i \]

where \( N \) is the number of events that occur within a year

\( C_i \) is the amount of each single loss event. The losses are assumed to be independent identically distributed (iid).

Moreover, \( N \) is assumed to be independent of \( C_i \)
With the independence between the frequency and the severity, we can write the expectation and variance of the annual loss in terms of those of the frequency and the severity:

\[
\mathbb{E}(X) = \mathbb{E}(N)\mathbb{E}(C)
\]

\[
\text{Var}(X) = \mathbb{E}(N)\text{Var}(C) + \text{Var}(N)(\mathbb{E}(c))^2
\]

In addition, the cumulative distribution function of the annual loss is given by:

\[
F_X(x) = \begin{cases} 
\sum_{n=1}^{+\infty} \mathbb{P}(N = n)F_C^k(x), & \text{for } x > 0 \\
\mathbb{P}(N = 0), & \text{for } x = 0
\end{cases}
\]

where \(F_C^k(x)\) is the cumulative distribution function of \(k\) losses \((C_i, \ldots, C_k)\) obtained by the convolution of order \(k\) of the severity distribution.

Usually, the frequency variable \(N\) is assumed to follow a Poisson distribution, but other distributions such as Bernoulli, negative Binomial or fixed frequency are sometimes considered. Depending on the characteristics of the levels of impacts, the severity variable \(C_i\) is assumed to follow a lognormal or a Pareto distribution. Again, other distributions could be considered, but the choice needs to be justified and documented. It is assumed that the frequency variable is independent from the severity variables.
variables, and the severity variables $C_i$ are assumed to be independent identically distributed. More details about the distributions are given in next chapter. The metrics chosen to measure the risks are the 99.5%-quantile and the expected loss of the distribution. They are computed through Monte Carlo simulations.

### 3.3.2 Aggregation

These risks are then aggregated using Monte Carlo methods to estimate the capital allocation needed to cover operational risk. A diversification effect is introduced between different categories of risk and different entities with correlation coefficients that measure the dependence of variables. The concept of correlation will be detailed in a dedicated chapter as well as the whole aggregation process that is dealt with in another section.

### 3.4 Monitoring & Reporting

Risk response strategies are the approaches to dealing with the risks identified and quantified. Risk response strategy is really based on risk tolerance. Risk tolerance in terms of severity is the point above which a risk is not acceptable and below which the risk is acceptable. Several strategies are available for dealing with risks: avoidance, acceptance, mitigation or transfer.

#### 3.4.1 Mitigation actions

An effective risk management should put in place a dedicated risk response strategy in front of major risk. The following types of actions can be considered within the business case:

- Improve process;
- Process formalisation: policy and procedures;
- Process implementation: application of policy and procedures within the organisation;
- Process design: sequence of tasks, owners...;
- Improve controls;
- Scope and frequency of control;
- Improve human resource allocation;
- Staffing: recruit additional headcounts;
- Skilling: recruit more senior staff, develop training;
- Improve tool;
- Implement a new tool to support the process;
- Enhance the tool functionalities...

Losses with a significant financial impact must lead to the definition of a risk mitigation strategy. In practise, the risk managers must fill in the description of the risk event (including timeline and total estimated loss amount), the control failings, the corrective actions and management response, and signals to be considered to prevent losses from occurring again. Once again, the objective is not to put in place a punitive process but to contribute to spread the risk culture across the Group.
3.4.2 Risk appetite and KRI

Risk appetite describes the types and degree of risk an organization is prepared to incur. As a result, risk appetite refers to an organization’s attitude towards risk taking and whether it is willing and able to tolerate either a high or a low level of exposure to risk. With mounting regulatory demands faced by many financial institutions, and other regulated industries, a robust operational risk appetite statement assists organizations to execute their strategy execution by clearly defining the boundaries within which they can operate. Risk appetite statement can be compared to losses, quantification and Key Risk Indicators and focus the attention of management to foster mitigation actions.

A Key Risk Indicator (or KRI) is a metric that provides information on the level of exposure to a given operational risk which the organization has at a particular point in time. The concept of a threshold or limit is to establish boundaries that, when exceeded, alert the organization to a significant change in risk exposure. It will provide an early-warning system to identify potential process failures and/or control issues. Indicators should have a set of thresholds or limits with an escalation structure attached to each threshold level. Those indicators should be used to support risk assessments and provide a way to track risk exposures between full updates of the risk assessment process. To be as efficient as possible, those indicators should be built intelligently. For instance, they should ideally predict what is going to happen rather than simply infer that something is changing. They should also be easy to monitor: a simple and cost effective collection should be coupled with an easy understanding of the indicator. This also implies that the KRIs needs to be capable of being measured with a high level of certainty and on a repeated basis.

Fixing risk appetite statement and alert on KRI indicators is requiring a deep involvement of management and business at inception. But it will prove its value by allowing them to be more systematically alerted when risk are falling outside limits and to take decisions around those risks.

4 Conclusion

Sun Tzu said “Invincibility lies in the defence”. As long as people, systems and processes remain imperfect, operational risk cannot be fully eliminated. Identify and implement efficient barriers and mitigating actions are crucial components of the operational robustness and firm strength resilience. Operational risk should not be viewed as an afterthought, but as an integral part of the strategic planning, business management and enterprise risk management processes.

The implementation of a strong operational risk framework is challenging. It requires management sponsorship and buy-in, appropriately skilled resources to challenge and advice the business and robust and organized processes. It also rests upon figures that help the management to evaluate the exposure of the company to its operational risks and therefore prioritize their actions. The accuracy of those figures is ensured by the use of a coherent and robust internal model for quantifying operational risks. This thesis will now focus on the quantification part of the operational risk framework.
CHAPTER 5
MODELLING STANDALONE SCENARIOS

Now that the principles and methodology of a Scenario Based Approach internal model have been introduced, this thesis will focus on the topic of quantification. As a first step, the quantification of scenarios will be presented on a standalone basis. This chapter introduces the standalone modelling of Operational risks by a frequency-severity approach. Then, it presents the distributions used for fitting purposes. The specificity of this modelling comes from the fact that the distributions are fitted on three assessed points. Therefore, the calibration exercise is very specific to the Scenario Based Approach model, because classical parameter estimation techniques are not consistent with such approach.

Statistical estimation involves the distribution of the data and typically uses method of moments or maximum likelihood functions. Indeed, two important conditions to the use of statistical estimation are that:

- the parameter values are assumed not to change quickly in time;
- we have access to sufficient (recent) historical data on the underlying process to provide reliable estimates of these parameter values\(^\text{21}\).

Unfortunately, the Operational risk profile of a company evolves a lot along with time. For instance, an entity might be highly exposed to the risk of a failure of the management of a project for the three years preceding its setting up. As soon as the implementation is finished, the risk does not exist anymore and therefore the potential risk of a failure in the project management that accounted for hundreds of millions of euros falls to zero. Moreover, the question of the availability of sufficient recent historical data is definitely in doubt when it comes to operational risks of an insurance company.

In this context, this chapter presents alternative techniques for the calibration of the chosen distributions, based on quantile-matching methods. In addition, the chapter suggest the use of an additional parameter in the calibration exercise in order to reduce the discrepancies of the model, by minimizing the inter-quantiles distance.

1 Theory: From the risk assessment to the standalone quantification

For each individual Operational Risk, a Loss Frequency Distribution and a Loss Severity Distribution are defined. Both distributions’ parameters are estimated from the risk assessment’s outputs. Both

\(^{21}\) MCLEISH DON L., Computer Intensive Methods for Stochastic Models in Finance, Chapter 7, 2005
distributions are then mixed through Monte Carlo simulations to obtain an estimated Loss Distribution of the considered risk.

Operational Risk quantification methodology is based on local Operational Risk assessments processed within each operating entity. The risk assessment exercise consists in defining the frequency of the event and the related severity. A common method is effectively to decompose the loss distribution into its two integral components: frequency and severity. Under this actuarial approach, frequency represents the number of events and severity represents the loss magnitude per event.

Therefore,

$$X = \sum_{i=1}^{N} C_i$$

where $N$ is the number of events that occur within a year

$C_i$ is the amount of each single loss event. The losses are assumed to be independent identically distributed (iid).

Moreover, $N$ is assumed to be independent of $C_i$

Combination of Poisson and lognormal is generally considered by financial institutions as a standard market practice in terms of Operational risk modelling. Characteristics of those distributions are in line with standard assumptions regarding Operational risks. Additional distributions are used to cope with specific risk situations: for high frequency risks using historical data, a Negative Binomial frequency distribution is used, for scenarios that could not happen more than once a year Bernoulli distribution is applied. Sensitivity testing is produced to fully capture and understand the behaviour of the distributions.

The annual loss is then computed using Monte Carlo simulations.\(^{22}\)

## 2 Calibration

One point for the frequency and three points for the severity are used to calibrate distribution parameters. This section aims to describe the calibration of the distribution.

### 2.1 Frequency distribution

Frequency refers to the number of events that occur within a given time period.

---

\(^{22}\) Panjer recursion and Fourier transformation methods are also very utilized in the operational risk theory but will not be detailed in this paper.
2.1.1 Poisson
The Poisson distribution is a discrete distribution used for modelling the number of events that occur within a year. The parameter of this distribution $\lambda$ is equal to the mean and the variance of the distribution. It corresponds to the average number of losses expected to happen within a year, and can be directly estimated from the risk assessment. Thus, it is particularly appropriate to retain this distribution when the observed mean and variance of the occurrence of the losses are relatively close. In the Supervisory Guidelines for the AMA\textsuperscript{23}, the Basel Committee indicates that “it is common for banks to use the Poisson distribution for estimating frequency”. However, other types are distributions can be chosen, especially when the mean is significantly different from the variance.

2.1.2 Bernoulli
This distribution is useful to model risks that cannot occur more than once in a year. For instance, this distribution can be appropriate for modelling the risk of an erroneous tax reporting, as it can only be detected once a year by the regulator and lead to financial sanctions for misstatement. The only parameter of this distribution $p$ corresponds to the probability of this event to occur.

2.1.3 Binomial
The binomial distribution constitutes an appropriate choice if the frequency variance is smaller than the mean. The latter is used to model the frequency of extreme events where there is relatively little variation and where the risk will not occur more than $n$ times within a year. The Binomial distribution correspond to the sum of an independent Bernoulli distributed variables, with a probability of occurrence of $p$. Those parameters can be estimated with the risk assessment.

2.1.4 Negative Binomial
The Negative Binomial has two parameters and is thus more flexible than the Poisson distribution. The Negative Binomial distribution is an appropriate choice if the frequency variance is larger than the mean which is often seen in empirical data. The latter is used to model the frequencies subject to a high degree of variability.

2.1.5 Fixed frequency
The fixed frequency constitutes an appropriate choice if the frequency of the event is well known. With fixed frequency the mean is equal to a constant with a variance of zero. If the fixed frequency is setup at 1, then the annual loss is only determined by the severity distribution.

\textsuperscript{23} Bank for International Settlement, operational Risk – Supervisory Guidelines for the Advances Measurement Approaches, article 11 (g), June 2011
2.2 Severity distribution

Severity estimation consists in giving the amount of an incurred loss. As mentioned on the Supervisory Guidelines for an AMA\textsuperscript{24}, the SBA model curves are predetermined and the scenario data are used only to estimate the parameters of those distributions, usually by percentile matching.

In our study, let’s consider the three assessed points being:

- Typical: considered as the mode of the severity distribution function (the loss that will occur most of the time)
- Serious: corresponds to the $\beta$-quantile (the worst loss out of $\frac{1}{1-\beta}$ losses)
- Extreme: corresponds to the $\gamma$-quantile (the worst loss out of $\frac{1}{1-\gamma}$ losses)

In order to find the relationship between the distribution parameters and those three assessed points, we need to have the analytical formulas for the mode and the quantiles of the distributions.

2.2.1 Lognormal

Modelling severity is the most challenging aspect especially regarding the lack of sufficient data. Lognormal, as a heavy tailed distribution, is particularly appropriate to model low frequency/high severity events. Lognormal is therefore very popular in modelling applications when the variable is skewed to the right.\textsuperscript{25}

Let $X$ be a random variable following a lognormal distribution. That is: $X = e^Y : Y \sim N(\mu, \sigma^2)$.

It is equivalent to: $\frac{\ln(X) - \mu}{\sigma} = \frac{Y - \mu}{\sigma} \sim N(0,1)$

Thus, we have the relationship between $X$ and its quantiles given by:

$$P\left(\frac{\ln(X) - \mu}{\sigma} < q_{\alpha}^{N(0,1)}\right) = \alpha \Leftrightarrow P\left(X < e^{\alpha q_{\sigma}^{N(0,1)} + \mu}\right) = \alpha$$

Moreover, the assessed impacts are such that $P(X < \text{serious}) = \beta$ and $P(X < \text{extreme}) = \gamma$

Finally, it comes: $\text{serious} = e^{\sigma q_{\beta}^{N(0,1)} + \mu}$ and: $\text{extreme} = e^{\sigma q_{\gamma}^{N(0,1)} + \mu}$

In order to obtain the relationship between $X$ and its mode, let’s calculate its density function using the normal density:

\textsuperscript{24} Bank for International Settlement, operational Risk – Supervisory Guidelines for the Advances Measurement Approaches, article 209, June 2011
\textsuperscript{25} Bank for International Settlement, operational Risk – Supervisory Guidelines for the Advances Measurement Approaches, article 210
\[ \mathbb{E}(h(X)) = \mathbb{E}(h(e^Y)) = \int_{-\infty}^{+\infty} h(e^y) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy = \int_{0}^{+\infty} h(x) \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln(x)-\mu)^2} \times \frac{dx}{x} \]

Thus the density of the lognormal is: \( f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(\ln(x)-\mu)^2} \mathbb{1}[x > 0] \).

The mode of the distribution is the maximum of this function. A condition for the mode being the maximum of the density function is \( f'(\text{mode}) = 0 \).

Yet \( f'(x) = \frac{1}{\sigma \sqrt{2\pi}} \left[ -\frac{1}{x^2} e^{\frac{1}{2\sigma^2}(\ln(x)-\mu)^2} + \frac{1}{x} \left( -\frac{1}{2\sigma^2} \right) \left( \frac{2\ln(x)}{x} - \frac{2\mu}{x} \right) e^{-\frac{1}{2\sigma^2}(\ln(x)-\mu)^2} \right] = 0 \)

Then \( f'(\text{mode}) = 0 \Leftrightarrow f(\text{mode}) \left[ 1 + \frac{\ln(\text{mode}) - \mu}{\sigma^2} \right] = 0 \Leftrightarrow \ln(\text{mode}) = \mu - \sigma^2 \Leftrightarrow \text{mode} = e^{\mu - \sigma^2} \)

In the end, the three equations that are used for the lognormal calibration are the following:

\[
\begin{align*}
\text{mode} &= e^{\mu - \sigma^2} \\
\text{serious} &= e^{\sigma q_0^{N(0,1)} + \mu} \\
\text{extreme} &= e^{\sigma q_2^{N(0,1)} + \mu}
\end{align*}
\]

The estimation of two variables (\( \mu \) and \( \sigma \)) needs two of the above equations. For instance, by using only the mode and the extreme quantile, the calibration gives:

\[
\begin{align*}
\ln(\text{mode}) &= \mu - \sigma^2 \\
\ln(\text{extreme}) &= \sigma q_2^{N(0,1)} + \mu \\
\sigma \Leftrightarrow & \left\{ \begin{array}{l}
\sigma^2 + \sigma q_2^{N(0,1)} + \ln \left( \frac{\text{mode}}{\text{extreme}} \right) = 0 \\
\mu = \sigma^2 + \ln(\text{mode})
\end{array} \right.
\]

The roots of the polynomial are given by:

\[ \sigma = \frac{-q_2^{N(0,1)} \pm \sqrt{\Delta}}{2} \]

where \( \Delta = \left( q_2^{N(0,1)} \right)^2 - 4 \ln \left( \frac{\text{mode}}{\text{extreme}} \right) \)

Thus

\[
\begin{align*}
\mu &= \left( \frac{-q_2^{N(0,1)} + \sqrt{\left( q_2^{N(0,1)} \right)^2 - 4 \ln \left( \frac{\text{mode}}{\text{extreme}} \right)}}{2} \right)^2 + \ln(\text{mode}) \\
\sigma &= \frac{-q_2^{N(0,1)} + \sqrt{\left( q_2^{N(0,1)} \right)^2 - 4 \ln \left( \frac{\text{mode}}{\text{extreme}} \right)}}{2}
\end{align*}
\]

To allow a better use of the experts’ inputs, a calibration with the three points collected is encouraged. For that, a choice has to be made for the model, taking into account the fact that the mode should be the most reliable assessment. Thus, it has been decided to integrate a weighting
factor between the Serious and Extreme impacts. The weighting factor \( a \) measures the weight allocated to the serious impact. Then, the equation used for the calibration is:

\[
\begin{align*}
\mu &= \sigma^2 + \ln(\text{mode}) \\
\frac{\ln(\text{serious}) - \mu}{\sigma} + (1 - a) \frac{\ln(\text{extreme}) - \mu}{\sigma} = a \times q^Y_{\beta} + (1 - a) \times q^Y_{\gamma}
\end{align*}
\]

The second equation gives:

\[
\mu = \sigma \left( -a \times q^Y_{\beta} - (1 - a) \times q^Y_{\gamma} \right) + \ln(\text{serious}) + (1 - a)\ln(\text{extreme})
\]

Then, we solve the second degree equation with respect of \( \sigma \):

\[
0 = \sigma^2 - \sigma \left[ -a \times q^Y_{\beta} - (1 - a) \times q^Y_{\gamma} \right] + \ln(\text{mode}^\text{extreme}) + \ln(\text{extreme}^\text{serious})
\]

It gives

\[
\sigma_{\pm} = \frac{-(-a \times q^Y_{\beta} - (1 - a) \times q^Y_{\gamma}) \pm \sqrt{\Delta}}{2}
\]

where

\[
\Delta = (a \times q^Y_{\beta} + (1 - a) \times q^Y_{\gamma})^2 - 4 \left( \ln(\text{mode}^\text{extreme}) + \ln(\text{extreme}^\text{serious}) \right)
\]

Finally, the lognormal parameters estimated from the experts’ assessment are:

\[
\begin{align*}
\mu &= \sigma^2 + \ln(\text{mode}) \\
-\left( a \times q^Y_{\beta} \times (1 - a) \times q^Y_{\gamma} \right) + \sqrt{(a \times q^Y_{\beta} + (1 - a) \times q^Y_{\gamma})^2 - 4 \left( \ln(\text{mode}^\text{extreme}) + \ln(\text{extreme}^\text{serious}) \right)}
\end{align*}
\]

### 2.2.2 Pareto

The Pareto distribution allows to model fat tails: it is useful for modelling values over a certain threshold.

Let \( X \) be a random variable that follows a Pareto distribution with parameters \( x_{\text{min}} > 0 \) and \( k > 0 \).

The probability density function is:

\[
f(x) = k \frac{x_{\text{min}}^k}{x^{k+1}} \mathbb{I}\{x > x_{\text{min}}\}
\]

Thus:

\[
f'(x) = k x_{\text{min}}^{k-1} \frac{-(k+1)}{x^{k+2}} \mathbb{I}\{x \geq x_{\text{min}}\}
\]

As \( x_{\text{min}} > 0 \) and \( k > 0 \), \( f'(x) < 0 \) for \( x \in [x_{\text{min}}; +\infty[ \). Then, \( \text{mode} = x_{\text{min}} \).

The cumulative distribution function is:
The system of equations is then

\[ F(x) = \mathbb{P}(X \leq x) = \int_{x_{\text{min}}}^{x} k \frac{y^k}{y^{k+1}} dy = 1 - \frac{x_m^k}{x} \]

Thus, \( q_\alpha^p \) is such that \( F(q_\alpha^p) = \alpha \Leftrightarrow 1 - \alpha = \left( \frac{x_m}{q_\alpha^p} \right)^k \Leftrightarrow k \ln \left( \frac{x_m}{q_\alpha^p} \right) = \ln(1 - \alpha) \).

Similarly to the lognormal calibration, a weighting factor is introduced between the serious and extreme impacts. That gives:

\[ a k \ln \left( \frac{x_m}{\text{serious}} \right) + (1 - a) k \ln \left( \frac{x_m}{\text{extreme}} \right) = a \ln(1 - \beta) + (1 - a) \ln(1 - \gamma) \]

In the end, the calibration exercise gives:

\[
\begin{align*}
x_{\text{min}} &= \text{typical} \\
1 - \alpha = & \left( \frac{x_m}{q_\alpha^p} \right)^k \Leftrightarrow k \ln \left( \frac{x_m}{q_\alpha^p} \right) = \ln(1 - \alpha) \]
\]

\[
\begin{cases}
\lambda = \text{typical} \\
1 - \alpha = e^{-\left( \frac{x_m}{\lambda} \right)^k} \Leftrightarrow k \ln \left( \frac{x_m}{\lambda} \right) = \ln(1 - \alpha) \\
1 - \alpha = e^{-\left( \frac{y}{\lambda} \right)^k} \Leftrightarrow k \ln \left( \frac{y}{\lambda} \right) = \ln(1 - \alpha)
\end{cases}
\]

2.2.3 Weibull

The Weibull distribution (with a shape parameter \( k < 1 \)) is part of the class of sub exponential distributions. The sub exponential distributions are those distributions whose tail decays slower than the exponential distribution. In consistency with Basel II recommendations, only cases with a shape parameter \( k < 1 \) should be considered to model Operational Risk.

The density function of the Weibull distribution with parameters \(( \lambda > 0; k > 0 )\) is defined as

\[
f(x) = \begin{cases} 
\lambda^k \frac{x^{k-1}}{\Gamma(k)} e^{-\left( \frac{x}{\lambda} \right)^k} & \text{for } x \geq 0, k > 1 \\
\lambda^k \frac{x^{k-1}}{\Gamma(k)} e^{-\left( \frac{x}{\lambda} \right)} & \text{for } x > 0, k < 1
\end{cases}
\]

In the case we are studying \(( k < 1 )\), the mode of the distribution is not defined. Then, only the quantiles are used for the calibration of the Weibull distribution.

The cumulative distribution function of the Weibull is:

\[
F(x) = \mathbb{P}(X \leq x) = \int_0^x \frac{k}{\lambda} \frac{y^{k-1}}{\Gamma(k)} e^{-\left( \frac{y}{\lambda} \right)^k} dy = 1 - e^{-\left( \frac{x}{\lambda} \right)^k}
\]

Thus, \( q_\alpha^w \) is such that \( F(q_\alpha^w) = \alpha \Leftrightarrow 1 - \alpha = e^{-\left( \frac{q_\alpha^w}{\lambda} \right)^k} \Leftrightarrow q_\alpha^w = \lambda (-\ln(1 - \alpha))^{1/k} \)

The system of equations is then

\[
\begin{align*}
\ln(\text{serious}) &= \ln(\lambda) + \frac{1}{k} \ln(-\ln(1 - \beta)) \\
\ln(\text{extreme}) &= \ln(\lambda) + \frac{1}{k} \ln(-\ln(1 - \gamma))
\end{align*}
\]
It leads to:

\[
\begin{align*}
\lambda &= \frac{\text{serious}}{(\ln(\frac{1}{1-\beta}))^{1/k}} \\
k &= \frac{\ln(\ln(\frac{1}{1-\beta})) - \ln(\ln(\frac{1}{1-\gamma}))}{\ln(\text{serious}) - \ln(\text{extreme})}
\end{align*}
\]

The choice of the distribution should be based on

3 Weighting factor

The distributions chosen for the severity depend on only two parameters. Therefore, the use of two assessed points would allow resolving a system of two equations with two unknowns. Say that the parameters are called $\mu$ and $\sigma$, and the assessed points are mode and quantile. The equation would be:

\[
\begin{align*}
\{ f(\mu, \sigma) &= \text{mode} \\
g(\mu, \sigma) &= \text{quantile}
\end{align*}
\]

In order to allow a better use of the Risk Assessment Questionnaire's inputs, it can be decided to utilize the three assessed points in the calibration exercise: the mode, the $\beta$-quantile and the $\gamma$-quantile. Then, the equation does not have solution as it has three equations with only two unknowns:

\[
\begin{align*}
\{ f(\mu, \sigma) &= \text{mode} \\
g(\mu, \sigma) &= \text{quantile}(\beta) \\
h(\mu, \sigma) &= \text{quantile}(\gamma)
\end{align*}
\]

A weighting factor is then introduced in the lognormal calibration between both the Serious and the Extreme impact. Actually, it measures the weight allocated to the serious value that corresponds to the $\beta$-quantile. Finally, the equation is:

\[
\begin{align*}
\{ f(\mu, \sigma) &= \text{mode} \\
 a \ g(\mu, \sigma) + (1 - a) \ h(\mu, \sigma) &= a \text{quantile}(\beta) + (1 - a) \text{quantile}(\gamma)
\end{align*}
\]

First, this section presents model discrepancies that arise after the calibration exercise. Then, an analysis has been conducted to determine the value of the optimal weighting factor.

3.1 Discrepancies in the model

During the risk assessment exercise, the expert measures three levels of impact. Nevertheless, the lognormal distribution has only two parameters\textsuperscript{26}. Thus, a weighting factor is introduced in order to

\textsuperscript{26} It is also true for Pareto and Weibull distribution but our study will only focus on lognormal distribution, as it is the distribution most widely used for quantifying operational risks.
use the three points assessed by the experts for the calibration exercise. The reverse exercise consists of computing the mode and the quantiles of a lognormal distribution with the parameters that have been calibrated on the risk assessment. Comparing the levels of impact assessed by the expert to the ones computed after calibration enables to back test the accuracy of the model.

Naturally, some information is lost when using only a weighted sum of the assessed impacts. Consequently, it is expected that there will be some discrepancies between the impact levels assessed by the experts and the ones computed by the model after calibration of the lognormal. The distribution of those discrepancies is shown in the figure below.

Figure 6: Discrepancies are inherent to the model
In most cases, the discrepancies are close to zero. Nevertheless the graph shows that for a minority of scenarios, the discrepancies are quite large. The following section presents a possible alternative for the calibration methodology that could reduce the error between the impacts assessed by the expert and the ones calibrated by the model.

3.2 Minimization of the model error

In order to have the best fit between the impacts assessed by the experts and those calibrated by the model, we want to find the weighting factor that allows minimizing the discrepancies between those points. Several approaches can be considered and will depend on the choice of the function to be optimized. Thus, it could be useful to understand the differences between those approaches to retain the one that is most adapted to our problem.

3.2.1 Inter-quantile distance minimization

First, one can choose to minimize the absolute error or the relative error: the relative error enables to reduce the instability due to extreme quantiles in the expression of the distance. Let’s recall how the relative error is calculated: it is the relative distance between the practical impacts (assessed by the expert) and the theoretical ones (computed by the model). That is:

$$relative\ error = \frac{|impact\ calibrated\ by\ the\ model - impact\ assessed\ by\ the\ expert|}{impact\ assessed\ by\ the\ expert}$$

For instance, for the Serious, it is:

$$relative\ error(Serious, a) = \frac{|Serious - e^{\mu(a)} + \sigma(a)q_B|}{Serious}$$
Secondly, various metrics can be retained to calculate the distance. We will use two of them:

- the Euclidean distance: \(d(x, y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}\)
- the Manhattan distance: \(d(x, y) = \sum_{i=1}^{n}|x_i - y_i|\)

### 3.2.2 Functions to be minimized

Below, the 4 possible functions to minimize:

<table>
<thead>
<tr>
<th>Error</th>
<th>Distance</th>
<th>Function to be minimized</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute</td>
<td>Manhattan</td>
<td>(\sqrt{(typical - e^{\mu(a)} - \sigma^2(a))} + (serious - e^{\mu(a) + \sigma(a)q_\beta})^2 + (extreme - e^{\mu(a) + \sigma(a)q_\gamma})^2)</td>
</tr>
<tr>
<td>Absolute</td>
<td>Euclidean</td>
<td>(</td>
</tr>
<tr>
<td>Relative</td>
<td>Manhattan</td>
<td>(\sqrt{(typical - e^{\mu(a)} - \sigma^2(a))} + \frac{(serious - e^{\mu(a) + \sigma(a)q_\beta})^2}{typical} + \frac{(extreme - e^{\mu(a) + \sigma(a)q_\gamma})^2}{extreme})</td>
</tr>
<tr>
<td>Relative</td>
<td>Euclidean</td>
<td>(\frac{</td>
</tr>
</tbody>
</table>

with:

\[
\begin{align*}
\mu(a) &= \ln(typical) + \sigma(a)^2 \\
\sigma(a) &= -a \times q_\beta^Y - (1 - a)q_\gamma^Y + \sqrt{(a \times q_\beta^Y + (1 - a)q_\gamma^Y)^2 - 4\left(\ln\left(\frac{typical}{extreme}\right) + a \times \ln\left(\frac{extreme}{serious}\right)\right)}
\end{align*}
\]

### 3.2.3 Optimization problems

Beside we have \(typical = e^{\mu(a) - \sigma^2(a)}\). Therefore, the optimization problems are:

- Method 1: SSE (Sum of the Square Errors: minimization of the Euclidean distance between the impacts)

\[
\min_{a \in [0; 1]} \text{error}(a) = \sqrt{(serious - e^{\mu(a) + \sigma(a)q_\beta})^2 + (extreme - e^{\mu(a) + \sigma(a)q_\gamma})^2}
\]
Quantification of Operational Risks using a Scenario Based Approach

- Method 2: SAE (Sum Absolute Errors: minimization of the Manhattan distance between the impacts)

\[
\min_{a \in [0;1]} \text{error}(a) = \left| \text{serious} - e^{\mu(a)+\sigma(a)q_\beta^Y} \right| + \left| \text{extreme} - e^{\mu(a)+\sigma(a)q_\gamma^Y} \right|
\]

- Method 3: SSRE\(^{27}\) (Sum of the Square Relative Error: minimization of the Euclidean norm of the discrepancies)\(^{28}\)

\[
\min_{a \in [0;1]} \text{error}(a) = \sqrt{\left( \frac{\text{serious} - e^{\mu(a)+\sigma(a)q_\beta^Y}}{\text{serious}} \right)^2 + \left( \frac{\text{extreme} - e^{\mu(a)+\sigma(a)q_\gamma^Y}}{\text{extreme}} \right)^2}
\]

- Method 4: SAPE\(^{29}\) (Sum Absolute Percentage Error: minimization of the Manhattan norm of the discrepancies)

\[
\min_{a \in [0;1]} \text{error}(a) = \left| \frac{\text{serious} - e^{\mu(a)+\sigma(a)q_\beta^Y}}{\text{serious}} \right| + \left| \frac{\text{extreme} - e^{\mu(a)+\sigma(a)q_\gamma^Y}}{\text{extreme}} \right|
\]

### 3.2.4 Results

To illustrate the optimization problem, the shape of the functions to be optimized is found below. Graphically, we can see that the approximate solutions of those 4 optimization problems are respectively 20%, 0%, 60% and 100%.

---


\(^{28}\) This method can also be called “minimization of the interquartile distance” with weights in the distance being the inverse of the squared quantile (cf. Alexis RENAUDIN, Modèle de capital économique pour le risque opérationnel bancaire: estimation, diversification, ISFA, 2012, pp. 41-47)

\(^{29}\) S. MAKRIDAKIS, M. HIBON, Evaluating accuracy (or error) measures, INSEAD, Fontainebleau, 1995
A numerical approximation was used to solve those optimization problems.

Methods 2 and 4 using the absolute error values lead to the total elimination of one of the three impacts, so that we finally work with only two impacts for the calibration. In order to allow a better use of the assessed points, it looks preferable to focus on the methods using the Euclidean distance and norm.

With the Euclidean metrics, the cases where there are small discrepancies on both the Serious and the Extreme are preferred to the ones where there is a perfect match for one of the impact and a large error for the other impact.

Indeed squaring the terms penalizes large errors. In fact, focusing on the relative error implies tolerating more absolute error between the assessed and the calibrated points for the Extreme than for the Serious. Thus, if one considers the relative error, then one would mechanically give more weight to the extreme values. This approach seems to be the most reasonable. Let’s take the example of a scenario that the expert assessed with a Serious equal to 100 000€ and an Extreme equal to 10 000 000€. With an absolute gap of 10 000€ for both impacts, the relative gap is 10% for the Serious and 0.1% for the Extreme. Of course, an error of 10 000€ is more serious for the Serious impact than for the Extreme impact. By using the method that minimizes the relative error, you give more accuracy to the smaller values and you tolerate more uncertainty on the large values.

The weighting factor minimizing the absolute error is likely to be relatively low: it over weights the Extreme in order to reach the tolerance level for the absolute value the corresponding weighting factor. Conversely, the weighting factor minimizing the relative error is likely to be larger: it gives more weight to the Serious in order to have similar relative errors for both impacts.
4 Conclusion

Fitting a distribution with only three points is delicate. However, those three calibration points are assumed to contain additional information: they give information about their location in the curve of the chosen distribution. Indeed, this makes those points more powerful than “classical” historical data points. Instead of proceeding to classical statistical estimation techniques, they enable to use the theoretical distribution functions and the calibration exercise is reduced to finding the solution of a system of equations that gives closed formulae for the parameters.

Nevertheless, the solution of a system composed of three equations with only two variables is not ensured. The removal of one of the calibration points would lead to the loss of some information. This motivates the introduction of a weighting factor between extreme quantiles: it enables to make better use of the risk assessment exercise. However, if the weighting was an arbitrary percentage set by default for all the risk scenarios, the fit might be poor sometimes. This is evidenced when comparing the mode and the quantiles of the distribution computed with the calibrated parameters to the initial assessed impacts.

Thus, the weighting factor can be considered as a third variable in the calibration exercise where it has to remain in the interval \([0; 1]\). As the weighting factor is the solution of a minimization problem there is no closed formula anymore which entails some additional computational difficulties for the calibration. However, with numerical methods, finding the optimal weighting factor that minimizes the discrepancies is very feasible. In practice, we performed the calibration with all of the above methods. They equally contribute to reducing the model error as they have been designed for different definitions of the error. However, the outcomes are very different, depending on the choice of the error function to be minimized. This uncertainty indicates a lack of robustness of this method and suggests being cautious when choosing one of the above methods. A further analysis on this topic might consist of studying the impact of the choice of one of the above method on the final results. Such study would belong to the sensitivity analyses, presented at the end of this thesis. But before that, we need to understand how the risks interact between one another. This will be discussed in next chapter.
CHAPTER 6
THE CORRELATION BETWEEN RISKS

Now that all the scenarios in the scope of operational risk have been quantified on a standalone basis, as detailed in previous chapter, the insurance company has to evaluate the amount of capital it should hold at aggregate level. Of course, the simple sum of all the standalone SCR is not conceivable. If perfect correlation between all risks of all risk categories was accepted, the capital charge would be way too big for the firm, and a fortiori larger than in the Standard Approach. Yet, regulators are pushing companies to compute their economic capital using internal models in order to better understand quantify and consequently manage their risks. Thus the assumption of a perfect correlation between operational risks would be disincentive and go against regulators’ will.

Therefore, insurance companies have to take into account the dependence that exists between operational risks and introduce this into their internal models. Yet supervisors expect the companies to be able to justify the diversification effect introduced in their internal models. In this context, it is mandatory for insurance companies to use some correlation whose soundness has to be demonstrated with a high level of confidence.

Moreover, supervisors want insurance companies to justify their underlying assumptions, in order to ensure the full adequacy of the internal model. Modelling operational risk with as frequency-severity compound approach implies independence between frequency and severity. This notion being used extensively, supervisors could expect the firm to conduct some studies on this topic in order to check whether it is proper or not.

This section gives an overview on the main dependence measures that are used in operational risk to quantify the degree of correlation between risks from different event categories that occur in different entities. This will be helpful to better understand the aggregation process presented further. Moreover, this chapter describes how the underlying model assumptions of independence can be back-tested in a Scenario Based Approach. The first method consists of considering that the SBA model captures the behaviour of the actual internal loss database. In this sense, the actual database could be used to verify that the assumption of independence between frequency and severity as well as the iid hypothesis are reasonable. Another method consist of considering the sets of parameters and perform tests on those sets to determine if the frequency parameters and independent from the impact parameters.

1 Measure of dependence

In statistics, dependence is any statistical relationship between two random variables. Formally, it refers to any situation in which random variables do not satisfy a mathematical condition of probabilistic independence. Random variables \((X_1, \ldots, X_n)\) are said to be independent if and only if the elements of the \(\pi\)-system generated by them are independent. In other terms, \(\forall (a_1, \ldots, a_n) \in \mathbb{R}^n : P(\bigcap_{i=1}^{n}(X_i < a_i)) = \prod_{i=1}^{n} P(X_i < a_i).\)
Increasingly, correlation is being used as a dependence measure in risk management. This section aims to highlight the existence of situations for which finding an alternative to the linear correlation as an all-purpose dependence measure becomes necessary. Those alternative options are used to capture non-linear relationships that exist between many real world risk factors, especially in the operational risk framework. The three basic types of measures are: linear correlation, rank correlation and tail dependence.

1.1 Linear correlation

By far, the most familiar dependence concept is the Pearson correlation coefficient that measures the linear correlation between a pair of variables \((X, Y)\). It is defined as:

\[
\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}
\]

where \(\mu_X\) and \(\mu_Y\) are the expected values and \(\sigma_X, \sigma_Y > 0\) denote the standard deviation of \(X\) and \(Y\) respectively. This definition can be extended to the multivariate case by introducing a symmetric positive definite matrix. It is well known that:

- \(\rho\) is a measure of linear dependence
- \(\rho\) is symmetric: \(\rho(X, Y) = \rho(Y, X)\)
- \(-1 \leq \rho \leq 1\)
- \(\rho\) is invariant with respect to linear transformations of the variables

Moreover, if \((X, Y)\) follows a bivariate normal distribution, then the correlation is fully informative about their joint dependence, and \(\rho(X, Y) = 0\) is equivalent to the independence of \(X\) and \(Y\).

However, two independent variables have a Pearson’s correlation coefficient equal to 0, but the converse is not true because the correlation coefficient detects only linear dependencies between two variables. This is a first limitation of this dependence measure that is highlighted by a famous example, when \(X \sim \mathcal{N}(0, 1)\) and \(Y = X^2\). In this case, \(X\) and \(Y\) are clearly dependent but

\[
\text{Cov}(X, Y) = \mathbb{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}(X^3) - (\mathbb{E}(X))^2 = 0.
\]

Thus, zero dependence does not require \(\text{Cov}(X, Y) = 0\) but \(\text{Cov}\left(f(X), g(Y)\right) = 0\) for any functions \(f\) and \(g\). Another weakness comes from the fact that it is only definite for variables that have finite second moments. Indeed, it is not the case for some heavy-tailed distributions, which are the ones that we are concerned with regarding operational risk. The third limitation of that correlation coefficient is that it is not invariant under strictly increasing nonlinear transformations. Given those limitations, alternative measures of correlations should be considered. Thus, other correlation coefficients have been developed to be more robust than the Pearson linear correlation coefficient. Next section will introduce the concept of rank correlations that are more sensitive to nonlinear relationships.
1.2 Rank correlation coefficients

In statistics, a rank correlation is any of several statistics that measure the similarity of the orderings of the data. A rank correlation coefficient measures the degree of similarity between two orderings, and can be used to assess the significance of the relation between them.

There are situations where measuring the correlation on the values is not appropriate. If the variables are ordinal or discrete, or if the results can be biased by outliers, it can be interesting to focus on the rank correlations. In this case, the values of the observations are discarded and only their rank is observed. With this kind of method, some information is lost and only enables to know if there exists a monotonic link between the variables, whether it is linear or non-linear. There are different techniques hence several correlation coefficients to validate the results. The most common are the Spearman’s rho and the Kendall’s tau. Both measures are based on the concept of concordance, which refers to the property that large values of one random variable are associated with large values of another, whereas discordance refers to large values of one being associated with small value of the other. By turning to rank correlations, certain of the theoretical limitations of standard linear correlation can be repaired.

1.2.1 Spearman’s rho

Fundamentally, Spearman’s rho is a special case of Pearson’s coefficient, computed from transformations of the original variables. The idea is to replace the values of the observations by their rank. Let \( \text{rank}(X) = (\text{rank}(x_i))_{i=1,n} \) and \( \text{rank}(Y) = (\text{rank}(y_i))_{i=1,n} \) be the vectors of the ranks of \( X \) and \( Y \). It means that \( \text{rank}(x_i) \) (resp. \( \text{rank}(y_i) \)) is the rank of the observation \( x_i \) (resp. \( y_i \)) in column \( X \) (resp. \( Y \)). The Spearman’s correlation coefficient is defined as:

\[
\rho_S(X, Y) = \rho(\text{rank}(X), \text{rank}(Y)) = \frac{\text{Cov}(\text{rank}(X), \text{rank}(Y))}{\sigma_{\text{rank}(X)} \sigma_{\text{rank}(Y)}}
\]

When the variables are both continuous, there is a simplified formula that comes from the fact that \( \overline{\text{rank}(X)} = \overline{\text{rank}(Y)} = \frac{\sum_{k=1}^{n} k}{n^2} = \frac{n(n+1)}{2} \). In that case, we have:

\[
\hat{\rho}_S = 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n^3 - n}
\]

where \( d_i = \text{rank}(x_i) - \text{rank}(y_i) \).

However, special attention must be put in the cases where the number of ex-aequo is large in the sample. Indeed, average rank is given to the observations that have the same value. The adjustment is as sensible as the number of identical values is high for \( X \) and \( Y \).

As a special case of Pearson’s coefficient, Spearman’s rho has the same main properties, namely \(-1 \leq \rho_S \leq 1\), as well as \( \rho_S(X, Y) = 0 \) when \( X \) and \( Y \) are independent. Moreover, it is a non-
parametric measure, therefore it does not need to make assumptions on the distributions of $X$ and $Y$. Still, when the couple $(X, Y)$ follows a normal bivariate distribution, it is almost as strong as the Pearson’s coefficient which entails that the same test of significance can be applied.

In practice, Spearman’s rho is very useful because it enables to capture any non-linear monotonic relationship between two variables, which avoids having to make a choice about the transformation function that would enable to have a linear relationship between those two variables. In general, significant difference between $\rho$ and $\rho_S$ should alert on the possible non-linear relationship that exists between the variables. Furthermore, this coefficient is less influenced by outliers than the Pearson’s coefficient that could be highly affected by the existence of an atypical point.

### 1.2.2 Kendall’s tau

Kendall’s correlation coefficient is not a proper variant of Pearson’s coefficient and therefore has its own principles. It is defined to measure the association between two variables and lies on the concept of concordant and discordant pairs:

- a pair of observations $i$ and $j$ is said to be concordant if and only if:
  \[(x_i > x_j \text{ and } y_i > y_j) \text{ or } (x_i < x_j \text{ and } y_i < y_j)\].

- a pair of observations $i$ and $j$ is said to be discordant if and only if:
  \[(x_i > x_j \text{ and } y_i < y_j) \text{ or } (x_i < x_j \text{ and } y_i > y_j)\].

For a sample of size $n$, let $C$ be the number of concordant pairs and $D$ the number of discordant pairs, then:

\[
\hat{\tau} = \frac{C - D}{\frac{1}{2} n(n - 1)}
\]

where the denominator is the total number of pairs, i.e.: $\frac{1}{2} n(n - 1) = \binom{n}{2}$.

Again, specific there is a specific formula in the case of the existence of ties in the sample. If $x_i = x_j$ or $y_i = y_j$ or both, the comparison is called a tie and is not counted as concordant or discordant. If there are a large number of ties, then the denominator has to be replaced by:

\[
\sqrt{\left(\binom{n}{2} - n_X\right)\left(\binom{n}{2} - n_Y\right)}
\]

where for each value $x_m$ of $X$, we introduce the number of occurrences $\text{nb}(x_m)$ such that $n_X = \sum_{m=1}^{n_X} \text{nb}(x_m)(\text{nb}(x_m) - 1)$ is the number of ties involving $X$ and $n_Y$ the ones involving $Y$.

In fact, the main different with Spearman’s coefficient is that Kendall’s correlation coefficient can be read as a probability. Hence, when $\tau = 0$, a pair of observations has as many chances to be concordant ou discordant. Thus, the Kendall’s tau is defined as:
\( \tau = \mathbb{P}(\text{concordant}) - \mathbb{P}(\text{discordant}) \)

### 1.2.3 Goodman and Kruskal’s Gamma coefficient

It can happen that the number of ties is high in the sample, especially when the values are integer. For instance, when studying the number of occurrence of an event modelled with a low frequency parameter, most of the Monte Carlo outcomes are likely to be equal to zero and one. In this case, the Goodman and Kruskal’s gamma is given as a measure of association that is highly resistant to tied data. It is a measure of rank correlation that measures the degree of association of the cross tabulated data with ordinal variables. Similarly to Kendall’s tau, Gamma coefficient uses the notion of concordant and discordant pairs and is defined as follows:

\[
Y = \frac{C - D}{C + D}
\]

### 1.2.4 Relationship between Pearson’s, Spearman’s and Kendall’s correlation coefficients

Both Spearman’s rho and Kendall’s tau are rank correlations. They lie on the same assumptions and use the same information. It is then straightforward that they have similar power (i.e. capacity to detect \( H_0 \) when the null hypothesis is verified). The main difference is in the interpretation of those figures: \( \rho^2 \) can be understood as a percentage of variance explained whereas \( \tau \) is a probability.

Nevertheless, it can be shown\(^{31}\) that:

\[-1 \leq 3\hat{\tau} - 2\hat{\rho}_S \leq 1\]

Also, when \( n \) is large enough and the coefficients are not too close to 1 or -1, we have\(^{32}\):

\[\hat{\rho}_S \approx \frac{3}{2}\hat{\tau}\]

Finally, when \((X, Y)\) follows a bivariate normal distribution (and also for all other elliptical copulas\(^{33}\)), we have interesting results that gives relation between those three correlation coefficients:

\[
\mathbb{E}(\hat{\rho}) = \rho \left[ 1 - \frac{1 - \rho^2}{2n} + O(n^{-2}) \right], \quad \text{which entails} \quad \hat{\rho} \xrightarrow{n \to \infty} \rho
\]

\[
\mathbb{E}(\hat{\rho}_S) = \frac{6}{\pi(n+1)} \left[ \sin^{-1} \rho + (n - 2) \sin^{-1} \frac{\rho}{2} \right], \quad \text{which entails} \quad \rho_S \xrightarrow{n \to \infty} \frac{6}{\pi} \sin^{-1} \frac{\rho}{2}
\]

\[
\mathbb{E}(\hat{\tau}) = \frac{2}{\pi} \sin^{-1} \rho
\]

---


\(^{32}\)Sergei AÏVAZIAN, Etude statistique des dépendances, Mir, Moscou, 1978 (p.114)

\(^{33}\)Normal and Student-t copula
The first result was given by Hotelling\(^{34}\) (1953), the second by Moran\(^{35}\) (1948), and the third one by Esscher\(^{36}\) (1954), also known as Greiner’s relation \(^{37}\).

Moreover, we have \(\sin^{-1} x \sim x\) when \(x \in \left[-\frac{1}{2}, \frac{1}{2}\right]\). Hence, as \(x \in [-1, 1]\), we have:

\[
\rho_S = \frac{6}{\pi} \sin^{-1} \frac{\rho}{2} \sim \frac{3\rho}{\pi}
\]

![Figure 9: Moran’s relation between Spearman’s \(\rho_S\) and Pearson’s \(\rho\) (for bivariate normal distributions)](image)


\[^{37}\] Maurice KENDALL, Jean DICKINSON GIBBONS, Rank Correlation Methods, Oxford university, September 1990.
Figure 10: Greiner’s relation between Kendall’s $\tau$ and Pearson’s $\rho$ (for bivariate normal distributions)

In order to illustrate those relations for bivariate normal distributions, the table below gives the estimated correlation coefficients, computed on a set of bivariate normal distribution with linear correlations ranging from $-1$ to $+1$. The table shows that, as expected, Spearman’s rank correlation coefficient is very close to Pearson’s. Moreover, as a result of the relation between Kendall’s and Pearson’s coefficients in addition to the proximity of Spearman’s and Pearson’s, Spearman’s rho is greater than Kendall’s tau. For instance, a Kendall’s $\tau$ of 59% could be transformed, using Greiner’s relation, to an “equivalent” Pearson correlation of 80%.

<table>
<thead>
<tr>
<th>Actual correlation</th>
<th>Pearson’s linear correlation $\rho$</th>
<th>Kendall’s rank correlation $\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ns=10</td>
<td>$\rho$</td>
<td>$\tau$</td>
</tr>
<tr>
<td>ns=100</td>
<td>$\rho$</td>
<td>$\tau$</td>
</tr>
<tr>
<td>ns=1000</td>
<td>$\rho$</td>
<td>$\tau$</td>
</tr>
<tr>
<td>ns=10000</td>
<td>$\rho$</td>
<td>$\tau$</td>
</tr>
</tbody>
</table>

The dependence measure chosen in this model for aggregation purpose is the Spearman rank correlation. The reason of using rank correlations instead of linear correlations is that linear correlations are not robust in case of non-elliptical distributions, which is the case for operational risk. In addition, it is very convenient, because Spearman’s rho is understandable and easily computable.
1.3 Coefficient of tail dependence

In some cases, the concordance between extreme values of random variables is of interest. For example, one may be interested in the probability that fraud losses in two countries exceed given levels. This requires a dependence measure for upper and lower tails of the distribution.

The concept of tail dependence provides, roughly speaking, a measure for extreme co-movements in the lower and upper tail of the joint distribution, that is to say between less probable values of variables. Such a dependence measure is essentially related to the conditional probability that one risk exceeds some value given that another exceeds that value.

Let \( X_1, X_2 \) be two random variables with marginal distributions given by \( F_1(\cdot), F_2(\cdot) \) and let \( u \) be the threshold value. Then, the upper tail coefficient \( \lambda_U \) is defined as:

\[
\lambda_U = \lim_{u \to 1^-} \mathbb{P}(F_1(X_1) > u \mid F_2(X_2) > u) = \lim_{u \to 1^-} \mathbb{P}(U_1 > u \mid U_2 > u)
\]

In a similar way, the lower tail coefficient is defined as:

\[
\lambda_L = \lim_{u \to 1^-} \mathbb{P}(F_1(X_1) \leq u \mid F_2(X_2) \leq u) = \lim_{u \to 1^-} \mathbb{P}(U_1 \leq u \mid U_2 \leq u)
\]

In practice, the coefficient of tail dependence is useful when the risks are aggregated with copulas. We will not go through this topic which has been largely treated in the extreme value theory. However, it is important to understand that there are improvements to measuring dependency with one single number: copulas contain the dependence structure of a joint distribution. Therefore, they can be used to detect specific dependence structures such as tail dependence. Different copulas imply different tail dependence structures.

Nevertheless, the implementation of the copula theory for aggregation purpose is not straightforward, especially in a SBA. Indeed, the choice of the copula would be fully based on its desired properties (e.g. tail dependence) because the parameters cannot be estimated with internal data. Moreover, because of the high technical skills needed to understand the copula theory, the estimation of its parameters using expert judgment is hardly conceivable. However, the normal copulas could be considered as a first approach because they are characterised by a correlation matrix that can be built on expert judgment. As it will be detailed in next chapter, the aggregation approach is based on the Normal copula approach: the aggregation consists of matching the ranks of a multivariate Gaussian vector with the ranks of the objects to be aggregated that have been previously modelled with a frequency-severity method.

2 Independence between frequency and severity

Now that we have presented the concept and the measures of dependence between variables, we can present the independence tests that could be performed in an Operational Risk framework. First, the theory on different test of independence is described. Secondly, we will see how in which extent it can be applied to the Operational Risks and try to decide which type of test is more appropriate in our studies.
2.1 Pearson’s chi-squared test

To test the hypothesis of independence between frequency and severity, the common practice is to use a Chi-squared test for independence: the Pearson’s chi-squared test. It is a statistical test applied to sets of categorical data to evaluate how likely it is that any observed difference between the sets arose by chance. In this test, an observation consists of the value of two outcomes and the null hypothesis is that the occurrence of these outcomes is statistically independent, i.e. the row variable is independent of the column variable. Each observation is allocated to one cell of a contingency table. This test of independence assesses whether the paired observations are independent of each other.

The value of the test statistic is:

\[
\chi^2 = \sum_{i=1}^{nrow} \sum_{j=1}^{ncol} \frac{(O_{i,j} - E_{i,j})^2}{E_{i,j}}
\]

where \( nrow, ncol \) are the number of rows and columns in the contingency table

\( O_{i,j} \) is the observed frequency

\( E_{i,j} \) is the expected frequency, that is:

\[
E_{i,j} = \frac{1}{nrow \times ncol} \left( \sum_{k=1}^{ncol} O_{i,k} \right) \left( \sum_{m=1}^{nrow} O_{i,m} \right)
\]

Under the null hypothesis, the test statistic is following a chi-squared distribution with \( nrow \times ncol - (nrow + ncol - 1) \) degrees of freedom.

In practice, the Pearson’s chi-squared test is designed for categorical variables. When working with discrete or continuous variable, we have to build classes. Special attention must be put in the determination of those classes distributions when building the contingency table. For instance, it is expected that frequencies would range in significantly different classes depending on the entity because all entities do not have the same risk profile. Moreover, the levels of impacts can be very different across operational risk categories. Hence, when conducting chi-squared tests on several entities and risk categories, the classes for the frequency and the impact should be designed so that they correspond in some ways with the case studied. An example of the sensitivity of the test outcome to the choice of the classes is provided in the figure below:

**Figure 11**: Chi-squared test applied to the same sample with different intervals for Y. The test performed on the left accepts the null hypothesis of independence whereas the one performed on the left rejects the hypothesis of independence with a confidence level of 1%.
Drawing on a conclusion on the inexistence of dependence of two variables based on a Chi-squared test would be highly imprudent. In addition to the Chi-squared test whose outcomes are very sensitive to the choices of the classes, non-parametric statistical tests could also be considered for testing the independence.

### 2.2 Significance tests

The first test that comes in mind is the significance of the correlation level. The idea is to answer the question: “is the correlation coefficient significantly different of 0?” As several measures of dependence have been defined, the corresponding tests for those coefficients are given below.

#### 2.2.1 Pearson’s correlation significance test

As it has been detailed above Pearson’s correlation coefficient is:

\[ \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\mathbb{E}[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \]

Under the null hypothesis of independence of \(X\) and \(Y\), the test statistic \(t = \rho \frac{n-2}{\sqrt{1-\rho^2}}\) follows a Student distribution with \(n - 2\) degree of freedom\(^{38}\). Thus, for a chosen confidence level \(\alpha\), we compare \(t\) to \(t_{1-\frac{\alpha}{2}}(n - 2)\) which is the \((1 - \frac{\alpha}{2})\)-quantile of the Student distribution with \(n - 2\) degrees of freedom, that can be read on a Student table. If \(t \leq t_{1-\frac{\alpha}{2}}(n - 2)\), then the null hypothesis \(H_0\) is accepted. If \(t > t_{1-\frac{\alpha}{2}}(n - 2)\), then \(H_0\) is rejected.

#### 2.2.2 Spearman rank correlation coefficient test

As it has been detailed above, Spearman’s rho is defined as Pearson’s correlation coefficient between the ranks of the variable, that is:

\[ \rho = \frac{\text{Cov}(\text{rank}(X), \text{rank}(Y))}{\sigma(\text{rank}(X)) \sigma(\text{rank}(Y))} \]

For a sample of size \(n\), the observations \((X_i, Y_i)\) are converted to ranks denoted \((r(X)_i, r(Y)_i)\) where \(r(X)\) is the mean average in the case of ties in the sample, and the Spearman rho is

\[ \hat{\rho} = \frac{\sum_i (r(X)_i - \bar{r}(X)_n)(r(Y)_i - \bar{r}(Y)_n)}{\sqrt{\sum_i (r(X)_i - \bar{r}(X)_n)^2 \sum_i (r(Y)_i - \bar{r}(Y)_n)^2}} \]

Under the null hypothesis of independence of \(X\) and \(Y\), the test statistic \(t = \rho \frac{n-2}{\sqrt{1-\rho^2}}\) follows a Student distribution with \(n - 2\) degree of freedom\(^{39}\). Thus, for a chosen confidence level \(\alpha\), we compare \(t\) to \(t_{1-\frac{\alpha}{2}}(n - 2)\) which is the \((1 - \frac{\alpha}{2})\)-quantile of the Student distribution with \(n - 2\)

\(^{38}\) [link](http://w3.mathinfolmd.univ-tlse2.fr/membres/chabriac/PY0010/PY0010_ch6.PDF) (consulted on July 26\(^{th}\) 2014)

\(^{39}\) [link](http://w3.mathinfolmd.univ-tlse2.fr/membres/chabriac/PY0010/PY0010_ch6.PDF) (consulted on July 26\(^{th}\) 2014)
degrees of freedom, that can be read on a Student table. If \( t \leq t_{1-\frac{\alpha}{2}}(n-2) \), then the null hypothesis \( H_0 \) is accepted. If \( t > t_{1-\frac{\alpha}{2}}(n-2) \), then \( H_0 \) is rejected.

### 2.2.3 Kendall rank correlation coefficient test

The Kendall rank correlation coefficient test is a non-parametric hypothesis test for statistical dependence based on the Kendall tau that is a measure of rank correlation. Let \((x_1, y_1), \ldots, (x_n, y_n)\) be a set of observations of the joint random variables \(X\) and \(Y\).

- if \( \{ \begin{align*} &x_i > x_j \text{ and } y_i > y_j \\ &x_i < x_j \text{ and } y_i < y_j \end{align*} \}
  \) then the observations \((x_i, y_i)\) and \((x_j, y_j)\) are concordant
- if \( \{ \begin{align*} &x_i > x_j \text{ and } y_i < y_j \\ &x_i < x_j \text{ and } y_i > y_j \end{align*} \}
  \) then the observations \((x_i, y_i)\) and \((x_j, y_j)\) are discordant

The Kendall tau coefficient is:

\[
\hat{\tau} = \frac{C - D}{\frac{1}{2} n(n - 1)}
\]

where \(C\) is the number of concordant pairs and \(D\) the number of discordant pairs.

Under the null hypothesis of independence of \(X\) and \(Y\), we have\(^{40}\):

\[
\frac{\hat{\tau} - \mathbb{E}_{H_0}(\hat{\tau})}{\sqrt{\text{Var}(\hat{\tau})}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)
\]

where \(\lim_{n \to +\infty} \mathbb{E}_{H_0}(\hat{\tau}) = 0\) and \(\lim_{n \to +\infty} \text{Var}(\hat{\tau}) = \frac{2(2n+5)}{9n(n-1)}\).

As long as \(n > 8\), it is possible to use the assumption that \(\hat{\tau}\)\(^{41}\) converges in distribution towards a normal distribution under the null hypothesis \(H_0\). Thus, under \(H_0\), the test statistic

\[
t = \frac{\hat{\tau}}{\sqrt{\text{Var}(\hat{\tau})}} = \frac{\hat{\tau}}{\frac{2(2n+5)}{9n(n-1)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)
\]

\(^{40}\) [http://stat.genopole.cnrs.fr/_media/members/cmatias/cours_stat_no_ensie.pdf](http://stat.genopole.cnrs.fr/_media/members/cmatias/cours_stat_no_ensie.pdf)

Catherine MATIAS, Introduction à la statistique non paramétrique, CNRS, Laboratoire Statistique & Génome, Evry, 2013/2014

\(^{41}\) AÏVAZIAN, S., Etude statistique des dépendances, Mir, Moscou, 1978
Thus, for a chosen confidence level $\alpha$, we compare $t$ to $q_{1-\frac{\alpha}{2}}$ which is the $(1 - \frac{\alpha}{2})$-quantile of the standard normal distribution. If $|t| \leq q_{1-\frac{\alpha}{2}}$, then the null hypothesis $H_0$ is accepted. If $|t| > q_{1-\frac{\alpha}{2}}$, then $H_0$ is rejected.

2.2.4 p-value

For those tests, the figure that will be of interest is the p-value. In statistical significance testing, it gives the likelihood of having a test statistic result at least as extreme as the one observed, assuming the null hypothesis is true. In practice, it means that the null hypothesis is rejected when the p-value is smaller than a given significance level. In this study, the level of confidence is set to 5%.

2.3 Independence tests in practice

2.3.1 Between frequency and severity of the losses

In the model, we assume the frequency distribution to be independent from the severity distribution. Assuming the losses are properly modelled through the SBA model – which implies that the internal loss data are coherent with the SBA model, this assumption could be back tested on the sample of internal losses. The study should be conducted within each homogeneous set, i.e. for each entity, and for each operational risk category. As the database is relatively young, working with annual observations is not appropriate, because the number of observations is very low. Then, it has been decided to work with the monthly observations of the severity and frequency of losses.

For the monthly observations of the severity, it has been decided to work with the log-median instead of the mean. Indeed, the quantiles of a distribution are preserved under monotone transformations of the data, unlike mean. Thus, the median of the log-losses is equal to the log of the median of the losses. In general, let $g(.)$ be a nondecreasing function. Then, for any random variable $Y$, the quantiles of $g(Y)$ coincide with the transformed quantiles of $Y$. This property comes from the fact that $\mathbb{P}(Y < q_a) = \alpha \Rightarrow \mathbb{P}(g(Y) < g(q_a)) = \alpha$, especially for $\alpha = 50\%$. Furthermore, if working with right-skewed distributions, the more observations you have, the more likely you are to have captured extreme events that would weight a lot in the mean calculation. Thus, conducting this study with the mean would increase the correlation (if any) between the frequency and the impact.

Using the actual internal data is an issue in practice because out of the 40 000 losses reported within the company, half of them indicate a null impact – either because the amount of the financial loss in not known or because there is no financial impact, but a reputational or strategic one. Those approximately 20 000 losses should then be divided up to more than 50 entities. Considering only the ones that are the most mature, and that have the thicker database, there are 10 000 losses within 10 entities. In order to study homogeneous risks within each of the entity, we should then work with the 100 risk categories, which would lead to 10 observations per risk category in each entity on average. In order to overcome this difficulty, we will study a higher level of granularity and only focus on risk categories for which the database is less precarious.

At the moment, only a few categories and entities enable to provide data for conducting independence tests. Below is an example of the outcomes. In the end, most of the tests performed enable to accept the null hypothesis of independence between the severity and the frequency.
Nevertheless, the gap between the p-value entails that those different methods lead to different levels of significance.

<table>
<thead>
<tr>
<th>Correlations</th>
<th>Test</th>
<th>Stat</th>
<th>q:</th>
<th>p-value:</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>P(</td>
<td>X</td>
<td>&lt;q)=5%</td>
</tr>
<tr>
<td>Linear</td>
<td>Chi2</td>
<td>-20,12%</td>
<td>Linear</td>
<td>-1,5368</td>
<td>2,0032</td>
</tr>
<tr>
<td>Log-linear</td>
<td></td>
<td>-7,89%</td>
<td>Log-linear</td>
<td>-0,5924</td>
<td>2,0032</td>
</tr>
<tr>
<td>Spearman rho</td>
<td></td>
<td>-16,87%</td>
<td>Spearman</td>
<td>-1,2811</td>
<td>2,0032</td>
</tr>
<tr>
<td>Kendall tau</td>
<td></td>
<td>-11,62%</td>
<td>Kendall</td>
<td>-1,2879</td>
<td>1,9600</td>
</tr>
</tbody>
</table>

Table 3: Results of different independence test on the same sample

2.3.2 Between frequency and severity parameters

Another type of independence could be linked to the expert bias: when assessing both frequency and severity, the expert could tend to adjust the severity with respect to the frequency, and vice-versa. Independence tests were applied to the sets of parameters for all the risk assessments, in all the entities. In this case, dividing up in sub-categories is not logical because this test intends at capturing a link between the parameter that is due to the expert judgement bias. Thus, this phenomenon should be observed at Group level, all categories taken together. Indeed, most the statistical tests of independence detailed above rejected the hypothesis of independence of the frequency parameters and the severity parameters.

That result seems consistent with the Operational Risk framework: it is a consequence of the fact the Scenario Based Approach focuses on low frequency/high severity risks. In the Operational Risk lognormal environment, the value of the frequency parameter $\lambda$ is in some extent conditional to the parameters for the impacts, that are $\mu$ and $\sigma$. Indeed, the model focuses on severe risks (low frequency/high severity). Moreover it is assumed that attritional risks (high frequency/low severity) are taken into account in the Profit and Loss attribution therefore those risks should not be quantified in the internal model in order to avoid double counting.

It is assumed that low frequency/low impact risks do not affect the business, and high frequency/high impact risks do not exist, because if the insurer had borne them before it would have already gone bankrupt. In the end, it is pretty much expected that if the expert first assesses frequency as high, he would them match it with low severity.
Here, the question is about the causality rather than existence of a correlation between the frequency and the severity parameters: assuming it is more likely to have a low impact in front of a high frequency, in which extent would the expert have tended to underestimate the severity regarding the frequency he has just assessed?

This question could be answered by asking the expert to assess the frequency of all the scenarios in a first time. Then, say 6 months later, the expert is asked to assess the impact of the scenario. Thanks to this procedure, the expert would be less influenced by the frequency parameter because he does not have their value in mind.

Also, it could be conceivable to ask different experts to do the exercise: one of them would assess the frequency and the other would assess the impacts. In the end, the obtained assessments could
be compared to the initial assessment in order to highlight the existence or not, of an unconsent
tendency that the expert has to match the frequency with the impact levels. However, when doing
this exercise, it is also expected that because of the inherent subjectivity of the expert judgement,
the results would significantly change from the initial assessment, done by a single expert.

3 Back testing the iid hypothesis

In the model, it is assumed that the impacts are independent and identically distributed. It could be
interesting to verify if the assumption of independence is reasonable. Other types of non-
parametrical tests are suited for such a study. They are presented in this section.

3.1 Non-parametric tests

3.1.1 Turning points test

The series \((X_i)_{i=1,...,n}\) is said to have a turning point at \(i\) if one of the following patterns exist:

- \(X_i > X_{i-1}\) and \(X_i > X_{i+1}\)
- \(X_i < X_{i-1}\) and \(X_i < X_{i+1}\)

The turning point test enables to test whether there is a positive or negative correlation at lag 1 in
the sample. Let \(T\) be the number of turning points in the sample:

\[
T = \sum_{i=2}^{n-1} \mathbb{I}\{ (X_i - X_{i-1})(X_i - X_{i+1}) > 0 \}
\]

Since the probability of a turning point is \(4/6\), under the null hypothesis \(H_0\) of \((X_i)_{i=1,...,n}\) being
independent identically distributed, we have:

\[
t = \frac{T - \mathbb{E}(T)}{\sqrt{\text{Var}(T)}} \overset{p}{\to} \mathcal{N}(0,1)
\]

with \(\mathbb{E}(T) = \frac{2}{3} (n - 2)\) and \(\text{Var}(T) = \frac{16n-29}{90}\).

Thus, for a chosen confidence level \(\alpha\), we compare \(t\) to \(q_{1-\frac{\alpha}{2}}\) which is the \((1 - \frac{\alpha}{2})\)-quantile of the
standard normal distribution. If \(|t| \leq q_{1-\frac{\alpha}{2}}\), then the null hypothesis \(H_0\) is accepted. If \(|t| > q_{1-\frac{\alpha}{2}}\),
then \(H_0\) is rejected.

3.1.2 Difference-sign test

The difference-sign test enables to determine if there is a trend in the sample. Let \(S\) be the number
of times \(X_i > X_{i-1}\) in the sample:

\[
S = \sum_{i=2}^{n} \mathbb{I}\{ X_i - X_{i-1} > 0 \}
\]
Under the null hypothesis $H_0$ of $(X_i)_{i=1,...,n}$ being independent identically distributed, we have:

$$t = \frac{S - \mathbb{E}(S)}{\sqrt{\text{Var}(S)}} \overset{D}{\rightarrow} \mathcal{N}(0,1)$$

with $\mathbb{E}(S) = \frac{n-1}{2}$ and $\text{Var}(T) = \frac{n+1}{12}$.

Thus, for a chosen confidence level $\alpha$, we compare $t$ to $q_{1-\frac{\alpha}{2}}$ which is the $(1-\frac{\alpha}{2})$-quantile of the standard normal distribution. If $|t| \leq q_{1-\frac{\alpha}{2}}$, then the null hypothesis $H_0$ is accepted. If $|t| > q_{1-\frac{\alpha}{2}}$, then $H_0$ is rejected.

### 3.1.3 Wald–Wolfowitz runs test

Still focusing on the differences, we could test the hypothesis that the elements of the difference sequence are mutually independent. For that, we apply the following transformation to the data:

$$Y_i = I\{X_i - X_{i-1} > 0\} - I\{X_i - X_{i-1} < 0\}; \ i = 2, ..., n$$

The run test is based on the null hypothesis that each element in the sequence is independently drawn from the same distribution. A “run” of a sequence is a maximal —non-empty segment of the sequence consisting of adjacent equal elements. For example, the sequence “++++++--+++++” consists of 5 runs, 3 of which consists of “+”, and the 2 others of “-”.

Under the null hypothesis, the number of runs in a series of $n$ elements is a random variable whose conditional distribution given the observation of $n_+$ positive values and $n_-$ negative values is approximately normal with mean $\mu = \frac{n_+n_-}{n} + 1$ and variance $\sigma^2 = \frac{(\mu-1)(\mu-2)}{n-1}$. Thus, we have:

$$t = \frac{nb(\text{runs}) - \mu}{\sigma} \overset{D}{\rightarrow} \mathcal{N}(0,1)$$

### 3.2 Back testing the iid hypothesis in practice

In practice, assuming the losses are properly modelled through the SBA model – which implies that the internal loss data should reflect the SBA, the iid assumption could be back tested on the sample of internal losses. In order to test the hypothesis of the impacts being independent identically distributed (iid), we should focus on homogeneous sets of data: losses that occurred in a given entity that belong to a given risk category. For this study, the retain granularity is the thinner as we seek for homogeneity. Nevertheless, the Group typology is made of a hundred of risk categories. Hence, only the most robust sets will be retained.

Moreover it has been decided to work with the series of impact losses per year. Indeed, the SBA model assumed that within the same scenario the losses are independent identically distributed (iid) within the year but not during the whole period of observation. For instance, it is very likely that the parameters of the scenario change from one year to another due to changes in the economic environment, in the internal controls or in the regulation. Similarly, it is expected that the loss amounts would have changed in such a situation from one year to another.
As it was said before, the internal loss database is still quite empty which entails that only relevant entities, year and risk categories have been selected for conducting this analysis. To decide which set should belong to the analysis, the number of 100 losses reported in the segment was set to be the minimum.

In the end, the iid hypothesis was tested on 38 event risk categories within 3 entities. Only one of the above tests rejected the null hypothesis of the impacts being iid at a confidence level of 5%. These results give a quite good level of comfort on the iid assumption of the model.

4 Conclusion

Even if those tests had contradicted our independence assumptions, it is important to recall that reductions of the problem have to be performed when designing a model. Frachot et al. (2004)\(^42\) clarified this point in their paper when they recall that the annual loss of a given event type is driven by two sources of randomness: frequency and severity. Therefore the correlation between two risks could be the result of correlation between frequencies, severities, or both. They admitted that the correlation between frequencies is acceptable if both frequencies share common dependence with respect to a variable such as the size of business or the economic environment. Nevertheless, they explain that a correlation between severities is way harder to tackle: “a basic feature of actuarial models requires assuming that individual losses are jointly independent within one specific business line/risk type class. Therefore it is conceptually difficult to assume simultaneously severity independence within each class and severity correlation between two classes”. Coppé et al. (2008)\(^43\) continue by identifying three methods to introduce dependence: (1) between frequencies, (2) between severities, (3) dependence is introduced only at the level of the aggregate losses for each unit of measure. They conclude that “Methods 1 and 3 are commonly used in practice, as they are the simplest to implement, and it is often difficult to formulate a coherent model that defines dependencies at the level of individual loss severities. Usually, measurements of frequency correlations are used to derive models for correlations at the aggregate loss level, while otherwise assuming independence of loss severities. Several authors, including (Nystrom and Skoglund 2002; Di Clemente and Romano 2003; Chapelle, Crama, Hubner, and Peters 2004; El Gamal, Inanoglu, and Stengel 2006), have proposed models in which correlations are only introduced at the aggregate loss level”. The methodology for modelling dependence between operational risks presented hereafter is aligned with the last statement: the aggregation process is based on the introduction of dependence at the aggregate loss level. It will be presented in the next chapter.

\(^42\) A. FRACHOT, T. RONCALLI, E. SALOMON, The Correlation Problem in Operational Risk, Groupe de Recherche Opérationnelle du Crédit Agricole, France, 2004

\(^43\) E. COPPE, G. ANTONINI, Observed Correlations and Dependencies Among Operational Losses in the ORX Consortium Database, IBM Zurich Research Lab, 2008
CHAPTER 7
AGGREGATION

Diversification represents a fundamental concept in almost all areas of insurance, banking and finance, from portfolio management to strategic decision-making at corporate level. Given the nature of operational risk a “common sense” suggests that operational risk events might be, at least partially, decorrelated. Indeed that all severe operational risk losses occur simultaneously and systematically in the same year is rather dubious and is hardly supported by empirical evidence. A correlation matrix is then introduced to take into account the diversification effects.

In practice, several levels of diversification are considered in the aggregation process. For instance, it is assumed that there exist diversification effects between two risks of a different nature. Similarly, geographical diversification is introduced as an event occurring on a Japanese entity should not be fully correlated with an event happening in Spain.

As it has been discussed in previous chapter, the dependence between risks is introduced at the aggregate loss level. The dependence structures are summarized into a correlation matrix that gives the correlation coefficient for each pair of risks. However, the expert-judgement matrix does not always have the required properties whereas usual aggregation techniques require the input matrix to be semi-definite positive in order to perform the Cholesky algorithm. This chapter reminds the reader of the desired properties of a correlation matrix that are needed for the validity of the input matrix. It also suggests a method that slightly modifies the input matrix in order to build a correlation matrix that is very close to the initial one.

Then, this chapter presents some usual aggregation techniques. The first one is the variance-covariance method that consists of multiplying the vector of risks by the covariance matrix. It discusses why and in which extent this method is not adapted for modelling Operational risks. The second method is based on the theory of copulas. This chapter will present aggregation techniques based on the ranks of multivariate distributions. We will study those methods with a Gaussian and a Student distribution and try to decide which approach is most adapted to a SBA internal model.

1 Correlation matrix

The correlation matrix is a positive semi-definite matrix whose element in the \((i,j)\) position is the correlation between the \(i^{th}\) and \(j^{th}\) elements of a random vector. As it has been discussed previously, the correlation coefficients of this study are defined as rank correlations: Spearman’s rho. In the SBA, most of the correlation levels are determined by expert-judgment. In some very specific cases, a correlation coefficient of 100% is conceivable. When considering assessments that are given by experts for this dependence structure, checking the validity of their hypothesis can be necessary.
1.1 Cholesky decomposition and positive definiteness

The Cholesky decomposition of a symmetric positive-definite matrix \( \Sigma \) is a decomposition of the form \( \Sigma = CC^T \) where \( C \) is a lower triangular matrix with real and positive diagonal entries and \( C^T \) denotes its conjugate transpose. It is commonly used in the Monte Carlo method for simulating systems with multiple correlated variables. For instance, it makes it very easy to simulate correlated normal variables. Indeed, let \((X_1, ..., X_n) \sim \mathcal{N}(0,1d_n)\). Then \( C(X_1, ..., X_n)^T \sim \mathcal{N}(0, \Sigma) \), where \( \Sigma = CC^T \).

There are various methods for calculating the Cholesky decomposition that we will not detail here. For instance, the Cholesky decomposition function is already implemented in several statistical software programmes such as Matlab or R. The computational complexity of commonly used algorithms is \( O(n^3) \) in general\(^{44}\).

Every real-value symmetric positive-definite matrix has a unique Cholesky decomposition\(^{45}\). The question of existence of a Cholesky decomposition when \( A \) is positive semi-definite has already been treated in the literature\(^{46}\). In this case, the uniqueness of the solution is not ensured – which does not prevent the algorithm from converging. However, the decomposition does not exist for matrices that are not symmetric positive semi-definite. Whereas correlation matrices are positive-semidefinite in theory, we will now see that it is not always the case in practice, notably when correlation matrices are built on expert opinion.

1.1.1 Definition of positive definiteness

In algebra, a symmetric real matrix \( M \in \mathbb{R}^{n \times n} \) is said to be positive semi-definite if \( x^T M x \geq 0, \forall x \in \mathbb{R}^n \). Positive semi-definiteness is a property that can be studied from the eigenvalues point of view.

Let \( M \in \mathbb{R}^{n \times n} \) be a positive semi-definite matrix, \( \lambda \in \mathbb{R} \) an eigenvalue of \( M \) and \( v \in \mathbb{R}^n \) an eigenvector of \( M \). As \( v \neq 0 \), we have \( v^T v = \Vert v \Vert_2^2 \geq 0 \). Thus:

\[
\lambda = \frac{v^T \lambda v}{v^T v} = \frac{v^T M v}{v^T v} \geq 0
\]

Therefore, the eigenvalues of \( M \) are non-negative.

Reciprocally, let \( M \in \mathbb{R}^{n \times n} \) be a matrix with non-negative eigenvalues. By the spectrum theorem, we have: \( M = PDP^{-1} = PDP^T \) is an eigendecomposition of \( M \) where \( P \in O_n(\mathbb{R}) \) is a unitary matrix whose rows comprise an orthonormal basis of eigenvectors of \( M \) and \( D \in \mathcal{D}_n(\mathbb{R}) \) is a diagonal matrix whose main diagonal contains the corresponding eigenvalues. Let \( v = (v_1, ..., v_n) \in \mathbb{R}^n \) and \( v' = P^T v = (v'_1, ..., v'_n) \in \mathbb{R}^n \).

\[
v^T M v = v^T PDP^T v = (P^T v)^T D (P^T v) = X^T D X' = \sum_{i=1}^n \lambda_i (x'_i) \geq 0
\]


\(^{46}\) A. HOUSEHOLDER, The theory of matrices in numerical analysis, USA, 1964
This entails that $M$ is positive semi-definite.

In the end, a matrix is positive semi-definite if and only if its eigenvalues are non-negative. Consequently, a Cholesky decomposition exists if and only if the eigenvalues of the matrix are non-negative.

### 1.1.2 Positive-definiteness of matrices with the same correlation coefficient

It is common practice for experts to set the same correlation coefficient in the whole matrix. If $\rho$ denotes the said correlation coefficient, then the matrix is like:

$$M_n(\rho) = \begin{pmatrix} 1 & \cdots & \rho \\ \vdots & \ddots & \vdots \\ \rho & \cdots & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$$

Let’s check the definiteness of this matrix by computing its eigenvalues $\lambda$:

$$M_n(\rho)v = \lambda v \quad \Leftrightarrow \quad \begin{pmatrix} 1 & \cdots & \rho \\ \vdots & \ddots & \vdots \\ \rho & \cdots & 1 \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

The system corresponds to:

$$\forall i = 1 \ldots n, \ v_i + \rho \sum_{j=1, j \neq i}^n v_j = \lambda v_i$$

$$\Leftrightarrow \forall i = 1 \ldots n, (1 - \rho)v_i + \rho \sum_{j=1}^n v_j = \lambda v_i$$

By summing the $n$ equations above, we obtain:

$$(1 - \rho) \sum_{i=1}^n v_i + n \rho \sum_{j=1}^n v_j = \lambda \sum_{i=1}^n v_i$$

- If $\sum_{i=1}^n v_i = 0$
  The above equation does not give any information, then we can use the previous one that leads to:
  $$\forall i = 1 \ldots n, (1 - \rho)v_i = \lambda v_i \Rightarrow 1 - \rho = \lambda$$

- If $\sum_{i=1}^n v_i \neq 0$
  We can now simplify the equation by $\sum_{i=1}^n v_i$ which leads to:
  $$(1 - \rho) + n \rho = \lambda$$

Then we have all possible eigenvalues for $M_n(\rho) \in \mathbb{R}^{n \times n}$ in $\{(1 - \rho); (1 - \rho) + n \rho\}, \forall n \geq 2$

Therefore, $M_n(\rho)$ is positive definite for: $\begin{Bmatrix} 1 - \rho > 0 \\ (1 - \rho) + n \rho > 0 \end{Bmatrix}$
Thus, \( M_n(\rho) \) positive definiteness is verified when \( \rho \in \left[ -\frac{1}{n-1}, 1 \right] = D_n, \forall n \geq 2 \).

In the end, as negative correlation coefficients are not utilized, we have \( M_n(\rho) \) positive definite when \( \rho \neq 100\% \). If \( \rho = 100\% \), then no diversification effect is to be taken into account. In that case, the aggregation is a simple sum and the computation is straightforward.

### 1.1.3 Positive definiteness of matrices with multiple correlation coefficients

All matrices are not composed of the same correlation coefficient. In this case, it happens that the property of definiteness is not verified anymore. For instance, let’s take the following correlation matrix:

\[
M = \begin{pmatrix}
1 & 1 & 0.75 \\
1 & 1 & 1 \\
0.75 & 1 & 1
\end{pmatrix}
\]

The function \( \text{eigen}(M) \) on R gives the following eigenvalues and eigenvectors:

\[
\lambda = \begin{pmatrix}
2.83808749 \\
0.25 \\
-0.08808749
\end{pmatrix}
\]

As there is one negative eigenvalue, this matrix is not semi-definite positive. Therefore, the Cholesky decomposition cannot work with this matrix. This illustration highlights the fact that building a correlation matrix is not straightforward. Let’s see which values would be eligible for a matrix of dimension 3 that has 100% correlation in all cases, except for one component, that has a correlation coefficient \( \rho \), that is:

\[
M(\rho) = \begin{pmatrix}
1 & 1 & \rho \\
1 & 1 & 1 \\
\rho & 1 & 1
\end{pmatrix}
\]

We compute the determinant of \( M(\rho) - \lambda I_d \) to have its eigenvalues:

\[
|M(\rho) - \lambda I_d| = \left| \begin{pmatrix}
1 - \lambda & 1 & \rho \\
1 & 1 - \lambda & 1 \\
\rho & 1 & 1 - \lambda
\end{pmatrix} \right|
\]

\[
= (1 - \lambda)^3 + \rho + \rho - (1 - \lambda) - (1 - \lambda) - \rho^2(1 - \lambda)
\]

\[
= (1 - \lambda)^3 + 2\rho - 2(1 - \lambda) - \rho^2(1 - \lambda)
\]

\[
= (1 - \rho - \lambda)((1 - \lambda)^2 + \rho(1 - \lambda) - 2)
\]

\[
= (1 - \rho - \lambda) \left( 1 - \lambda + \frac{-\rho + \sqrt{\rho^2 + 8}}{2} \right) \left( 1 - \lambda + \frac{-\rho + \sqrt{\rho^2 + 8}}{2} \right)
\]
The roots of this equation are then: 

\[(1 - \rho)(1 - \frac{-\rho - \sqrt{\rho^2 + 8}}{2})(1 - \frac{-\rho + \sqrt{\rho^2 + 8}}{2})\]. In order to have the positive semidefiniteness, we need to have 

\[
\begin{align*}
(1 - \rho) & \geq 0 \\
\left(1 + \frac{\rho + \sqrt{\rho^2 + 8}}{2}\right) & \geq 0 \text{(OK)} \\
\left(1 + \frac{\rho - \sqrt{\rho^2 + 8}}{2}\right) & \geq 0
\end{align*}
\]

\[\Leftrightarrow \begin{cases} 
\rho \leq 1 \\
\left(1 + \frac{\rho - \sqrt{\rho^2 + 8}}{2}\right) \geq 0
\end{cases}\]

Let \(f(\rho) = \left(1 + \frac{\rho + \sqrt{\rho^2 + 8}}{2}\right)\). Then \(f'(\rho) = \frac{1}{2}\left(1 - \frac{2\rho}{2\sqrt{\rho^2 + 8}}\right) \geq 0\). Therefore, \(f\) increases with \(\rho\). Yet \(f(1) = 0\). Then \(\rho = 1\) is the only value for \(\rho\) that satisfies the inequality, and for which \(M\) is semi-definite positive (and therefore a correlation matrix). The strict inequations are never verified, and there is no solution such that \(M\) is positive definite.

### 1.2 Nearest correlation matrix methods

In order to apply the Cholesky decomposition, the matrix must be definite semi-positive. For that, it happens that the correlation coefficients have to be changed.

A first approach would be to change one by one the coefficients, so that the obtained matrix satisfies the required properties. To be conservative, the coefficients should be increased rather than decreased.

A second method consists of modifying Cholesky algorithm instead of the correlation matrix itself. Indeed, the algorithm involves the square root of a sum that might be negative. If this happens, the idea is to change the last used correlation value in a way to make this sum zero instead of negative.

Otherwise, workaround could be developed to modify slightly the correlation coefficients. Those methods create a matrix that is definite semi-positive and has coefficients that are very close to the ones of the initial matrix. This section aims to present such methods.

#### 1.2.1 Method of the spectral decomposition

The method, developed in Ricardo Rebonnato and Peter Jäckel’s paper\(^{47}\), builds a correlation matrix that satisfies the required conditions and best matches a given not positive-semidefinite matrix. It provides a backwards calculation of the origin matrix with given eigenvalues and given eigenvectors.

Given a real and symmetric matrix \(M\), we define:

\(^{47}\) REBBONATO R., JÄCKEL P., "The most general methodology to create a valid correlation matrix for risk management and option pricing purposes", 1999
v = eigenvectors(M),

and Λ = diag(λ_i), where λ_i = eigenvalues(M)

First the negative eigenvalues have to be replaced by zero. There is at least one negative eigenvalue, because the matrix is not semi-definite, so Λ’ is defined as:

\[ Λ’ = \text{diag}(λ_i') : λ_i' = \begin{cases} λ_i, & λ_i \geq 0 \\ 0, & λ_i < 0 \end{cases} \]

Second we define a scaling matrix as:

\[ T = \text{diag}(t_i): t_i = \left(\sum_m v_i^2 λ_m'\right)^{-1} \]

The row vectors of the matrix \( B' = \sqrt{v} \Lambda' \) can be normalized by \( T \) in such a way that we can calculate a slightly changed, but positive semi-definite correlation matrix:

\[ \tilde{M} = BB^T, \]

with \( B = \sqrt{T}B' = \sqrt{T}v\sqrt{Λ'} \)

Thanks to this procedure, we obtain a correlation matrix which is intuitively similar to the initial one. Besides, the fewer eigenvalues have to be replaced by zero, the closer the modified matrix is from the initial one.

In order to illustrate this method on a small matrix, let’s see the detailed calculations of the non-semi-definite positive matrix given in previous section. The initial correlation matrix is:

\[ M = \begin{pmatrix} 1 & 1 & 0.75 \\ 1 & 1 & 1 \\ 0.75 & 1 & 1 \end{pmatrix} \]

The function eigen(M) on R gives the following results:

\[
\begin{align*}
\text{eigenvalues}(M) &= \begin{pmatrix} 2.83808749 & 0.25 \\ -0.08808749 & 0.7071068 & 0.4311881 \\ -0.6097921 & 2.78179e-16 & -0.7925614 \end{pmatrix} \\
\text{eigenvectors}(M) &= \begin{pmatrix} -0.5604256 & 0.7071068 & 0.4311881 \\ -0.6097921 & 2.78179e-16 & -0.7925614 \\ -0.5604256 & -0.7071068 & 0.4311881 \end{pmatrix}
\end{align*}
\]

As described in the methodology, we will replace the negative eigenvalue by zero, which entails:

\[ Λ' = \begin{pmatrix} 2.83808749 & 0 & 0 \\ 0 & 0.25 & 0 \\ 0 & 0 & 0 \end{pmatrix} \]

It comes:
\[ B' = v\sqrt{\Lambda'} = \begin{pmatrix} -0.94412795 & 0.3535534 & 0 \\ -1.02729383 & 1.390896e - 16 & 0 \\ -0.94412795 & -0.3535534 & 0 \end{pmatrix} \]

The computation of the scaling matrix \( T \) is the following:

\[
\begin{align*}
t_1 &= 1/ \sum_{m=1}^{3} v_{1m}^2 \lambda'_m \\
t_2 &= 1/ \sum_{m=1}^{3} v_{2m}^2 \lambda'_m \\
t_3 &= 1/ \sum_{m=1}^{3} v_{3m}^2 \lambda'_m \\
\end{align*}
\]

\[
\begin{align*}
\Rightarrow t_1 &= 1/(0.5604256^2 \times 2,83808749 + 0.7071068^2 \times 0.25 + 0.4311881^2 \times 0) \\
t_2 &= 1/(0.6097921^2 \times 2,83808749 + (2.78179e - 16)^2 \times 0.25 + 0.7925614^2 \times 0) \\
t_3 &= 1/(0.5604256^2 \times 2,83808749 + 0.7071068^2 \times 0.25 + 0.4311881^2 \times 0) \\
\end{align*}
\]

It gives:

\[
T = \begin{pmatrix} 0.98388631 & 0 & 0 \\ 0 & 0.94756854 & 0 \\ 0 & 0 & 0.98388631 \end{pmatrix}
\]

Then:

\[
B = B' \sqrt{T} = \begin{pmatrix} -0.93649036 & 0.35069330 & 0 \\ -1 & 1.35394e - 16 & 0 \\ -0.93649036 & -0.35069330 & 0 \end{pmatrix}
\]

Finally, we have:

\[
\hat{M} = BB^T = \begin{pmatrix} 1 & 0.936490365 & 0.75402840 \\ 0.936490365 & 1 & 0.936490365 \\ 0.75402840 & 0.936490365 & 1 \end{pmatrix}
\]

Let’s verify its semi-definite positivity by checking its eigenvalues:

\[
eigenvalues(\hat{M}) = \begin{pmatrix} 2.754028 & 0 & 0 \\ 0 & 0.2459716 & 0 \\ 0 & 0 & 2.936162e - 16 \end{pmatrix}
\]

In conclusion, we obtain a semi-definite positive matrix that does not deviate too much from the original and for which the Cholesky decomposition converges. We can see that it remains relatively close to the original matrix. The changes in the correlation coefficient are:

\[
\begin{align*}
\frac{0.75402840}{0.75} - 1 &= +0.54\% \\
\frac{0.936490365}{1} - 1 &= -6.35\%
\end{align*}
\]
1.2.2 Method of the projection

This method, developed in Nicholas J. Higham’s paper, consists in projecting from the set of symmetric matrices onto the correlation matrices, which is the intersection of the following sets:

- \( S = \{ Y \in \mathbb{R}^{n \times n}, Y = Y^T, Y \geq 0 \} \): symmetric semi-definite positive matrices
- \( U = \{ Y \in \mathbb{R}^{n \times n}, Y = Y^T, \forall i = 1: n, y_{ii} = 1 \} \): symmetric matrices with unit diagonal elements

We denote \( P_U \) the projection onto \( U \) and \( P_S \) the projection onto \( S \).

\[
P_U(A) = A - \text{diag}(A - I)
\]

In other terms, if \( A = (a_{i,j})_{i,j=1:n} \), then

\[
P_U(A) = \begin{pmatrix}
1 & a_{1,2} & \ldots & a_{1,n} \\
a_{2,1} & 1 & \ldots & a_{2,n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n,1} & a_{n,2} & \ldots & 1
\end{pmatrix}
\]

Projection onto \( S \) is what is done in the spectral decomposition method therefore we use similar notations. For a symmetric \( A \in \mathbb{R}^{n \times n} \) with spectral decomposition:

\[
A = v\Lambda v^T
\]

where \( \Lambda = \text{diag}(\lambda_i) \), let:

\[
P_S(A) = v\Lambda' v^T
\]

where \( \Lambda' = \text{diag}(\lambda'_i) \):

\[
\lambda'_i = \begin{cases} 
\lambda_i, & \lambda_i \geq 0 \\
0, & \lambda_i < 0 
\end{cases}
\]

To find the nearest matrix at the intersection of the sets \( S \) and \( U \) we iteratively project by repeating the operation:

\[
A \leftarrow P_U(P_S(A))
\]

The algorithm to apply is as follows:

Given a real and symmetric matrix \( M \):

\[
\Delta S_0 = 0 ; Y_0 = M ; \varepsilon = M ; k = 1 ; \text{tol} = 0.0001
\]

while (norm(\( \varepsilon \)) > tol){

\[
R_k = Y_{k-1} - \Delta S_{k-1}
\]

\[
X_k = P_S(R_k)
\]

\[48 \] HIGHAM Nicholas J., “Computing the Nearest Correlation Matrix – A Problem from Finance”, Manchester Institute for Mathematical Sciences, 2002
\[
\Delta S_k = X_k - R_k \\
\text{tmp} = Y_k \\
Y_k = P_U(X_k) \\
\varepsilon = \text{tmp} - Y_k \\
k = k + 1
\]

The algorithm was implemented on R. In order to illustrate this method on a small matrix, let’s see the first steps of calculation of the algorithm applied to the non-semi-definite positive matrix given in previous section. The initial correlation matrix is:

\[
M = \begin{pmatrix}
1 & 1 & 0.75 \\
1 & 1 & 1 \\
0.75 & 1 & 1
\end{pmatrix}
\]

\( k = 1 \)

\[
R_1 = \begin{pmatrix}
1 & 1 & 0.75 \\
1 & 1 & 1 \\
0.75 & 1 & 1
\end{pmatrix}
\]

\[
X_1 = \begin{pmatrix}
1.0163 & 0.9699 & 0.7664 \\
0.9699 & 1.0553 & 0.9699 \\
0.7674 & 0.9699 & 1.0164
\end{pmatrix}
\]

\[
\Delta S_1 = \begin{pmatrix}
0.0163 & -0.0301 & 0.0164 \\
-0.0301 & 0.0553 & -0.0301 \\
0.0164 & -0.0301 & 0.0164
\end{pmatrix}
\]

\[
Y_1 = \begin{pmatrix}
1 & 0.9699 & 0.7664 \\
0.9699 & 1 & 0.9699 \\
0.7674 & 0.9699 & 1
\end{pmatrix}
\]

\( k = 2 \)

\[
R_2 = \begin{pmatrix}
0.9836 & 1 & 0.75 \\
1 & 0.9447 & 1 \\
0.75 & 1 & 0.9836
\end{pmatrix}
\]

\[
X_2 = \begin{pmatrix}
1.0072 & 0.9560 & 0.7736 \\
0.9560 & 1.0265 & 0.9560 \\
0.7736 & 0.9560 & 1.0072
\end{pmatrix}
\]

\[
\Delta S_2 = \begin{pmatrix}
0.0236 & -0.0440 & 0.0236 \\
-0.0440 & 0.0819 & -0.0440 \\
0.0164 & -0.0440 & 0.0236
\end{pmatrix}
\]

\[
Y_2 = \begin{pmatrix}
1 & 0.9560 & 0.7736 \\
0.9560 & 1 & 0.9560 \\
0.7736 & 0.9560 & 1
\end{pmatrix}
\]
In the end, after 18 loops, we obtain:

\[
\hat{M} = \begin{pmatrix}
1 & 0.9433981 & 0.7799990 \\
0.9433981 & 1 & 0.9433981 \\
0.7799990 & 0.9433981 & 1
\end{pmatrix}
\]

### 1.2.3 Comparison of those methods

The second method is iterative, hence longer to implement than the first one that can be easily computed on Excel (except for the eigenvalues and eigenvectors that have to be computed on R, or thanks to a VBA macro). To have an insight on the accuracy of both method, it can be interesting to have a look at the discrepancies between the initial and the modified matrix. The distance used to measure it is the Frobenius norm, that is:

\[
\| M - \hat{M} \|^2 = \sum_{i,j=1}^{n} (m_{ij} - \hat{m}_{ij})^2
\]

The discrepancy between the initial matrix and the matrix modified with Rebbonato’s method is 1.62% with this metric. It is 1.46% with the Higham’s method.

In practice, another issue comes from the numerical implementation of those methods. Their principle is to build a correlation matrix that has similar coefficients to the initial one, and non-negative eigenvalue. It replaces the null eigenvalues by zero. When performing such methods, an issue could arise because of the numerical approximation. It happens that some eigenvalue that are zero in theory are approximated by a multiple of \(-1.10^{-16}\). This practical difficulty could be easily circumvented by changing 0 by \(+1.10^{-10}\) when changing the values of negative eigenvalues.

### 2 Aggregation methods

#### 2.1 Variance-covariance method

As stated in EIOPA’s Advice papers that describe the SCR standard formula, a simple technique to aggregate capital requirements is the use of correlation matrices. Indeed, the capital requirements on sub-risk level are aggregated in order to derive the capital requirement for the overall risk. For instance, the overall SCR is calculated as follows:

\[
SCR_{\text{overall}} = \sqrt{\sum_{i,j}^n Corr_{i,j} SCR_i SCR_j}
\]

where \(i,j\) run over all sub-risks and \(Corr_{i,j}\) denotes the entries of the correlation matrix.

This aggregation method is based on a closed formula that is the consequence of the assumption that the annual losses follow a normal distribution: it is the Gaussian approximation. This model assumes that the vector of annual losses \((X_1, ..., X_n)\) is a Gaussian vector:
\[(X_1, \ldots, X_n) \sim \mathcal{N}(\mu, \Sigma)\]

where \(\mu = (\mu_1, \ldots, \mu_n)'\) and \(\Sigma = (\rho(i,j)\sigma_i\sigma_j)_{i,j \in [1:n]}\)

This method has a very large advantage: it is easy to put in place. It only requires to have quantified the standalone risks, defined the aggregation structure, and estimated the matrices of correlation. It does not use simulation which reduces the computation time.

On the other hand, aggregating risks using a linear aggregation formula might impose a normal structure to risks. However, operational risks are modelled in a Poisson-lognormal environment. Therefore, aggregation with the variance-covariance matrix is not an appropriate choice and other methods, such as copulas, have to be considered.

### 2.2 Gaussian copula

The most common alternative to the variance-covariance method is the copula-based approach to catch dependencies between operational risk scenarios, or entities. One of the main advantages of copulas is to model the dependency structure between variables while preserving the marginal distributions of individual operational risk scenarios.

This thesis does not intend to cover the whole topic of copula, which has already been widely treated in the literature.\textsuperscript{49} It only presents the basics on this mathematical tool in order to understand one numerical application derived from this theory.

#### 2.2.1 Theory

A copula is a multivariate probability distribution function for which the marginal probability distributions of each variable are uniform on the interval \([0; 1]\). Thus, it is defined on \([0; 1]^n\) by:

\[C(u_1, \ldots, u_n) = \mathbb{P}(U_1 \leq u_1, \ldots, U_n \leq u_n), \quad \forall (u_1, \ldots, u_n) \in [0; 1]^n.\]

The applications of copulas are based on the following theorem, known as Sklar’s Theorem: it is a fundamental theorem that enables to link the multivariate cumulative distribution function \(F = F_{X_1, \ldots, X_n}\) to the marginal cumulative distribution functions \(F_{X_1}, \ldots, F_{X_n}\). There exists an \(n\)-copula \(C\) such that for all \((x_1, \ldots, x_n) \in \mathbb{R}^n\), \(F(x_1, \ldots, x_n) = C(F_{X_1}(x_1), \ldots, F_{X_n}(x_n))\). Conversely, if \(C\) is an \(n\)-copula and \(F_{X_1}, \ldots, F_{X_n}\) are cumulative distribution functions, then the function \(F = F_{X_1, \ldots, X_n}\) is an \(n\)-dimensional distribution function with margins \(F_{X_1}, \ldots, F_{X_n}\). If \(F_{X_1}, \ldots, F_{X_n}\) are continuous, then \(C\) is unique.

From Sklar’s theorem we see that for continuous multivariate distribution functions, the univariate margins and the multivariate dependence structure can be separated, and the dependence structure can be represented by a copula.

\textsuperscript{49} The reader could refer to Nelsen (1999).
We denote $F^{-1}(t) = \inf \{ x \in \mathbb{R}, F(x) \geq t \}$ for all $t$ in $[0; 1]$, using the convention $\inf \emptyset = -\infty$. Therefore we have the following corollary:

Let $F = F_{x_1, \ldots, x_n}$ be a $n$-dimensional distribution function with continuous margins $F_{x_1}, \ldots, F_{x_n}$ and copula $C$, such that for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$, $F(x_1, \ldots, x_n) = C \left( F_{x_1}(x_1), \ldots, F_{x_n}(x_n) \right)$. Then for any $(u_1, \ldots, u_n) \in [0; 1]^n$ : $C(u_1, \ldots, u_n) = H \left( F_{x_1}^{-1}(u_1), \ldots, F_{x_n}^{-1}(u_n) \right)$.

2.2.2 Monte Carlo integration for aggregating operational risks

The aggregation of several standalone operational risk scenarios to a higher level is calculated based on the full set of $n_S$ Monte Carlos runs for each scenario in the aggregation.

The aggregation takes into account the correlation matrix between the risks to be diversified. These correlations are defined as rank correlations (Spearman’s rho). As opposed to linear correlations, ranks correlations are non-parametric and can be used for any kind of distribution. Indeed, linear correlations are not robust in case of non-elliptical distributions, which is the case for operational risk. The rank correlation is independent of the marginal distribution.

The aggregation is based on a Normal Copula Method. This aggregation method uses Cholesky decomposition of the correlation matrix to simulate correlated Gaussian vectors with the same correlation as the one assessed by the expert. The ranks of the correlated Gaussian vectors are reproduced in the standalone risks simulation strips so that they have the same rank correlations.

The procedure for aggregating $n$ risks together and for the determination of ranking table is as follows:

1. Simulate $n_S$ Monte Carlo realizations of each the risks $(X_1, \ldots, X_n)$ to be aggregated

   $$ (X_1, \ldots, X_n) = \begin{bmatrix} X_1(1) & \cdots & X_n(1) \\ \vdots & \ddots & \vdots \\ X_1(n_S) & \cdots & X_n(n_S) \end{bmatrix} \in \mathbb{R}^{n_S \times n} $$

2. Simulate vectors $(Y_1, \ldots, Y_n) = \begin{bmatrix} Y_1(1) & \cdots & Y_n(1) \\ \vdots & \ddots & \vdots \\ Y_1(n_S) & \cdots & Y_n(n_S) \end{bmatrix} \in \mathbb{R}^{n_S \times n}$ with the Gaussian copula:
   - Simulation of $Z = (Z_1, \ldots, Z_n) = \begin{bmatrix} Z(1) & \cdots & Z_n(1) \\ \vdots & \ddots & \vdots \\ Z_1(n_S) & \cdots & Z_n(n_S) \end{bmatrix} \in \mathbb{R}^{n_S \times n}$ of independent standard normal variables
   - Multiplication of the resulting vector $Z$ with the triangular of the Cholesky decomposition of the correlation matrix $(n \times n$ matrix) resulting in a matrix $Y = (Y_1, \ldots, Y_n) = ZC \in \mathbb{R}^{n_S \times n}$ of the standard normal variables with approximately the same correlation as the defined correlation matrix for the aggregation, i.e. where:

   $$ \forall i \in [1; n_S] : (Y_1(i), \ldots, Y_n(i)) \sim \mathcal{N}(0, CC^T) = \mathcal{N}(0, \Sigma) \in \mathbb{R}^n $$
3. Give a rank to each of the vectors of correlated standard normal variables $Y_j : j \in [1:n]$

$$(R_1, \ldots, R_n) = \begin{bmatrix}
    \text{rank}(Y_1(1)) & \cdots & \text{rank}(Y_n(1)) \\
    \vdots & \ddots & \vdots \\
    \text{rank}(Y_1(ns)) & \cdots & \text{rank}(Y_n(ns))
\end{bmatrix} \in \mathbb{R}^{ns \times n}$$

with $\text{rank} ( Y_j(i) )$ being the rank of $Y_j(i)$ in the vector $\begin{bmatrix} Y_1(1) \\ \vdots \\ Y_n(ns) \end{bmatrix}$

4. For each risk $j$, link the rank of $Y_j$ with the rank of $X_j$

   o Rearranging the values of the $ns$ Monte Carlo realizations of the analysis objects to be aggregated such that they have the same sort order as in the vectors of the correlated standard normal variables;

   $$(X_1, \ldots, X_n) = \begin{bmatrix}
    X_1(R_1(1)) & \cdots & X_n(R_n(1)) \\
    \vdots & \ddots & \vdots \\
    X_1(R_1(ns)) & \cdots & X_n(R_n(ns))
\end{bmatrix}$$

5. Summing up the rearranged values of the Monte Carlo realizations row by row. Each row gives one Monte Carlo run of the aggregation.

This procedure ensures that the rank correlation of the Monte Carlo runs of scenarios to be aggregated is approximately equal to the specified correlation.

An example is provided next page, with $ns = 10$, $n = 2$, $\Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$.

Below is the R-code for the aggregation process with two standalone scenarios that have their frequency modelled through a Poisson and their impacts modelled through a lognormal:

```R
C=chol(M)
# standalone simulations of the risks to be aggregated
Vpois1<-rpois(ns,lambda1)
Vpois2<-rpois(ns,lambda2)
X1<-sapply(Vpois1,function(Vpois1){sum(rlnorm(Vpois1,mu1,sigma1))})
X2<-sapply(Vpois2,function(Vpois2){sum(rlnorm(Vpois2,mu2,sigma2))})
# simulation of normal variables w/ desired correlation
U1<-rnorm(ns,0,1)
U2<-rnorm(ns,0,1)
U=t(rbind(U1,U2))%*%C
U1=U[,1]
U2=U[,2]
# rearrange X1, X2 in the same order as U1, U2
X1=sort(X1)[rank(U1)]
X2=sort(X2)[rank(U2)]
```
Figure 14: Illustration of the aggregation process with a normal copula approach where the risks to be aggregated are modeled with fixed frequency = 1 and lognormal severity.
2.3 Alternatives to the Gaussian copula

As it has been discussed, normal copulas could be considered as a first approach for the use of copulas in a SBA internal model because they are characterised by a correlation matrix that can be built on expert judgment. Student copulas that belong to the family of elliptical copulas (as well as Gaussian copulas) are also characterised by a correlation matrix. They have an additional parameter: the number of degrees of freedom. Indeed, the Student copula is an option to introduce tail dependence, which means that extreme events are more likely to come together.

The aggregation process would be similar to what has been presented before, but instead of simulating from the multivariate Gaussian distribution, the simulation is done from the multivariate Student with the selected degrees of freedom $\nu$.

The aggregation procedure is the same as for the Gaussian copula method for step 1, 3, 4 and 5. Step 2 is as follows:

2. Simulate vectors $(Y_1, \ldots, Y_n) = \begin{bmatrix} Y_1(1) & \ldots & Y_n(1) \\ \vdots & \ddots & \vdots \\ Y_1(ns) & \ldots & Y_n(ns) \end{bmatrix} \in \mathbb{R}^{ns \times n}$ with the Gaussian copula:

   - Simulation of $Z=(Z_1, \ldots, Z_n) = \begin{bmatrix} Z(1) & \ldots & Z_n(1) \\ \vdots & \ddots & \vdots \\ Z_1(ns) & \ldots & Z_n(ns) \end{bmatrix} \in \mathbb{R}^{ns \times n}$ of independent standard normal variables
   - Multiplication of the resulting vector $Z$ with the triangular of the Cholesky decomposition of the correlation matrix ($n \times n$ matrix) resulting in a matrix $Y = (Y_1, \ldots, Y_n) = ZC \in \mathbb{R}^{ns \times n}$ of the standard normal variables with approximately the same correlation as the defined correlation matrix for the aggregation, i.e. where:

     $$\forall i \in [1; ns] : (Y_1(i), \ldots, Y_n(i)) \sim \mathcal{N}(0, CC^T) = \mathcal{N}(0, \Sigma) \in \mathbb{R}^n$$

   - Simulation of a vector of $ns$ Chi-squared random values independent of with $\nu$ degrees of freedom: $\mathcal{X} = \begin{pmatrix} \mathcal{X}(1) \\ \vdots \\ \mathcal{X}(ns) \end{pmatrix}$

   - Multiplication of the $n$ correlated standard normal variable $Y_i(j)$ by $\frac{\sqrt{\nu}}{\sqrt{\chi(j)}}$ so that

     $$\hat{Y}_i(j) = Y_i(j) \frac{\nu}{\sqrt{\chi(j)}}$$

     is distributed as a multivariate Student with parameters $\Sigma$ and $\nu$.

In order to better understand the differences in the structures of dependence, the figure below gives an illustration of the Gaussian and the Student copulas, in the bivariate case, where $\rho = 50\%$ and $\nu = 1$. The figure also illustrates how the copula influences the aggregation of two variables distributed with Poisson-Lognormal.
Figure 15: Simulation of independent Poisson-Lognormal variables

Figure 16: Simulation of a Gaussian copula (left) and a Student copula (right), $p = 0.5$; $\nu = 1$

Figure 17: Resulting Poisson-Lognormal variables correlated with a Gaussian (left) and a Student (right) copula
The alternative copula option would be introduced as a method to see the impact of alternative assumptions of dependence and check the robustness of the calculation. At this stage, we don’t see a method to explicitly determine the level of degrees of freedom parameter. When used as a back-testing tool to check the robustness of the calculations performed with the Gaussian copula method, the degrees of freedom could be set to something which is seen in financial markets, but this remains arbitrary. As long as the definition of the degree of freedom is not clearly transposed to a SBA internal model, using Gaussian would remain the more reasonable.

3 Conclusion

In the end, we have seen that the drawbacks of a covariance matrix built on expert judgment that might have wrong algebraic properties can be easily circumvented by nearest correlation problems algorithms. In this context, the use of aggregation methods that appeal to Cholesky algorithm is ensured.

When it comes to the choice of an aggregation technique, the decision should be driven by the identified structure of dependence. The variance-covariance method should be dismissed straightaway because operational risks are not modelled with the Gaussian distribution. Furthermore, a short study compared the result of the aggregation of two standalone risks obtained with the Variance-Covariance method to the aggregation of the same risks performed with the Gaussian copula. Those results were compared for different frequency parameter values. It appeared that on a standalone basis, the Gaussian copula is more conservative when the frequency parameter is smaller than 100. As most of the scenarios in the scope of operational risk have a low frequency parameter, it could come to the conclusion that the using the Gaussian copula is more prudent than using the Variance-Covariance method.

Copula method should definitely be preferred. Information about the tail of the distribution could be obtained by computing the coefficient of tail dependence for risks that have sufficient available data. As it is mostly not the case, the choice of the copula will be driven by required properties of the dependence structure, evaluated by the expert. The study mentioned above also compared results obtained with the Gaussian and the Student copulas. It showed that the results of the aggregation performed with the Student copula do not diverg too much from the ones obtained with the Gaussian approach. Because of the uncertainty caused by the calibration of the Student copula in a SBA internal model, we can conclude that the Gaussian assumption seems reasonable.

Unfortunately, this study is not taken over in this thesis. However, this could be subject to further analysis in order to gain better understanding the aggregation process, and therefore evaluate the sensitivity of the model to a change in the aggregation assumptions.
CHAPTER 8
COMPUTATION OF THE CAPITAL CHARGE

Now that the risks have been assessed, quantified and aggregated, the computation of the capital charge only depends on basic statistics of the loss distribution. The calculation of the Operational Risk SCR figure is based on a Monte Carlo approach and performed as follows:

$$SCR = VaR_{99.5\%} - EL$$

where $VaR_{99.5\%}$ is the 99.5% percentile of the n simulations of the aggregate loss

$EL$ is the expected loss of the n simulations

This formula may raise a few questions. First of all, let’s recall how Solvency II defines the SCR: “it shall correspond to the Value-at-Risk of the basic own funds of an insurance or reinsurance undertaking subject to confidence level of 99.5% over a one-year period”. The use of the Value-at-Risk to measure the risk borne by the insurer is discussed in this section. As this topic has been widely treated in the literature, this thesis does not intend to go into this in depth but it will only give a brief overview on this topic in order to ensure that the reader is aware of the limits of the usage of Value-at-Risk.

Moreover, the common understanding of Solvency II’s definition is that SCR is the smaller x such that: $P(Loss > SCR) \leq 0.5\%$. This section will also explain the notion on “unexpected loss” developed in Basel II Accord. This leads to the deduction the expected loss of the distribution from the 99.5%-quantile. The understanding of expected loss will be presented in this chapter. An alternative to the common interpretation of the expected loss by the mean of the loss is suggested. Both approaches are compared in order to determine the most prudent approach.

Another concern on this formula for computing the SCR relates to the use of Monte Carlo simulations. It is necessary to ensure that the level of convergence of the method is satisfying. This topic will be discussed in the last section of this chapter both at the standalone and aggregate level.

1 Risk measure

To determine the amount of economic capital that an insurance undertaking must hold in order to cover from a risk, it is mandatory to decide how to evaluate the level of this risk. Hence, the risk manager should decide which risk measure is more appropriate. Many of them exist and have been discussed in the literature\(^50\) and this section will not go through all of the existing risk measures. As regulators are working on the harmonization of their regulatory requirements, we can see a trend

\(^{50}\) DENUIT Michel, Actuarial theory for Dependent Risks: Measures, Orders and Models, Wiley, Belgium, July 2005
towards convergence between banks and insurance companies. Solvency II Directive uses the Value-at-Risk (VaR) for computing the SCR\textsuperscript{51}, which implies that this risk measure is widely used in risk management.

1.1 Value-at-Risk

This risk measure was introduced in the banking industry in the end of last century, when banks developed models to measure their own risk. For a given portfolio and time horizon, the VaR or Value-at-Risk is the maximal loss that the company is likely to suffer with a certain probability, that is:

\[ \text{VaR}_\alpha(X) = \inf \{ x \mid F_X(x) \geq \alpha \} = q_\alpha \]

This risk measure is widely used because it is easy to compute (estimations methodologies are based on normal assumption, historical data or stochastic methods) and have many interesting characteristics\textsuperscript{52}:

- VaR is a common risk measure: enables to compare different positions and risk factors. For instance, in the insurance industry, VaR enables to compare operational and financial risks that do not have much in common, whereas duration or Greek measures only apply to fixed-income positions and derivative positions respectively.
- VaR captures the effect of the correlations between risks. Moreover, it is a holistic measure that takes into account the effect of all risk factors instead of looking at the effect of risk factors one at the time. Thus, it can be used to measure aggregate risks.
- VaR is a probabilistic measure that gives information on the likelihood associated with certain loss amounts.
- VaR is easily understood as an amount of money lost.

Nevertheless, VaR suffers from limitations that conducted to the development of more sophisticated models. Those limits were formalized by introducing the desired properties that define a “good” risk measure.

1.2 Coherent risk measure

In order to assess risk, one must define what a measure of risk means. To consider risk measures more formally, Artzner et al (1999) postulated a set of risk-measure axioms that are called axioms of coherence. Let \( X \) and \( Y \) be two risky positions. The risk measure \( \rho(.) \) is said to be coherent if it satisfies the following properties:

- Monotonicity: \( X \geq Y \Rightarrow \rho(X) \leq \rho(Y) \)
- Subadditivity: \( \rho(X + Y) \leq \rho(X) + \rho(Y) \)
- Positive homogeneity: \( \rho(aX) = a\rho(X), \forall a > 0 \)

\textsuperscript{51} “Solvency Capital Requirement shall correspond to the Value-at-Risk of the basic own funds of an insurance or reinsurance undertaking subject to a confidence level of 99.5 % over a one-year period.” Solvency II Directive, art. 101.3 : Calculation of the Solvency Capital Requirement

\textsuperscript{52} DOWD Kevin, BLAKE David, After VaR: The Theory, Estimations, and Insurance Applications of Quantile-Based Risk Measures, University of Nottingham, 2006
Translational invariance: $\rho(X + a) = \rho(X) - a$

The second axiom means that it is expected that the risk of a portfolio is not larger than the sum of the risks of each of its sub-portfolios. In other terms, we expect that the aggregation of individual risks – whether correlated or not - does not increase the overall risk.

1.3 Limitations of the VaR

In fact, it is easy to demonstrate that the VaR does not respect the subadditivity result by giving a counter-example. Suppose we have two risks $X$ and $Y$ that are independent identically distributed such that $\mathbb{P}(X = 100) = \mathbb{P}(Y = 100) = 90\%$ and $\mathbb{P}(X = 1000) = \mathbb{P}(Y = 1000) = 10\%$. Then we have $VaR_{95\%}(X) = VaR_{95\%}(Y) = 100$. By independence, we have $\mathbb{P}(X + Y = 200) = 90\% * 90\% = 81\%$, $\mathbb{P}(X + Y = 1100) = 2 * 90\% * 10\% = 18\%$ and $\mathbb{P}(X + Y = 2000) = 10\% * 10\% = 1\%$. Therefore, $VaR_{95\%}(X + Y) = 1100 > 200 = VaR_{95\%}(X) + VaR_{95\%}(Y)$. This violates the property of subadditivity.

In fact, the subadditivity is verified in some particular cases, notably when $X$ and $Y$ have elliptical distributions. As elliptical distributions are the normal and the Student distributions, we do not expect that this result holds in an operational risk framework. Nevertheless, it has been shown that for distributions that have tails “reasonably” heavy, there exists a level of confidence $\alpha_0$ such that $\forall \alpha \in ]\alpha_0, 1[, VaR_\alpha(X + Y) \leq VaR_\alpha(X) + VaR_\alpha(Y)$. In practice, the confidence level that is used in Solvency II framework is $\alpha = 99.5\%$ at a one-year horizon. It is sufficiently close to 100\% to have this property.

2 Expected loss

Commonly in Risk Management, the expected loss is the expected value of a distribution or the arithmetic mean of the distribution. The expected loss is considered as part of the cost of doing business and is already included in the Profit and Loss attribution. In order to avoid double-counting, the expected loss is deduced from the 99.5\% percentile in the SCR computation. However, Supervisory Authorities expect companies to be able to demonstrate that expected loss is adequately captured in its business practices. In a letter published in 2005, Basel Committee gives its insight on the treatment of expected operational risk losses in AMA. It clarifies the notion of expected loss by stating that: “the banks must be able to demonstrate that losses corresponding to the expected loss are highly predicable and reasonably stable and that the estimation process is consistent over time”.

Usually, expected loss is the expected value (mean) of the aggregate loss distribution. However, other metrics could be considered to measure the expected loss. For instance, the median is a

---

53 EMBRECHTS et al., 1998
54 BASEL COMMITTEE ON BANKING SUPERVISION, The treatment of expected losses by banks using the AMA under the Basel II Framework, Basel Committee Newsletter No.7, November 2005
55 “mean is the expected loss”

http://books.google.fr/books?id=KmrVKnlLTSsC&pg=PA140&ljpg=PA140&dq=expected+loss+basel+II+mean+loss+distribution+operational+risk&source=bl&ots=hn7pUU0FRZ&sig=3e5FANaCQUAzuTV3pN7Fb3AleXY&hl=fr&
possible alternative to the choice of the mean for estimating the expected loss. Two explanations could be used as arguments in favour of the use of the median for computing the expected loss.

First of all, it has no impact for low frequency risks modelled through a Poisson distribution because there is a threshold for $\lambda$ under which the median of the loss is zero. Secondly, it is a preventive behaviour as using the median value, instead of the mean, generally underestimates the cost of doing business. Indeed, with the right skewed distributions used for modelling operational risks, the mean is greater than the median.

2.1 Frequency threshold for the null median

When the median is zero, its deduction has no impact on the calculation of the SCR. This section aims to demonstrate the existence of threshold for the frequency parameter under which the median of the aggregate loss is null, on a standalone basis.

First, we computed the SCR values with different values for $\lambda$. We chose to study the median and the SCR figures for a set of $\lambda$ ranging from 0.65 to 0.75. In order to understand the effect on the SCR we also kept the values of the 99.5% quantile and the median. We only used 20 different simulations here, because we just needed to make an idea on the underlying phenomenon.

![Figure 18](http://example.com/image.png) Computation of the SCR (left) and median (right) with different values for $\lambda$ ranging between 0.65 and 0.75

The SCR plot does not enable to make any conclusion, but the median plot lets us think that the threshold for $\lambda$ is around 0.69. Secondly, we decided to focus around the interval $[0.69;0.70]$. For this analysis, we used 500,000 runs for Monte Carlo simulation, which enable to produce accurate results.

![Graphs showing SCR and median plots with different values for $\lambda$ ranging between 0.69 and 0.70](image1)

*Figure 19: Computation of the SCR (left) and median (right) with different values for $\lambda$ ranging between 0.69 and 0.70*

Obviously, there is some kind of a threshold effect on the median. In order to better capture its effect, we proceeded to the same analysis with different values for the lognormal parameters.

![Graphs showing SCR and median plots with different values for $\lambda$ ranging between 0.69 and 0.70](image2)

*Figure 20: Computation of the SCR (left) and median (right) with different values for $\lambda$ ranging between 0.69 and 0.70*

Here, we observe the same effect of a threshold for the frequency parameter above which the value of the median jumps. Moreover, as the parameter $\sigma$ is relatively small, the extreme quantiles do not fluctuate that much. Therefore, with a large mean and small standard deviation, it is clear that the frequency threshold affects the value of the SCR that decreases when the frequency exceeds the threshold.
In order to understand where this phenomenon comes from, let’s first recall that the aggregate loss we are studying is computed with a compound Poisson distribution. This mean than the loss amount is given by:

\[ X = \sum_{i=1}^{N} C_i \]

where \( N \) denotes the number of events

\( C_i \) denotes the amount of the \( i \)th loss

A Monte Carlo run will be zero when there is no event occurring, i.e. when \( N = 0 \). Here, we are focusing on the risks for which \( N \) is following a Poisson distribution, that is: \( N \sim \mathcal{P}(\lambda) \). We have:

\[ \mathbb{P}(X = 0) = \mathbb{P}(N = 0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda} \]

Let \((X_1, ..., X_n)\) be the outcomes of \( n \) runs of Monte Carlo simulations. If we order the sample, the median will be given by the \( \left( \frac{n}{2} \right) \)-th value of the order statistic. The median equal to zero is equivalent to the fact that there are more simulations with an outcome equal to zero than non-zero outcomes. That is:

\[ \{\text{median}(X_1, ..., X_n) = 0\} \iff \sum_{k=1}^{n} \mathbb{I}(X_k = 0) > \sum_{k=1}^{n} \mathbb{I}(X_k > 0) \]

Yet:

\[ \mathbb{P}(X = 0) > \mathbb{P}(X > 0) \iff e^{-\lambda} > \sum_{k=1}^{+\infty} \frac{-\lambda^k}{k!} \iff 1 > \sum_{k=1}^{+\infty} \frac{\lambda^k}{k!} \iff 1 > \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} - \frac{\lambda^0}{0!} \iff 2 > \sum_{k=0}^{+\infty} \frac{\lambda^k}{k!} \]

\[ \iff 2 > e^\lambda \iff \lambda < \ln 2 \approx 0.6931 \]

Hence, we have a null median for standalone risks for which frequency is computed with a Poisson distribution with parameter \( \lambda < \ln 2 \).

In practice, the scenarios in the scope of operational risk are low frequency / high impact risks. Therefore, the median remains equal to zero in a large majority of the cases encountered on a standalone basis. In all those cases, the SCR is equal to the 99.5%-percentile of the aggregate loss distribution.

2.2 Median is more prudent than the mean

As the median is less sensitive to outliers, deducing the median rather than the mean could be more prudent. Indeed, when the loss is modelled with a heavy-tail distribution, the mean is larger than the median.
In effect, we have seen in the calibration section (chapter 5) that the quantiles of the lognormal distribution were given by: $q_\alpha^{\lognormal} = e^{\sigma q_\alpha^{\lognormal} + \mu}$. The median of the lognormal distribution is equal to the 50%-quantile of the distribution, that is: median = $e^{\sigma q_{50\%}^{\lognormal} + \mu} = e^0 + \mu = e^\mu$. The moments of the lognormal distribution can be computed from the moment generating function of the normal distribution, that is: $\mathbb{E}(e^X) = e^{\mu + \sigma^2/2}$, where $X \sim \mathcal{N}(\mu; \sigma^2)$. Hence: mean = $e^\mu + \sigma^2/2$. Therefore, we have median < mean.

We have seen that the severity of operational risks is mostly modelled with the lognormal distribution, followed by the Pareto and Weibull distributions. Those mathematical proofs hold for the severity distribution. However, we can intuitively understand that the property holds for sums of heavy-tails distributions, for instance for the aggregate loss distribution. Indeed, we have just shown that the median is null when the frequency parameter of the standalone scenario is below ln(2). On the other, when $\lambda > \ln 2$ the statement because there is no closed formula for the median of a random sum of random variables. We leave the floor to the intuition of the reader based on the understanding that the median is less sensitive to extreme values than the mean. We suggest proving this statement thoroughly using the Panjer algorithm or the Fourier transform. However, the result could be noticed with Monte Carlo simulations.

3 Monte Carlo: Study of the convergence of the VaR

3.1 Standalone scenarios

For this study, I wanted to highlight the fact that a Monte Carlo simulation with 500,000 runs enables to reach the convergence of the SCR for the two types of risks that we encounter: high frequency/low severity and low frequency/high severity risks. In order to examine the convergence of the Value-at-Risk, we will compute the SCR value using an increasing number of simulations and then plot the obtained figures: that would enable to graphically determine the number of Monte Carlo runs needed to have a satisfying convergence. Then, we will try to build an estimator for the interval of confidence by computing various times the SCR value obtained with 500,000 simulations. A statistic descriptive table will let us focus on the variance of the sample.

3.1.1 Low Frequency/High severity scenario

First, we will plot the successive values of the SCR during the simulation process. This analysis will keep the values of the SCR computed based on a k x 1,000 simulation strip, with k ranging from 1 to 1,000 (so that the last SCR will be based on a simulation strip obtained with ns=1,000,000 runs). The R code for this function is:

```R
Vpois<-rpois(ns/k,lambda)
X<-sapply(Vpois,function(Vpois){sum(rlnorm(Vpois,mu,sigma))})
SCR=rep(0,k)
SCR[1]=quantile(X,0.995)-EL(X)
```
for (j in 2:pas){
    Vpois<-rpois(ns/k,lambda)
    Xtemp<-sapply(Vpois,function(Vpois){sum(rlnorm(Vpois,mu,sigma))})
    X=c(X,Xtemp)
    SCR[j]=quantile(X,0.995)-EL(X)
    rm(Xtemp)
}

Figure 21: Monte Carlo convergence for a low frequency/high impact scenario is reached after 200 000 runs

In this figure, we can see that the SCR reaches an acceptable level of convergence when the number of simulations is larger than 200,000. Therefore, simulating standalone scenarios with 500,000 scenarios enables to have reached a stable level for the SCR value.

Let’s now simulate 1,000 times the SCR (obtained with 500,000 simulations), and see if the outputs remain close to the mean of the sample.

<table>
<thead>
<tr>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Max</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 898 307</td>
<td>4 982 262</td>
<td>5 007 050</td>
<td>5 030 982</td>
<td>5 114 801</td>
<td>5 006 893</td>
<td>35 139</td>
</tr>
</tbody>
</table>

Table 4: Descriptive statistics of a set of 1,000 simulations for the SCR (computed with 500,000 runs) computed for a low frequency/high impact scenario

We can see that the sample does not deviate too much from its mean, as its coefficient of variation is given by:

\[ cv = \frac{\sigma(SCR)}{\mu(SCR)} = \frac{35 139}{5 006 893} = 0.70\% \]

These results are illustrated on the boxplot below:
3.1.2 High frequency/Low impact

We can see that the convergence is quicker for high frequency/low impact risks than for the previous ones. It comes from the fact that there are more realizations of losses in this case. Indeed, when the frequency is low, most of the Monte Carlo simulations will be equal to zero. Then, the convergence of non-zero realizations takes less time, and therefore less runs. This observation can be confirmed with the following table:

<table>
<thead>
<tr>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Max</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 248 568</td>
<td>3 273 560</td>
<td>3 281 493</td>
<td>3 288 766</td>
<td>3 315 392</td>
<td>3 281 155</td>
<td>11 022</td>
</tr>
</tbody>
</table>

Table 5: Descriptive statistics of a set of 1000 simulations for the SCR (computed with 500,000 runs) computed for a high frequency/low impact scenario

We can see that the standard deviation relatively to the mean of the sample is even smaller than in the low frequency / high impact case:
Quantification of Operational Risks using a Scenario Based Approach

\[ c_v = \frac{\sigma(SCR)}{\mu(SCR)} = \frac{11\,022}{3\,281\,155} = 0.34\% \]

These results are illustrated on the boxplot below:

![Boxplot](image)

Figure 24: Boxplot of a set of 1,000 simulations for the SCR (computed with 500,000 runs) computed for a high frequency/low impact scenario

3.2 Aggregated scenarios

We want to highlight the fact that when conducted at the aggregated level, the SCR convergence is reached quicker than at the standalone level. Therefore, simulations with 100,000 runs would seem reasonable in terms of convergence.

We want to compare the convergence of the standalone SCR to the aggregated SCR. For that, we will compare the convergence seen at standalone level to the one seen at aggregated level. In order to make the comparison possible, we have to consider a fictional entity that is composed of scenarios that have the same characteristics.

The fictional entity has quantified 33 scenarios. All of them have a Poisson-lognormal distribution with \( \lambda=0.5 \), \( \mu=13 \), \( \sigma=1 \). Those 33 scenarios are split into 11 categories, each of those categories containing 3 risk scenarios. Aggregation is done at two levels.

- The first level of the aggregation reflects the benefits from diversification within each category. For all the risk categories, it is assumed that there is a correlation of 75\% between the 3 scenarios that belong to this risk category. For instance the first category being Internal Fraud, there are 3 scenarios that belong to this category and it is assumed that there is a correlation of 75\% between those 3 scenarios.
The second level of the aggregation reflects the effects of diversification across the 11 categories: it is performed with a 11x11 correlation matrix with all coefficients set to 50%\(^{56}\). The overall aggregation process is as follows:

![Diagram illustrating two levels of diversification for the aggregation](image)

Figure 25: Illustration of the two levels of diversification for the aggregation

The standalone SCR behaves as described as follows\(^{57}\):

<table>
<thead>
<tr>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Max</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 774 136</td>
<td>4 951 502</td>
<td>5 006 381</td>
<td>5 060 939</td>
<td>5 249 933</td>
<td>5 006 022</td>
<td>80 397</td>
</tr>
</tbody>
</table>

Table 6: Descriptive statistics of a set of 1000 simulations for the standalone SCR (computed with 100,000 runs)

\[ c_v = \frac{\sigma(SCR)}{\mu(SCR)} = \frac{80 397}{5 006 022} = 1.61\% \]

The aggregated SCR behaves as follows:

\(^{56}\) Those figures were chosen arbitrarily for illustration purpose.

\(^{57}\) Note that the statistics are different from the ones given in the standalone section, because they are computed on a sample of SCR computed with 100,000 runs instead of 500,000 runs in the previous section.
### Table 7: Descriptive statistics of a set of 1000 simulations for the aggregated SCR (computed with 100,000 runs)

<table>
<thead>
<tr>
<th>Minimum</th>
<th>1st Quartile</th>
<th>Median</th>
<th>3rd Quartile</th>
<th>Max</th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>75 980 940</td>
<td>77 645 090</td>
<td>78 203 252</td>
<td>78 647 701</td>
<td>80 548 602</td>
<td>78 160 638</td>
<td>750 838</td>
</tr>
</tbody>
</table>

\[
\frac{\sigma(SCR)}{\mu(SCR)} = \frac{750 838}{78 160 638} = 0.96\%
\]

The reduction of the coefficient of variation between the standalone and the aggregated SCR illustrates the fact that the diversification enables to reduce the volatility of the Monte Carlo simulations and therefore to reach the convergence with a smaller number of Monte Carlo runs.

Moreover, this example illustrates the effect of diversification on the SCR. Indeed, the undiversified SCR is the sum of the standalone SCR of all scenarios. Taking the average value for the standalone SCR, we have \( SCR_{undiv} = 33 \times 5 006 022 = 165 198 726 \). Thus, the effect of diversification is a reduction of more than a half of the economic capital as \( SCR_{div} = 78 160 638 = 47\% \times SCR_{undiv} \).

### 4 Conclusion

Eventually, the robustness of the computation of the capital charge is highly dependent on the convergence of the Monte Carlo simulations. Increasing the number of simulations is always an option but it naturally implies an increase of the computational time, as well as the data storage. The accuracy of the results and the computational speed have to counterbalance each other in order to reach an acceptable compromise. This chapter has graphically showed that, as expected, the convergence is reached more quickly at the aggregate level than on a standalone basis. Practically speaking, this would be in favour of a method that sets a larger number of simulations on a standalone basis than at the aggregated level. The Monte Carlo section also highlighted the capital reduction after proceeding to the aggregation which illustrated the diversification effects.

Concerning the deduction of the expected loss whose sound justification is asked by the Supervisory Authorities, the choice of the median seems to be more prudent. In the context of regulators
demanding very high level of justification, it seems more reasonable to retain the median that remains null most of the time—and therefore saves some of the fastidious work of comparing the expected loss to what has been written in the Profit and Loss attribution. In fact, whatever the purpose, when it comes to summarizing the core of the loss distribution with only one figure, the median should be preferred as it is less sensitive to outliers and therefore more robust than the mean.

The first consideration in this chapter concerned the use of the Value-at-Risk as a risk measure. Yet, not all regulators have chosen the VaR as a risk measure. For instance, Swiss Solvency Test considers the Tail-Value-at-Risk (TVaR) at a 99% confidence level on a one-year horizon. Although its limits have been largely pointed out, among which its absence of subadditivity, other papers have insisted on the robustness of the Value-at-Risk. Cont et al. (2007)\textsuperscript{58} compared to the robustness of the VaR to the robustness of recently proposed alternatives such as the Expected Shortfall (CVaR) and the class of spectral risk measures. In their paper, they “reveal a conflict between robustness and subadditivity” which “shows that one cannot achieve robust estimation while preserving subadditivity”. They also mention some papers that provide reasons for not choosing subadditivity over robustness, against a strict adherence to the coherence axioms of Artzner et al. (1999)\textsuperscript{59}: For instance, Danielsson et al. (2005)\textsuperscript{60} report that VaR is sub-additive in most of their practical applications when Ibragimov and Walden (2001)\textsuperscript{61} state that diversification does not necessarily decrease tail risk and therefore measures for heavy-tailed risks do not require subadditivity.

However, the existence of a threshold under which the median is null certainly exists for extreme quantiles as well. The intuition that the 99.5%-percentile remains null for risks that have a frequency below the one-per-two hundred year threshold could be a serious problem for the sake of robustness. This intuition can be verified when performing sensitivity analyses that will be presented in next chapter. In some way, it could put in doubt the level of confidence chosen by Solvency II for a one-year horizon, especially within a SBA model context, where the focus is on low frequency high impact risks, for which frequency parameter may be set below 0.5%.

\textsuperscript{58} CONT R., DEGUEST R., SCANDOLO G., Robustness and sensitivity analysis of risk measurement procedures, Center of Financial Engineering, Columbia university, Financial Engineering Report No. 2007-6., 2007
\textsuperscript{59} ARTZNER P., DELBAEN F., EBER J.M., HEATH D., Coherent measures of risk, Mathematical Finance, 9, No. 3, 2009, pp.203-228
\textsuperscript{60} DANIELSSON J., JORGENSEN B. N., SARMA M., SAMORODNITSKY G., DE VRIES C. G., Subadditivity re-examined: the case for Value-at-Risk, Working Paper, 2005
CHAPTER 9
ALTERNATIVE APPROACH

Now that the principles of a SBA model have been introduced, the main steps of the quantification presented, their advantages and drawbacks discussed, and alternatives suggested when needed, this last chapter tries to give further analysis on the quantification on operational risk. This part aims to take a step backwards in order to gain a better overall understanding of the model.

In this respect, a way to better comprehend the model is to study its robustness. The robustness of a model is its ability to remain stable against external disturbances. This issue is fundamental in conditions where many factors can vary. Therefore, this chapter aims to test the robustness of the model by conducting sensitivity analyses, whose objective is to study how the uncertainty in the output of the model can be apportioned to different sources of uncertainty in its inputs. In this thesis, sensitivity analyses will be conducted on the frequency and severity parameters as well as on the choice of distribution. They will be treated by the simplest approach that consists of changing one factor at a time (OFAT) and to observe the effect on the output.

Performing those analyses can be useful for a range of purposes that are directly or indirectly close to testing the robustness of the model. For instance, they can evidence some errors in the model, by encountering unexpected relationships between inputs and outputs. When it is observed that some inputs have no effect of the model output, they can also lead to model simplifications by removing some redundant parts of the model. We will see that sensitivity analyses also lead to reduce the model uncertainty: indeed, it can enable to identify which inputs cause significant uncertainty in the output and should therefore be the focus of attention for further analysis, if the robustness of the model was to be increased. In other terms, when some limitations of the model have been identified, some workaround can be performed in order to circumvent the uncertainty and build more stable results. Consequently, performing those studies is mandatory in order to ensure the buy-in of the model by increasing the understanding of the relationships between the inputs and the results.

In this chapter, we will see that there exist some parameters for which the model is highly sensitive. In fact, when studying more in details those situations, we can identify that the model is highly sensitive to cases where the design of the scenario itself is poor. This chapter suggests some workaround to be performed in those cases in order to reduce the model uncertainty.

Still with the objective of injecting a new perspective into a SBA model, this chapter proposes an alternative to the use of available data when feeding the scenario risk assessment. It consists of developing a scaling method that can help the expert to benchmark its measurements by comparing them to actual losses that occurred in another entity of the Group. This technique does not aim to replace the expert judgement which is fundamental in a SBA model. However, it provides the expert with a tool that could be of interest if he has never experienced any similar loss: he could hold on to objective verified data by adjusting the value of the losses so that they gain in comparability.
1 Sensitivity of the model

Discrepancies are inherent to the model. The most important thing is to know where the limits are. For that, sensitivity tests can be conducted in order to determine to what and in which extent the model is the more sensitive. It is notably of interest, for the actuary, to be aware of the sensitivity of the model to changes in its parameters. This section aims to present such tests performed on operational risks modelled with a Poisson-lognormal actuarial approach. They give an insight on the sensitivity to the frequency parameter, the severity parameter, as well as on the sensitivity to a change in the severity distribution. Finally, this section suggests an alternative for helping the expert in the assessment exercise: the objective is to find a way to minimize the sensitivity of the model.

1.1 Sensitivity to frequency parameter

Anyone would expect that the more a risk is likely to occur, the riskier it is for the company. Thus, when the frequency increases, the risk measure should increase too. As it has been discussed before, the model presented in this thesis measures the unexpected loss. Thus, the median of the loss distribution is interpreted as the expected loss and therefore deduced from the 99.5% percentile to compute the SCR. As the frequency increases, the extreme quantiles grow and the expected loss also becomes larger, which could affect the SCR computation. Hopefully, the figure below shows that the increase in the expected loss is weak compared to the highest quantiles’.

![Figure 27: Increase in the 99.5%-percentile and expected loss in terms of lambda](image)

However, specific attention should be put to the metric chosen for computing the expected loss. Indeed, we have seen that the median starts to increase when the frequency parameter is moved from 69% to 70%, because of the existence of a threshold under which the median is null (see chapter 8, section 2.2). For some given severity parameters, the median value can jump which can entail a decrease of the SCR in the surrounding of $\lambda = \ln 2$. This phenomenon is seen when $\sigma$ is
Notwithstanding the SBA view of operational risks focuses on low frequency high impact risks. Therefore, most of the scenarios have their frequency parameter below the threshold. For those scenarios, the sensitivity of their SCR to the frequency parameter is intuitive: the SCR increases as the frequency of the event increases, with the median remaining null therefore the SCR being equal to the 99.5%-percentile of the loss distribution.

The scenarios that are above the threshold belong to the high frequency low severity category of risks. As it has been presented in chapter 6, correlation exists between the frequency and the severity parameters. That correlation is due to the scope of the risks to be quantified. Indeed, risks that are certain to occur and for which the amount of the loss is almost certain are not quantified for the computation of economic capital. In effect, those risks are to be put in provisions in the P&L attribution. In the end, scenarios that have their frequency parameter above the frequency threshold have their parameter \( \sigma \) that tends to be larger and their parameter \( \mu \) smaller. This can be seen in the following graphs that represent the parameters \( \mu \) and \( \sigma \) in function of the frequency parameter \( \lambda \).

Figure 28: \( \mu \) (in ordinate) against \( \lambda \) (in abscissa)

Figure 29: \( \sigma \) (in ordinate) against \( \lambda \) (in abscissa)

However, it is important to recall that those scenarios remain in minority.

This corresponds to risks that have relatively high frequency parameter there is low volatility in the severity distribution.
The extreme percentile are highly influenced by the parameter $\sigma$. It is pretty intuitive: risk is more sensitive to the volatility of a phenomenon than to its mean$^{64}$. Therefore, scenarios that have a large $\sigma$ should have large extreme percentile. Quite the opposite, the expected loss refers to the core of the distribution: it is more influenced by the mean parameter of the distribution. Hence it is expected that the ratio $EL/99.5\%$-percentile for high frequency low impact risks in scope is very low.

Another concern that should be tested in the sensitivity analysis of the frequency parameter is the existence of a threshold for the frequency parameter under which the 99.5\% percentile is null. We proceed similarly to what has been done in the section Expected loss of previous chapter by illustrating the existence of a threshold for the frequency parameter for which the quantile becomes non-null with Monte Carlo simulations.

Similarly to the effect observed on the median, we observe that the 99.5\%-percentile of the aggregate loss distribution is null when the frequency parameter is low. Intuitively, risks that have a frequency below 0.5\% are not expected to occur for the one-year horizon loss with a 99.5\% level of confidence. A zoom on the interval enable to confirm this assumption:

$^{64}$ In the traditional sense of Markowitz theory
Those graphs evidence the fact that the VaR is not robust anymore for scenarios that have a very low frequency parameter. Indeed, similarly to what was observed with the median, a certain combination of the severity parameters might lead to a jump for the VaR if the frequency parameter is moved from 0.4% to 0.5%: the SCR would initially be null and then equal to 5 billion. This issue should be taken into consideration in the scenario quantification.

Moreover, it illustrates one of the weaknesses of Solvency II. In order to resist to a one-in-two hundred years event, the company is to consider only those events that are expected to occur within the horizon of time with the 99.5% level of confidence. In this context, some risks representing a real risk of bankruptcy are totally dismissed from the economic capital computations.

1.2 Sensitivity to severity parameters

Here again, the sensitivity of the SCR to the severity parameters is intuitive: anyone would expect that an increase in both the mean and the volatility parameters entails an increase in the SCR. Indeed, this effect is easily verified. However, the interesting part of conducting those analyses is to understand in which extent does the increase of each of the severity parameter contributes to the increase of the SCR. For this purpose, an interesting view of the problem is to determine what would be the effect of a combined decrease in the severity mean parameter and an increase in the volatility parameter, ceteris paribus.

For instance, let’s take the following example (the frequency parameter being unchanged):

<table>
<thead>
<tr>
<th>Impact (Lognormal)</th>
<th>Capital Charge$^{65}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical</td>
<td>20 000</td>
</tr>
<tr>
<td>Serious</td>
<td>350 000</td>
</tr>
<tr>
<td>Extreme</td>
<td>5 000 000</td>
</tr>
</tbody>
</table>

$^{65}$ Figures are random. They are only given for illustration purpose.
Indeed, it is pretty unintuitive that a decrease in the Typical and Serious impacts would result in an increase in the capital charge. Actually, the explanation of that phenomenon comes from the change in the parameters of the lognormal distribution. As we have seen in the Modelling section, the formula for the calibration of the parameters of the Lognormal is:

\[
\begin{align*}
\mu &= \ln(\text{typical}) + \sigma^2 \\
\sigma &= \sqrt{\left(\frac{q_\beta^N + q_\gamma^N}{2}\right)^2 + 4 \left( \ln \left(\frac{\text{typical}}{\text{extreme}}\right) + \frac{1}{2} \ln \left(\frac{\text{extreme}}{\text{serious}}\right) \right) - q_\beta^N + q_\gamma^N}
\end{align*}
\]

Between those scenarios, there has been:

- A decrease in the lognormal parameter \(\mu\) (because all the impacts have decreased, or remained unchanged)
- An increase in the lognormal parameter \(\sigma\) (because the gap between the mode and the extreme quantiles has increased)

Intuitively, the decrease in the parameter \(\mu\) would result in a decrease of the capital charge whereas an increase in the parameter \(\sigma\) would lead to an increase in the capital charge. When these effects are combined, one cannot guess the change in the capital charge. In fact, it depends on the sensitivity of the capital charge to these shifts of the parameters. In order to illustrate the effect of these changes in the parameters of the distribution on the capital charge, we compute the capital charge deviation by shifting parameters separately:

![Figure 32: Variation of the standalone SCR in terms of the severity parameters](image-url)
It appears clearly that the variation of the parameter $\mu$ has a larger impact than the variation of $\sigma$. In fact, the variation of the quantile is proportional to the exponential of the variation of $\mu$:

$$q_{\gamma}^{\text{log} \mathcal{N}(\mu, \sigma)} = \exp(\sigma \times q_{\gamma}^{\mathcal{N}(\mu, \sigma)} + \mu)$$

$$\Rightarrow q_{\gamma}^{\text{log} \mathcal{N}(\mu+\Delta, \sigma)} = \exp(\sigma \times q_{\gamma}^{\mathcal{N}(\mu, \sigma)} + (\mu + \Delta)) = \exp(\Delta) \times q_{\gamma}^{\text{log} \mathcal{N}(\mu, \sigma)}$$

In this example, $\mu = 12$ and $\Delta = 10\% \times 12 = 1.2$ Therefore, the SCR is multiplied by $\exp(1.2) \approx 3.32$. Nevertheless, a +10% change for $\mu$ is equivalent to all the impacts being multiplied by 3.32. A more reasonable sensitivity test would have consisted (for instance) in shifting all the impacts by +10%, which means a +0.8097% on $\mu$, and leads to an increase of +10% of the SCR.

Thus, in previous example, the variation of $\mu$ is really small, compared to the variation of $\sigma$. This explains why the effect of an increase of $\sigma$ wins compared to the effect of a decrease of $\mu$.

### 1.3 Sensitivity to the choice of distribution for the severity

As it has been discussed, there are some discrepancies that are inherent to the model. Since the mode, $\beta$-quantile and $\gamma$-quantile correspond respectively to the Typical, Serious and Extreme impacts, the parameters (calibrated on the assessed impacts) can enable to perform the reverse exercise. It is possible to compare the impacts assessed by the expert to the ones computed by the model. It gives the discrepancies between the assessed and the calibrated impacts. Usually, the discrepancies remain admissible. Nevertheless, it may also happen that there are huge differences between assessed and calibrated impacts (see table below).

<table>
<thead>
<tr>
<th>Discrepancies with the Lognormal distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical</td>
</tr>
<tr>
<td>0.00%</td>
</tr>
</tbody>
</table>

However, discrepancies depend on the choice of the distribution. Such large discrepancies might indicate that the chosen distribution is not adequate for the scenario. Hence, other distributions might also be considered for modelling this risk: the Pareto distribution and the Weibull distribution should be considered. For this particular example, it turns out that the discrepancies are significantly lower when using the Pareto distribution. This indicates that the Pareto distribution may be a better choice for modelling this specific risk.

<table>
<thead>
<tr>
<th>Discrepancies with the Pareto distribution</th>
<th>Discrepancies with the Weibull distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Typical</td>
<td>Serious</td>
</tr>
<tr>
<td>0.00%</td>
<td>-1.87%</td>
</tr>
</tbody>
</table>

Nevertheless the corresponding SCR are the following:

<table>
<thead>
<tr>
<th>Capital Charge</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lognormal</td>
</tr>
</tbody>
</table>

Confidential - 109
By changing the choice of distribution, the SCR can be multiplied by more than 500. That is un conceivable and underlines the fact that there might be something wrong in the assessment of the scenario. At this stage, we have only focused on the effects of changes in the parameters of the scenario, but we have not seen how the scenario was built. When risk is modelled through a SBA internal model, the design of the scenario itself is of interest.

1.4 Possible alternative: split the scenarios between HFLI and LFHI

To ensure the strength of a SBA model, it is necessary to ensure that the list of scenarios covers all the risks in the scope of operational risk. It is also mandatory to make sure that the design of the scenario itself is appropriate. The relevance of the figures used for the quantification of the scenarios can be back tested using internal data, or external data. We have said that we should be cautious when using historical data in order to cope with the scarcity of the database. The actual database reflects the “real life” and its intrinsic imperfection: it can be misleading sometimes.

Let’s take the example of a scenario that has the following historical data available for this scenario and draw our conclusions:

<table>
<thead>
<tr>
<th>Mean</th>
<th>Sd</th>
<th>Min</th>
<th>First quartile</th>
<th>Median</th>
<th>Third quartile</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>325 533</td>
<td>461 567</td>
<td>0</td>
<td>3 000</td>
<td>20 000</td>
<td>150 000</td>
<td>7 000 000</td>
</tr>
</tbody>
</table>

With these basic statistics on the historical data available, we can see that extreme losses have a huge weight in the mean of the sample. Thus, the standard deviation is high, which lets us think that the sample is not homogeneous. There might be a bias in this sample coming from the fact that small losses might not be detected as often as big ones by internal controls. Hence, major losses might tend to have a higher probability to be detected and consequently to be recorded in the internal loss...
database. If used to help the expert in the risk assessment exercise, the heterogeneity of the data might lead to a precarious scenario design. This might be evidenced by large discrepancies between the assessed and calibrated impacts. In fact, large discrepancies may be observed when the serious and extreme impacts (that should correspond to the extreme quantiles) are too far from the typical (that should correspond to the mode of the distribution).

In order to have more homogeneity and therefore less discrepancies between the assessed and calibrated impacts, the scenario could be split into two scenarios that would have the same underlying story but different scopes:

- High frequency/Low impact
- Low frequency/High impact

For instance, an internal fraud scenario might be split into one scenario of frequent but not severe frauds -by tied agents- and one scenario of rare and extreme internal fraud events -similar to the Kerviel's affair. Those two scenarios would not be assessed and quantified the same way. For instance, the low frequency / high impact scenario could be modelled with different (and more heavy-tailed) distributions, such as the Pareto instead of the lognormal. In order to split the scenarios, using well-known Extreme Value Theory techniques could be of interest. For instance, graphical methods such as the mean excess plot\(^66\) or the Hill plot could help to decide which threshold to choose. For a fixed threshold, the value of the Mean Excess Function is the mean of the losses which exceed this threshold, but losses which we have removed the threshold before. In other words, it's the mean of the excess. We can plot the graph of this function and the objective will be to find a “break” in its evolution such that after this break (we look by increasing values of thresholds), we observe some kind of linearity. In our particular example, the lack of available data does not enable to decide exactly the level of the threshold. Nevertheless, it gives an idea: the threshold might be set to 1,000,000.

![Figure 33: An illustration of the Mean Excess Plot](image)

However, drawing such a conclusion based on very few observations is not very reasonable. This illustrates once again the problem of the lack of historical data in operational risk and reinforces the cleverness of a SBA model. However, when performing the assessment of scenarios, experts tend to try to rely on historical objective verifiable data.

\(^{66}\) The mean excess plot is a tool widely used in the study of extreme values.
It is understandable that when assessing the scenarios, the expert feels the need to hold on to some historical data in order to benchmark its opinion and make sure he does not over or under estimate the possible impact of the event. As the recourse to internal data can be a problem sometimes, an alternative for the benchmarking exercise would be to use scaled external data. Next section provides an overview on scaling techniques and offers a model for scaling internal loss data from one entity to another.

2 Benchmarking: scaling methods

To meet the Solvency II standard of estimating two hundred-year losses, the analysts must bring as much relevant data as possible into the estimation process. Basel II agreement insists on the use of four data elements, one of them being external data. It is clear that external data cannot be directly used in an internal model, because not all companies are comparable. Thus it is pretty intuitive that losses that occur in a large American banking company are likely to be larger than the ones that would occur in a small French union. Hence, some treatment has to be considered in order to make projections for a particular insurance company based on the losses recorded in the banking and insurance industry. The analyst typically has two methods available for dealing with the heterogeneity of loss databases: selecting and pooling relevant data within and across categories, and scaling the data using statistical models to make the data more representative of the types of losses experienced by their institution.

This section presents succinctly the scaling method developed by Dahen and Dionne (2008) that inspired our approach. For a deeper overview of the different approaches available in the literature, the reader should refer to Dahen et al. (2008) and Cope et al. (2008).

2.1 Literature review

The relationship between the size of operational losses and some exposure factors using regression models has already been treated in the banking industry. The studies fit linear models of log-losses onto some exposure factors that are linked to the size of the company. In those models, the losses are interpreted as the product of a “common” and an “idiosyncratic” component. The first component is common to all the banks and reflects macroeconomic environment whereas the “idiosyncratic” component captures the behaviour of losses that have specific characteristics, related to the bank in which they occur, the country, the business line or the type of risk. In the end, the models assume that the losses are such that:

\[ \text{loss}_i = f(\text{Component}_{\text{idiosyncratic}}) \times \text{Component}_{\text{common}} \]

where \( \text{loss}_i \) is a loss that is characterised by the bank, the geographical region, the business line, the event type in which it occurs.

---

67 DAHEN Hela, DIONNE Georges, Scaling Models for the Severity and Frequency of External Operational Loss Data, January 2008

68 COPE Eric and LABBI Abderrahim, Operational Loss Scaling by Exposure Indicators : Evidence from the ORX Database, October 2008
Component_{idiosyncratic} is a scaling factor that depends on the bank, geographical region, business line, event type or other exposure factors.

Component_{common} is the common component: the baseline loss value that corresponds to the exponential of the intercept of the regression.

Several variables can be considered for explaining the severity of losses. It is likely that the size of losses may be correlated with the size of the company. Indeed, Shih et al. (2000)\(^{69}\) showed that the institution size is the main scaling factor: there is a non-linear relation between operational losses and the size of the company. They considered a regression model based on the logarithm of the size of the company, explained in terms of logarithm of assets:

\[
\log Y_i = a_o + a_1 \log A_i + e_i
\]

where \( Y_i \) is the loss amount

\( A_i \) denotes the assets of the banks as the indicator of the firm size

\( e_i \) is the error term that appears to increase with the firm size and encouraged the authors to use a weighted least squares fitting procedure in order to correct the heteroscedasticity

\( a_o, a_1 \) are the coefficients to fit, with \( a_o = Component_{common} \) being the common component and \( a_1 = Component_{idiosyncratic} \) being the idiosyncratic component is the so-called scaling factor.

After fitting values to the parameters Component_{idiosyncratic}, the model can be used to scale losses from one firm \( i \) to another \( j \) with the following relation:

\[
\begin{align*}
\log Y_i &\approx a_o + a_1 \log A_i \\
\log Y_j &\approx a_o + a_1 \log A_j
\end{align*}
\Rightarrow \frac{Y_j}{A_j a_1 j} \approx \frac{Y_i}{A_i a_1 i} \approx e^{a_o} = Component_{common}
\]

Total revenue which is an indicator of the firm size is the only risk factor considered in this model. Nevertheless, the results of this study show that the total assets only explain 5% of the variance. Indeed, it is expected that other factors should be considered to explain the size of the losses.

To improve this model, Dahen and Dionne (2008) introduced scaling factors other than firm size: the geographical region, the business line and the risk type. They decided to keep the power relationship between the loss amount and the firm size and added some factors, so that:

\[
\log Y_i = a_o + a_1 \log A_i + \sum_{j\geq 2} a_j E_{ij} + e_i
\]

\(^{69}\) SHIH J., SAMAD-KHAN A., MEDAPA P., Is the size of an operational loss related to firm size?, Operational Risk Magazine 2,1, 2000
http://www.opriskadvisory.com/docs/is_the_Size_of_an_Operational_Loss_Related_to_Firm_Size_(Jan_00).pdf (visited on 5 Sept 2014)
where $Y_i$ is the loss amount

$A_i$ denotes the assets of the banks as the indicator of the firm size

$EF^j_i$ is the value of the j-th exposure factor for loss $i$.

To estimate the parameters, they used the Ordinary Least Squares (OLS). With this method, the adjusted $R^2$ is around 10%. They report that though this value is low, it is better than the 5% of Shih et al. (2000) and recall how difficult it is to capture certain non-observable factors related to the internal controls that are not present in the database and therefor difficult to quantify.

2.2 Scaling Internal data

A similar approach is presented below. It aims to scale losses that occurred in different entities of the Group to make them comparable.

In this section, we will proceed as follows:

- study of the correlation between some possible candidates for the exposure factors and the size of the los-losses: only those with a p-value smaller than 5% are retained for the regression;
- robustness tests for the retained variables: only those that have a p-value smaller than 5% will be used as explanatory variables in the regression model;
- validation of the model by scaling all the losses on one given entity for one losses with specific characteristics.

Practically speaking, the outcomes of scaling studies highly depend on choices made by the modeller, whether it comes from the choice of a regression model, or the choice of preliminary treatment to be applied to filter the data.

2.2.1 Presentation of our data

The statistician has a lot of preparatory work to do before scaling its data. In our study, the internal loss database is made of almost 40 000 losses. Some preparatory work was performed to clean the data:

- Delete all losses with a financial impact smaller than 1. The purpose is to work on log-losses therefore losses have to be larger than 1.
- Delete losses that occurred prior to 2010. As the control environment evolves significantly with time, only recent years results should be retained for the study.
- Reduce the scope of the study to the major entities. Some entities are not mature enough to contribute correctly to the loss database. There losses may be outliers and may prevent the regression from capturing the right phenomenon. Moreover, only insurance companies are retained. In this study, it has been decided to keep entities that have reported more than 50 losses.
- Investigate for size indicators variables. In most of the regression models, total assets are used as size indicators. In this model, it has been decided to use a combination of other
Quantification of Operational Risks using a Scenario Based Approach

size indicators such as the number of employees, the amount of provisions, the total assets and the total liabilities.

The scaling consists of doing a regression on the size of losses in terms of other variables. In the internal loss database, we have access to the following information:

- entity
- location
- business line
- event type
- date
- description
- gross loss
- net loss

In order to be consistent with what has been done in the literature, we will work with the gross loss amounts. Qualitative information has to be translated into quantitative variables via dichotomisation. For the sake of simplicity, it has been decided to keep a limited number of variables that were considered to best capture the information contained in the database:

- location: each location is associated with one of the following variables: Asia, Europe, USA.
- business lines: events are split between insurance, investment management and banking.
- event types: it has been decided to work with Basel II classification in order to be consistent with what was done in the literature. Therefore, events are split into 7 categories called event types. Those variables are introduced as dichotomous variables. 7 variables are created for capturing the event type effect: Internal Fraud (IF), External Fraud (EF), Employment Practices and Workplace Safety (EPWS), Clients Product and Business Practices (CPBP), Damage to Physical Assets (DPA), Business Disruption and System Failures (BDSF) and Execution Delivery and Project Management (EDPM).

However, the description of the event has been disregarded. Any ‘big-data’ specialist could have tried to draw something from this variable by introducing a set of key words and connected losses that contain those key words in their description. In the end, 17 variables are created. The correlation between those 17 variables is given in table 8.
As expected, the size variables are highly correlated. The relation between those factors and the loss size has to be tested in order to retain only representative variables in the study. As shown on table 9, the linear correlations between the size of the losses and the explanatory variables are always close to zero. However, the linear correlations between the explanatory variables and the log-losses are far larger. Indeed, this result was confirmed by Pearson’s test that gave very low p-values for all of the variables. There exists undoubtedly a relation between the log of the size of operational losses and the size of the company, region, business lines, event type and date.

Table 8: Linear correlations between the explanatory variables

Table 9: Pearson’s correlation coefficient between the losses and the explanatory variables

Table 10: Spearman’s correlation coefficient between the losses and the explanatory variables

The regions and size of the firm are the variables most correlated with the log-loss size. In addition, the size variables are highly correlated with log-loss size, with the number of employees showing the
strongest relationship. Moreover, the Spearman coefficients being very close to the Pearson’s could indicate that the relation is almost linear. Shih et al. (2000) stated that “the logarithm of the scale variables showed a stronger relationship to [log]-losses than did the raw variables”, which is not the case for our sample. Therefore, we will not use the same model to perform the regression.

2.2.2 Choosing a regression model

The regression model implemented for internal losses was based on the method developed by Dahen and Dionne (2008). After studying the levels of correlation, the exposure factors that we retained in our study were the following:

- geographical region
- size of the entity
- event type of the loss
- business line in which the loss occurred
- year of the event

In most of the regression models, the logarithm of the total assets is used as size indicator. In this model, it has been decided to use a raw combination of other size indicators such as the number of employees, the amount of provisions, the total assets and the total liabilities. As stated in Dahen and Dionne, these size variables are correlated, but our study shows that all of those variables are non-null significant, which encouraged us to keep all those indicators.

The significance of the variables will be tested on the following regression model:

\[
\log Y_i = a_o + \sum_{j=1}^{1} a_j EF_j^i + e_i
\]

where \( Y_i \) is the loss amount

\( EF_j^i \) is the value of the \( j \)th exposure factor (among the above 17 variables) for loss \( i \).

2.2.3 Regression model

This section presents the robustness tests for the exposure factors. For each exposure factor introduced in the regression variable, the table gives the corresponding coefficient and the confidence level of significance:

- *** coefficient significant at the 99% confidence level
- ** coefficient significant at the 95% confidence level
- * coefficient significant at the 90% confidence level

The first table gives the significance of the exposure factors. First model (column 1) was designed with only region variables: it enables to explain 38% of the overall variance. The variable USA is a linear combination of the other region variables therefore it will be withdrawn from the analysis. The next models (columns 2-3-4-5) test which of the size factors best explains the size of the log-losses: those 4 models have coefficient of determination between 0.7% and 8.2%. Combining those size factors (column 6) enables to improve the percentage of variance explained by the model to 31.1%
with all factors being significant: this will encourage us to keep all those size exposure factors in the regression. The event type variables explain 7.1% of the overall variance (column 7), with event type 7 withdrawn from the analysis as a linear combination of the other event types. The variable year has a R² of 0.3% which is definitely low (column 8), but the variable remains significant at the 99% confidence level. Business Lines variables have a coefficient of determination of 0.1%, while being poorly significant (column 9). We will see that those variables will not be retained as explanatory variables.

Then, we introduce one by one the explanatory variables in the analysis and observe how their combination enables to improve the regression model.

Introducing the size variables in addition to the region variables (column 1) increases the coefficient of determination to 41.7%. With the event type, the adjusted R² reaches 43%, but some event types are not significant (event types 1 (IF), 2 (EF) and 5(DPA)): they will be withdrawn from the analysis for the final regression model. It still increases a little with the introduction of the year variable (R² adjusted = 43.6%). As expected, the business lines do not give anything interesting: the R² adjusted decreases and both insurance and investment variables are not significant. In the end, we performed the regression with only variables significant at the 95% confidence level and withdrawn the variable Year (column 5). The model explains 43% of the overall variance with 9 explanatory variables.

In the end, the retained model in the last column of the table (column 5). The variable Year was withdrawn from the analysis because it gives a very high value for the constant, which could introduce a lot of variability in the scaling model. With those assumptions, the scaling model has 9 variables that explain around 43% of the overall variance, which is considerably higher than in Dahen and Dionne’s paper.

Hence, our regression model is the following:

\[ \log Y_i = a_0 + \sum_{j \geq 1} a_j EF_j + e_i \]

With the figures given in the table below, that is:

\[ \log Y_i = 7.059 - 2.440 \times Asia + 1.984 \times Europe \\
+ 3.15 \times 10^{-4} \times HR - 8.09 \times 10^{-4} \times Equity \\
- 1.16 \times 10^{-5} \times Liabilities + 2.94 \times 10^{-3} \times Provisions \\
+ 0.678 \times EPWS + 0.698 \times CPBP + 1.180 \times BDSF \]
Quotation of Operational Risks using a Scenario Based Approach

<table>
<thead>
<tr>
<th>Region+Size</th>
<th>Region + Size + EventType</th>
<th>Region + Size + EventType + Year</th>
<th>Region + Size + EventType + Year + BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7,013</td>
<td>7,21</td>
<td>435,9</td>
<td>425,5</td>
</tr>
<tr>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Asia</td>
<td>-2,336</td>
<td>-2,588</td>
<td>-2,625</td>
</tr>
<tr>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Europe</td>
<td>2,255</td>
<td>1,91</td>
<td>2,009</td>
</tr>
<tr>
<td>***</td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>USA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>HR</td>
<td>3,31E-04</td>
<td>3,08E-04</td>
<td>3,22E-04</td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Equity</td>
<td>-9,43E-04</td>
<td>-8,11E-04</td>
<td>-8,42E-04</td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Liabilities</td>
<td>-7,14E-06</td>
<td>-1,00E-05</td>
<td>-9,83E-06</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Provisions</td>
<td>3,05E-03</td>
<td>2,79E-03</td>
<td>2,86E-03</td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Event Type1</td>
<td>-0,2021</td>
<td>-0,2931</td>
<td>-0,2979</td>
</tr>
<tr>
<td>Event Type2</td>
<td>-0,1701</td>
<td>-0,1422</td>
<td>-0,2176</td>
</tr>
<tr>
<td>Event Type3</td>
<td>0,6291</td>
<td>0,4916</td>
<td>0,49</td>
</tr>
<tr>
<td></td>
<td>**</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>Event Type4</td>
<td>0,6775</td>
<td>0,6128</td>
<td>0,6027</td>
</tr>
<tr>
<td></td>
<td>***</td>
<td>***</td>
<td>***</td>
</tr>
<tr>
<td>Event Type5</td>
<td>0,5933</td>
<td>0,5167</td>
<td>0,5192</td>
</tr>
<tr>
<td>Event Type6</td>
<td>1,156</td>
<td>1,175</td>
<td>1,172</td>
</tr>
<tr>
<td>Event Type7</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Investment</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>41,7%</td>
<td>43,2%</td>
<td>43,8%</td>
</tr>
<tr>
<td>R² adj</td>
<td>41,6%</td>
<td>43,0%</td>
<td>43,6%</td>
</tr>
</tbody>
</table>

Table 11: Robustness tests by introducing step by step all of the variables

Confidential - 119
2.2.4 Validation

In Dahen and Dionne’s paper, they apply the scaling to the whole database so that they obtain the equivalent losses of the full set of losses adjusted to one bank in order to concretize the scaling model. It means that they calculate the loss amount that would suffer another bank that has different size characteristic and operates in a different region for the same type of risk, in the same year than the one from the loss data.

To evaluate the accuracy of our model, scaled losses can be compared to those that have really occurred in this given entity. Dahen and Dionne compared the losses occurred in Merrill Lynch to the ones scaled to Merrill Lynch from the whole external loss database. Their results enable to conclude that their method is satisfying. In their case, they are working with Fitch’s OpVaR database, which is an external database where operational losses over USD1,000,000 are reported. Thus, only major losses are recorded in this database which makes them similar enough to be compared. The largest loss in the database is USD16 billion, which is $10^5$ times higher than the minimum loss.

With internal data, losses range from 1 euro to hundreds of millions of euros, that is $10^9$ times higher than the minimum loss. Thus, losses are very different intrinsically, therefore not easily comparable. In practice, we proceeded to the scaling in a different way: the whole database is scaled on one entity that operates in one country for losses of one given event type that occurred during one given year.

For a loss that occurred in region $i_A$, entity $j_A$, that belongs to the event type $k_A$ and during the year $l_A$, we have:

$$\text{Component}_{\text{common}} \approx \frac{\text{loss}_{i_A, j_A, k_A, l_A}}{f \left( \text{Component}_{\text{idiosyncratic}}(i_A, j_A, k_A, l_A) \right)}, \quad \forall i_A, j_A, k_A, l_A$$

Once the coefficients of the explanatory variables have been computed by the regression, we can compute the idiosyncratic component for each loss recorded in the database: it is the intercept of the regression (i.e. the constant). To scale this loss to another region, in another entity that belongs to another type of event in another year, we assume that the unexplained part of the regression model (that depends on unobservable factors such as internal controls) are the same for every bank and proceed with the following normalization formula:

$$\text{loss}_{i_B, j_B, k_B, l_B} = \frac{f \left( \text{Component}_{\text{idiosyncratic}}(i_B, j_B, k_B, l_B) \right)}{f \left( \text{Component}_{\text{idiosyncratic}}(i_A, j_A, k_A, l_A) \right)} \text{loss}_{i_A, j_A, k_A, l_A}$$

where $f \left( \text{Component}_{\text{idiosyncratic}}(i, j, k, l) \right) = \exp \left( \sum_{i, j, k, l} a_{i, j, k, l} E F_{i, j, k, l} \right)$

To compare the scaled losses to the actual losses, we compared the median from both sample for each entity and each event type. The normalized results give the ratio:

$$\frac{\text{median(scaled losses)}}{\text{median(actual losses)}}$$
The average ratio is 1.08 with a standard deviation of 1.15. The purpose of the scaling is to provide a method for benchmarking the losses occurred in another entity in order to better perform the scenario assessment. For instance, an expert working for an entity that has never measured any similar scenario before could use the scaling method to scale losses to have an idea of the size of the losses that could occur in his entity. In this context, the aim of the study is to give order of magnitude more than an exact figure for the scaled loss. Therefore, as long as the ratio is not larger than 10 or smaller than 0.1, the scaling model could be considered satisfying.

However, results are not satisfying for entity 5. Indeed, the model is fairly accurate for event type 7 losses but it largely underestimates losses that belong to the event types 1 to 6. In fact, this entity widely contributes to the loss database (around 30% of all losses were reported by this entity) and it reports mostly small losses of type Execution Delivery and Project Management (EDPM). Moreover, it is the only entity located in Asia. When looking at the coefficient of the model, we can see that the coefficient of variable Asia is largely negative. This indicates that the model has interpreted that losses occurring in Asia would always be low. Even though it is the case for EDPM losses, the other types of losses reported by this Asian entity are relatively a lot more severe. However, the scaling will always give small values to losses occurring in Asia. As it is true for a certain type of risk, which represents the majority of risks reported by this entity, the variable region explains a large proportion of the overall variance. But we can understand that this is biased because not all entities have the same reporting threshold.
In effect, we have seen that the coefficient of determination $R^2$ is way larger in our study than in what can be found in the literature. However, when only focusing on losses above a certain threshold\(^7\), the coefficient of determination drops significantly: the adjusted $R^2$ even becomes negative for a threshold set at 1,000,000€. This clearly alters on the soundness of the scaling model. Therefore, the scaling should be subject to further analysis in order to possibly integrate this technique as part of the methodology for the measurement operational risk scenarios.

### 3 Conclusion

This chapter intended to be more open than the others in this thesis. In fact, the objective was to highlight the fact that there are limits to the model as well as limits to the regulation. Indeed, illustrations have evidenced that the VaR is not robust anymore for scenarios that have a very low frequency parameter. Indeed, similarly to what was observed with the median, a certain combination of the severity parameters might lead to a jump for the VaR if the frequency parameter is moved from 0.4% to 0.5%. Nevertheless, if those risks are sufficiently numerous, they may cause a large increase of the aggregated SCR because of the non-subadditivity of the Value-at-Risk.

However, Solvency II only considers taking into account the one-in-two hundred years event. Indeed, it illustrates one of the weaknesses of Solvency II. In order to resist to a one-in-two hundred years event, the company is to consider only those events that are expected to occur within the horizon of time with the 99.5% level of confidence. In this context, the risk that has a frequency set to 0.4% should not be taken into account. In this context, some risks representing a real risk of bankruptcy are totally dismissed from the economic capital computations.

The sensitivity analyses presented in this chapter also highlighted the fact that the model might be very sensitive to change in the severity parameters and distribution. This example enables show the wide range of purposes of the sensitivity analyses. In addition to the identification of inputs that may cause significant uncertainty in the results, the study leaded to notice the importance of the design of the scenarios. Indeed, it is pretty intuitive that the accuracy of the design of the scenario itself highly contributes to the model robustness in a Scenario Based Approach. This observation enabled to focus of attention on those shaky scenarios in order to perform some workaround and build more stable results. The first alternative was to consider splitting those scenarios between a high frequency low impact and a low frequency high impact scenario.

Another alternative to make those precarious scenarios stronger is to propose a tool that can help the expert in the risk assessment exercise. In this context, this chapter has suggested a scaling method that could be used to make internal data more comparable from one entity to another. The method is based on the analysis that the size of the log-losses is highly correlated with the

\(^{7}\) This study has been conducted with the following thresholds: 10,000€; 20,000€; 100,000€; 500,000€; 1,000,000€.
geographical region and size of the company where the loss occurred. The regression model aims to capture the characteristics of the occurred losses that can be forecasted by some explanatory variables. Nevertheless, the results of the validation of the model put in doubt the soundness of the model.

If adequately calibrated, scaling methods could be used to calibrate an operational risk internal model based on the Loss Distribution Approach. Indeed, scaling external data enables to get information on the tail of the distribution that is not included in a company’s internal loss database. Nevertheless, scaled external data should not be used as it is. Chappelle et al. (2005)\textsuperscript{71} proposed a manner to combine internal and external data to compute an overall loss distribution for a given business line and a given event type category. In this model, external data is only used to model losses above a high threshold, while internal data is used for the core of the distribution.

In a SBA model, external loss data are used in the scenario identification process to back test internal exposure to operational risk with what occurred in the other entities and check whether similar situations are relevant for a given entity in a given context. It is understandable that when assessing the scenarios, the expert feels the need to hold on to some historical data in order to benchmark its opinion and make sure he does not over or underestimate the possible impact of the event.

With the development of the external database ORX to insurance companies, it is expected that external data will be more utilized in the future by insurers modelling operational risks through an internal model, whether it is via a LDA or a SBA. Therefore this scaling approach could be implemented for an external database. Moreover, some further analysis could be conducted in order to implement a scaling model for the frequency.

\textsuperscript{71} CHAPELLE A. et al., Measuring and Managing Operational Risk in the Financial Sector: An Integrated Framework, February 2005
CONCLUSION

This thesis has described one of the three main approaches employed for quantifying operational risk with an internal model. The aim of this thesis has been to scrutinize the steps of the quantification process in a SBA model in order to discourse about the robustness of the model. With the objective of injecting a new perspective into a SBA model, alternatives that may enable to circumvent some of the hurdles encountered have been presented throughout the development.

When examining closely the quantification process, the angle of this thesis has been to determine to which extent actuarial methods can be transposed into a more qualitative model that is essentially motioned through expert judgment, although a great of objective data (internal as well as external historical data) is used to feed the risk assessment. When the use of those actuarial methods has been considered not to be feasible in the Scenario Based Approach (e.g. use of Student copula for the aggregation), we have tried to give intuition about the underlying reasons why this given method was not appropriate for our problem. As it has been discussed throughout this thesis, the strength of a SBA internal model comes from the fact that it is a forward-looking approach that adequately represents the risk profile of the company. Therefore, it is more likely to capture the risks that the insurance company might encounter in the coming year.

This thesis has also tried to give possible alternatives to practical issues that could have arisen when performing the quantification of operational risks within a risk management team. For instance, it has suggested solutions to day-to-day problems (e.g. non-convergence of Cholesky algorithm when performing the aggregation) as well as possible alternative for the calibration of lognormal parameters (e.g. minimization of the inter-quantile distance). It has also mentioned some aspects of the model sensitivity and highlighted the importance of the design of the scenario itself. In this context, it has suggested a first approach for a scaling method that could be used in the context of a SBA model that is to say in order to help the experts to benchmark their measurements in the assessment of risk scenarios. In the end, this thesis has intended to take a step backwards in order to gain a better overall understanding of the model.

Indeed, each of the Advanced Measurement Approach techniques has its own advantage. For instance, quantifying operational risks with a Loss Distribution Approach reduces the level of subjectivity of the model by relying on objective and verifiable information. As for the Bayesian Approach, it enables to trace the causality links between events which leads to a better understanding of the possible interactions that exist between risks and therefore a pertinent modelling of correlations. Nevertheless, some of their drawbacks should also be raised. As the loss Distribution Approach depends on historical data, it is a backward-looking approach: it does not take into account the effect of a change in the economic environment or the level of internal controls put in place within the company. Most of the criticisms of Bayesian Approach and SBA come back to the question of objectivity. Indeed, both approaches highly rely on expert judgment.

After all, the existence of some flaws in the models is inherent to their nature. Indeed, a model is never perfect, or totally representative of the reality. In Operational risk, the major stake is to succeed in finding a compromise between the necessary pragmatism of the approach and the level
of abstraction of the mathematical model. Indeed, a model that quantifies operational risks cannot be fully disconnected from the operational risk management teams that put in place controls and actions to manage and mitigate the identified risks.

In many cases, the quality of a scientific field depends on how well the mathematical models developed on the theoretical side agree with observations of repeatable experiments. Lack of agreement between theoretical mathematical models and experimental measurements often leads to important advances as better theories are developed. This is the reason why insurance companies have to be aware of the imperfections of their model so that they make effort in order to perpetually try to find ameliorations and alternatives that may improve their model.
APPENDIX

Proof of median < mean

Pareto distribution

We have seen in the calibration section (chapter 5) that the quantiles of the Pareto distribution were given by: \( q_\alpha^{\text{Pareto}} = x_m(1 - \alpha)^{-\frac{1}{k}} \). The median of the Pareto distribution is equal to the 50%-quantile of the distribution, that is: median = \( x_m \sqrt[2]{k} \).

The Pareto distribution has the following density function: \( f(x) = \frac{kx_m^k}{x^{k+1}}\mathbb{I}\{x > x_m\} \). That gives:

\[
\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f(x) dx = \int_{x_m}^{+\infty} k \left(\frac{x_m}{x}\right)^k dx = k x_m^k \int_{x_m}^{+\infty} x^{-k} dx = \begin{cases} k x_m^{\frac{k}{k-1}} & \text{if } k > 1 \\ +\infty & \text{if } k < 1 \end{cases}
\]

Therefore, the mean of the Pareto distribution is \( \mathbb{E}(X) = \begin{cases} x_m \frac{k}{k-1} & \text{if } k > 1 \\ +\infty & \text{if } k < 1 \end{cases} \)

Let’s prove that \( \forall k > 0 : \text{median} < \text{mean} \)

- For \( k < 1 \): median < mean = +\infty
- For \( k > 1 \):

\[
\text{median} < \text{mean} \quad \iff \quad x_m \sqrt[2]{k} < x_m \frac{k}{k-1}
\]

\[
\iff 2 < \left(\frac{k}{k-1}\right)^k
\]

\[
\iff 2 < \left(1 + \frac{1}{k-1}\right)^k
\]

Let \( f(k) = \left(1 + \frac{1}{k-1}\right)^k \)

We have \( f'(k) = k \left(1 + \frac{1}{k-1}\right)^{k-1} \cdot \left(-\frac{1}{(k-1)^2}\right) < 0 \)

Therefore, \( f \) is a decreasing function, and \( \forall k \in ]0; 1[ : f(k) > \lim_{x \to +\infty} f(x) \)

Furthermore, we know that: \( e = \exp (1) = \lim_{k \to +\infty} \left(1 + \frac{1}{k}\right)^k = \lim_{k \to +\infty} \left(1 + \frac{1}{k-1}\right)^{k-1} \)
Hence, \[ \lim_{k \to +\infty} f(k) = \lim_{k \to +\infty} \left(1 + \frac{1}{k-1}\right)^k \]
\[ = \lim_{k \to +\infty} \left(1 + \frac{1}{k-1}\right)^{k-1} \times \left(1 + \frac{1}{k-1}\right) \]
\[ = \lim_{k \to +\infty} \left(1 + \frac{1}{k}\right)^k \times \left(1 + \frac{1}{k}\right) \]

Yet, \[
\begin{aligned}
\lim_{k \to +\infty} \left(1 + \frac{1}{k}\right)^k &= e \\
\lim_{k \to +\infty} \left(1 + \frac{1}{k}\right)^k &= 1
\end{aligned}
\]

Hence, \[ \lim_{k \to +\infty} f(k) = e \]

Finally, it comes: \[ \forall k \in ]0; 1[ : f(k) > \lim_{x \to +\infty} f(x) > 2 \]

In conclusion, \( \forall k > 0 : \text{median} < \text{mean} \)

**Weibull distribution**

The Weibull distribution has the following density:
\[
f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} e^{-\left(\frac{x}{\lambda}\right)^k} \mathbb{I}\{x \geq 0\}.
\]

We have seen in the calibration section that we would only consider the Weibull distribution when it belongs to the sub-exponential class, that is when the shape parameter \( k < 1 \).

We have seen in the calibration section (chapter 5) that the quantiles of the Weibull distribution were given by: \( q_{\alpha}^{\text{Weibull}} = \lambda(-\ln(1 - \alpha))^{1/k} \). The median of the Weibull distribution is equal to the 50%-quantile of the distribution, that is: median = \( \lambda(\ln 2)^{1/k} \).

The moment generating function of a Weibull distributed random variable\(^{72} \) is given by \( \mathbb{E}(e^{t\ln X}) = \lambda^t \Gamma\left(1 + \frac{1}{k}\right) \). In particular, the mean of the Weibull is \( \mathbb{E}(X) = \lambda \Gamma\left(1 + \frac{1}{k}\right) \)

\[ \text{median} < \text{mean} \quad \Leftrightarrow (\ln 2)^{\frac{1}{k}} < \Gamma\left(1 + \frac{1}{k}\right), \quad \forall \ 0 < k < 1 \]
\[ \Leftrightarrow (\ln 2)^{x} < \Gamma(1 + x), \quad \forall \ x > 1 \]

We have: \( \ln 2 < 1 \Rightarrow (\ln 2)^{x} \) decreases with \( x \). Hence: \( (\ln 2)^{x} < \ln 2 < 1 \)

The Gamma function is defined as: 

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt.$$

We have: 

$$\Gamma(1 + x) = \int_0^{+\infty} t^x e^{-t} dt.$$

One of the famous properties of the Gamma function is that $$\Gamma(1 + x) = \Gamma(x)^73$$. Especially, for integers, we have: $$\Gamma(n) = (n - 1)!$$

It can be shown also that $$\Gamma^{(k)}(x) = \int_0^{+\infty} (\ln t)^k t^{x-1} e^{-t} dt$$.

In particular, $$\Gamma''(x) = \int_0^{+\infty} (\ln t)^2 t^x e^{-t} dt > 0$$ when $$x > 0$$, which implies that $$\Gamma''$$ is a strictly increasing function and can only be equal to zero once when $$x > 0$$. Yet, we have $$\Gamma(2) = \Gamma(1 + 1) = 1 \times \Gamma(1) = \Gamma(1) = (2 - 1)! = 1$$. By Rolle theorem, we know that there exist $$x$$ in $$]1, 2[$$ such that $$\Gamma'(x) = 0$$. This is the only value for which $$\Gamma'$$ is equal to zero when $$x > 0$$ (as it is strictly increasing). Hence, $$\Gamma$$ is strictly increasing on $$]2; +\infty[$$. In other terms, we have: $$\Gamma(x) > 1 > \ln 2 > (\ln 2)^x, \forall x > 1$$. This concludes the demonstration.

---

73 This property can be shown using integration by parts
# Quantification of Operational Risks using a Scenario Based Approach

<table>
<thead>
<tr>
<th>Region</th>
<th>HR</th>
<th>Equity</th>
<th>Liabilities</th>
<th>Provisions</th>
<th>All Size Indicators</th>
<th>Event Type</th>
<th>Year</th>
<th>BL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>9,355</td>
<td>6,943</td>
<td>7,251</td>
<td>7,803</td>
<td>7,573</td>
<td>6,343</td>
<td>7,51721</td>
<td>-300,6</td>
</tr>
<tr>
<td>Asia</td>
<td>-4,2115</td>
<td>***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Europe</td>
<td>0,1866</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>USA</td>
<td>NA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HR</td>
<td>2,41E-04</td>
<td></td>
<td>7,05E-04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equity</td>
<td>3,37E-04</td>
<td></td>
<td>1,04E-04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Liabilities</td>
<td>5,14E-06</td>
<td></td>
<td>-8,90E-05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provisions</td>
<td>8,68E-04</td>
<td></td>
<td>5,71E-03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,05054</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,29045</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2,41233</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2,54012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1,6354</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0,73769</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Event Type7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NA</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0,15346</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insurance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-1,5261</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invesment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-0,706</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Residuals</th>
<th>F</th>
<th>R²</th>
<th>R² adj</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1066</td>
<td>38,3%</td>
<td>38,2%</td>
</tr>
<tr>
<td></td>
<td>306,1</td>
<td>8,2%</td>
<td>8,2%</td>
</tr>
<tr>
<td></td>
<td>177,3</td>
<td>4,9%</td>
<td>4,9%</td>
</tr>
<tr>
<td></td>
<td>22,9</td>
<td>0,7%</td>
<td>0,6%</td>
</tr>
<tr>
<td></td>
<td>91,69</td>
<td>2,6%</td>
<td>2,6%</td>
</tr>
<tr>
<td></td>
<td>388,3</td>
<td>31,1%</td>
<td>31,1%</td>
</tr>
<tr>
<td></td>
<td>44,75</td>
<td>7,3%</td>
<td>7,1%</td>
</tr>
<tr>
<td></td>
<td>11,02</td>
<td>0,3%</td>
<td>0,3%</td>
</tr>
<tr>
<td></td>
<td>3,321</td>
<td>0,0%</td>
<td>0,1%</td>
</tr>
</tbody>
</table>

**Robustness tests for the variables**
BIBLIOGRAPHY

Regulatory texts


BASEL COMMITTEE ON BANKING SUPERVISION (2011) Supervisory Guidelines for the Advanced Measurement Approaches, June 2011


Publications

AÏVAZIAN Sergeï (1978) Etude statistique des dépendances, Mir, Moscou


COPE Eric and LABBI Abderrahim (2008) Operational Loss Scaling by Exposure Indicators: Evidence from the ORX Database


DAVIS Ellen (2005) Operational Risk: Practical Approaches to Implementation, Riskbooks


DOWD Kevin, BLAKE David (2006) After VaR: The Theory, Estimations, and Insurance Applications of Quantile-Based Risk Measures, University of Nottingham


HIGHAM Nicholas J. (2002) Computing the Nearest Correlation matrix – A Problem from Finance, The University of Manchester


HOUSEHOLDER A. (1964) The theory of matrices in numerical analysis, USA


MAKRIDAKIS S., HIBON M. (1995) Evaluating accuracy (or error) measures, INSEAD, Fontainebleau

MATIAS Catherine (2013) Introduction à la statistique non paramétrique, CNRS, Laboratoire Statistique & Génome, Evry

MILLIMAN (2012) A brief overview of current approach to operational risk under Solvency II


RAKOTOMALALA Ricco (2012) Analyse de corrélation, Etude des dépendances – Variables quantitatives, Université Lumière Lyon 2, France

REBBONATO R., JÄCKEL P. (1999) The most general methodology to create a valid correlation matrix for risk management and option pricing purposes


Actuarial theses


RENAUDIN Alexis (2012) Modèle de capital économique pour le risque opérationnel bancaire: estimation, diversification, ISFA

Websites

http://news.bbc.co.uk
http://w3.mathinfoolmd.univ-tlse2.fr/
http://www.bis.org/
http://www.duclert.org/
http://www.ffsa.fr/
http://www.investopedia.com/
http://www.lesechos.fr
http://www.mathworks.fr/
http://www.milliman.com/
http://www.optimindwinter.com/
http://www.risk.net
LIST OF TABLES

Table 1: Modular view of operational risks split into 7 Event Lines and 8 Business Lines .................. 30
Table 2: Estimation of different correlation coefficients in function of the sample size ...................... 62
Table 3: Results of different independence test on the same sample ............................................. 68
Table 4: Descriptive statistics of a set of 1,000 simulations for the SCR (computed with 500,000 runs) computed for a low frequency/high impact scenario .......................................................... 97
Table 5: Descriptive statistics of a set of 1000 simulations for the SCR (computed with 500,000 runs) computed for a high frequency/low impact scenario .......................................................... 98
Table 6: Descriptive statistics of a set of 1000 simulations for the standalone SCR (computed with 100,000 runs) ........................................................................................................ 100
Table 7: Descriptive statistics of a set of 1000 simulations for the aggregated SCR (computed with 100,000 runs) ........................................................................................................ 101
Table 8: Linear correlations between the explanatory variables ..................................................... 116
Table 9: Pearson’s correlation coefficient between the losses and the explanatory variables ........ 116
Table 10: Spearman’s correlation coefficient between the losses and the explanatory variables .... 116
Table 12: Robustness tests by introducing step by step all of the variables ................................. 119
Table 13: Ratio between the median of scaled and actual losses, per entity in scope and per event type ..................................................................................................................... 121
Table 14: Distribution of the loss database per entity and per event type ........................................ 122
LIST OF FIGURES

Figure 1: Repartition of the main risks in the SCR ................................................................. 18
Figure 2: Illustrative boxplot of internal loss data for the 7 event risk categories .................. 24
Figure 3: Illustration of a Bayesian Network ........................................................................... 31
Figure 4: SCR computation based on the aggregate loss distribution curve ......................... 33
Figure 5: Annual loss as a combination of frequency and severity ......................................... 39
Figure 6: Discrepancies are inherent to the model .................................................................. 50
Figure 7: Distributions of the discrepancies for the Serious and the Extreme impacts ........... 51
Figure 8: Shape of the function to be minimized with the 4 different methods ..................... 54
Figure 9: Moran’s relation between Spearman’s $p_S$ and Pearson’s $p$ (for bivariate normal distributions) .............................................................................................................. 61
Figure 10: Greiner’s relation between Kendall’s $\tau$ and Pearson’s $p$ (for bivariate normal distributions) ......................................................................................................................... 62
Figure 11: Chi-squared test applied to the same sample with different intervals for $Y$. The test performed on the left accepts the null hypothesis of independence whereas the one performed on the left rejects the hypothesis of independence with a confidence level of 1% ..................................................................... 64
Figure 12: Traditional view of assessing risks ........................................................................ 69
Figure 13: SBA view of operational risks ................................................................................ 69
Figure 14: Illustration of the aggregation process with a normal copula approach where the risks to be aggregated are modelled with fixed frequency $= 1$ and lognormal severity ........................................... 86
Figure 15: Simulation of independent Poisson-Lognormal variables ...................................... 88
Figure 16: Simulation of a Gaussian copula (left) and a Student copula (right), $\rho = 0.5$; $\nu = 1$ ...... 88
Figure 17: Resulting Poisson-Lognormal variables correlated with a Gaussian (left) and a Student (right) copula ...................................................................................................................... 88
Figure 18: Computation of the SCR (left) and median (right) with different values for $\lambda$ ranging between 0.65 and 0.75 ........................................................................................................ 93
Figure 19: Computation of the SCR (left) and median (right) with different values for $\lambda$ ranging between 0.69 and 0.70 ............................................................................................................ 94
Figure 20: Computation of the SCR (left) and median (right) with different values for $\lambda$ ranging between 0.69 and 0.70 ............................................................................................................ 94
Figure 21: Monte Carlo convergence for a low frequency/high impact scenario is reached after 200,000 runs ......................................................................................................................... 97
Figure 22: Boxplot of a set of 1,000 simulations for the SCR (computed with 500,000 runs) computed for a low frequency/high impact scenario .............................................................................. 98
Figure 23: Monte Carlo convergence for a high frequency/low impact scenario is reached after 100,000 runs ........................................................................................................................................ 98
Figure 24: Boxplot of a set of 1,000 simulations for the SCR (computed with 500,000 runs) computed for a high frequency/low impact scenario .............................................................................. 99
Figure 25: Illustration of the two levels of diversification for the aggregation ........................ 100
Figure 26: Boxplot of a set of 1000 simulations for the aggregated SCR (computed with 100,000 runs) ...................................................................................................................................... 101
Figure 27: Increase in the 99.5%-percentile and expected loss in terms of lambda .................. 104
Figure 28: $\mu$ (in ordinate) against $\lambda$ (in abscissa) .................................................................... 105

Confidential - 135
Figure 29: $\sigma$ (in ordinate) against $\lambda$ (in abscissa) ................................................................. 105
Figure 30: Evidence of the existence of a threshold for the frequency parameter under which the 99.5%-percentile is null ................................................................. 106
Figure 31: Evidence of the existence of a threshold for the frequency parameter under which the 99.5%-percentile is null (zoom around 0.5%) ................................................................. 107
Figure 32: Variation of the standalone SCR in terms of the severity parameters .................... 108
Figure 33: An illustration of the Mean Excess Plot ................................................................. 111