Avril 2011 N°14

SCOR Papers

Modélisation du risque opérationnel dans le secteur de l'assurance

Par **Julie Gamonet**
Centre d'études actuarielles
Lauréate du prix du jeune actuaire 2010

Résumé

Une des grandes nouveautés de Solvabilité II est l'obligation, pour les compagnies d'assurance, de mobiliser une partie de leurs fonds propres en couverture de leur exposition aux risques opérationnels. Pour calculer ce montant de capital, le régulateur propose deux approches : une approche standard et une approche avancée. L'approche standard est une approche simplifiée calculée comme un pourcentage des primes ou des provisions. L'approche avancée est un modèle interne où le risque correspond réellement à la situation de l'entreprise. Le CEIOPS incite les compagnies à opter pour le modèle interne en rendant l'approche standard beaucoup plus consommatrice en fonds propres comme nous l'avons constaté dans le QIS5.
Cette note propose une approche qui distingue les risques de fréquence des risques de gravité. Les risques de fréquence sont définis comme étant des risques de pertes de montants peu élevés mais fréquentes : ils sont modélisés par la méthode Loss Distribution Approach. Les risques de gravité sont des risques de pertes de montants très importants mais très rares : ils sont modélisés par les réseaux Bayésiens.
Remarque : la modélisation séparée des risques de fréquence et de gravité rend la modélisation des risques de fréquence beaucoup moins sensibles aux choix techniques (lois ajustées, agrégation).

Ce SCOR Paper est une synthèse du travail pour lequel l'auteur a obtenur le prix du Jeune actuaire décerné par SCOR en partenariat avec l'Institut des actuaires.

SCOR Global Risk Center

1. Modélisation des risques de fréquence : méthode Loss Distribution Approach

La méthode Loss Distribution Approach consiste à ajuster des lois statistiques à des données de pertes, plus précisément de modéliser d'une part, la fréquence des incidents opérationnels et d'autre part, leur sévérité, puis de les combiner pour obtenir la distribution de pertes totales. Cette approche est souvent utilisée pour modéliser la charge totale des sinistres dans le cadre de la tarification ou du provisionnement.

La condition essentielle pour appliquer cette méthode sera la disponibilité de données de pertes historiques afin de calibrer le modèle.

Etape 1 : choix et calibrage des lois de fréquence et de sévérité

La modélisation de la distribution des pertes se fera pour chaque risque k :

$$S_k = \sum_{j=1}^{N_k} X_k^{(j)}$$

où N_k est la variable aléatoire représentant le nombre de pertes pour le risque k,

$X_k^{(j)}$ est la variable aléatoire représentant le coût de la perte j pour le risque k,

S_k est la somme des pertes du risque k.

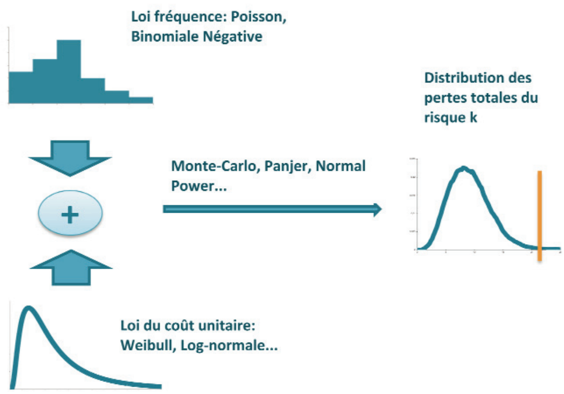
Les principales lois utilisées pour la fréquence sont usuellement Poisson, Binomiale et Binomiale Négative.

Il y a un quasi consensus de place pour utiliser la méthode du Maximum de vraisemblance afin d'estimer les paramètres des lois.

Le choix du modèle est ensuite validé par des tests statistiques.

Etape 2 : Construction de la distribution de pertes totales

Il est possible d'obtenir une bonne approximation de la fonction de répartition de la distribution de pertes S_k avec la méthode de simulation de Monte-Carlo.

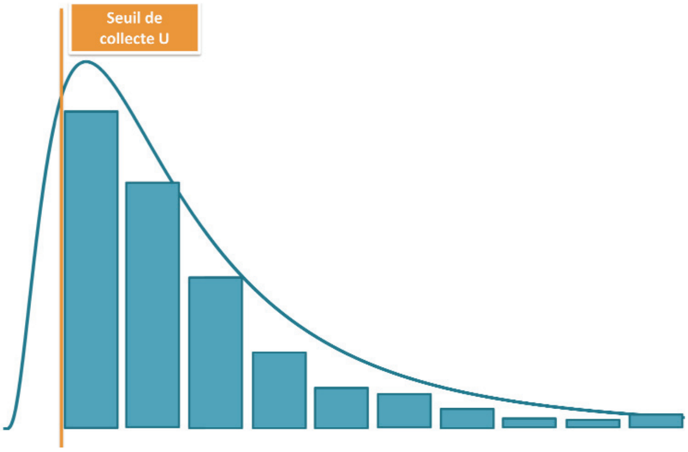


Problème pratique : le seuil de collecte

Généralement, les opérationnels ne déclareront pas toutes les pertes : il faudra alors fixer un seuil de déclaration. Les observations dont nous disposerons seront alors les pertes dont le montant dépasse ce seuil. Il faudra ainsi prendre en compte cette troncature dans nos ajustements et modifier l'estimation des paramètres et des tests d'adéquation en conséquence. En effet, si nous ne tenons pas compte du fait que les données sont tronquées, nous allons sous-estimer la fréquence et ajuster la loi des coûts individuels de façon erronée.

Loi du coût individuel :

Les données en dessous du seuil ne sont pas remontées mais doivent être prises en compte dans l'estimation des paramètres : il faudra estimer les paramètres de la loi en tenant compte du manque de données à gauche de l'histogramme tel que nous le montre le graphique ci-dessous (ce graphique représente le coût des pertes individuelles et l'adéquation de la loi théorique réajustée pour tenir compte des données tronquées) :

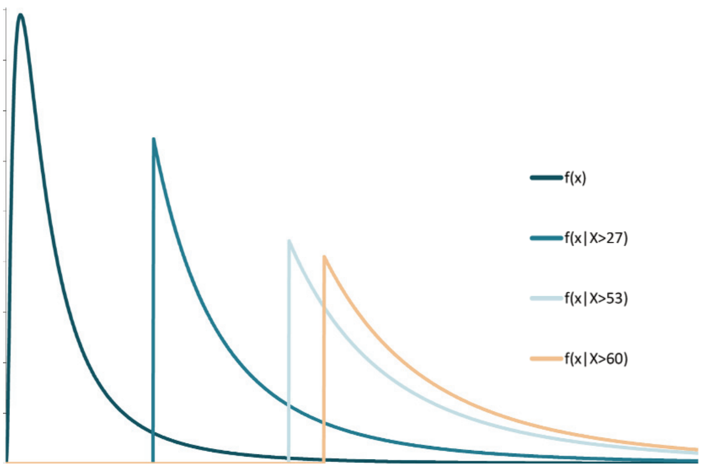


Soit U le seuil de collecte des pertes.
Soit X la variable aléatoire représentant les coûts des pertes non tronquées (c'est-à-dire les pertes dont le montant peut être inférieur à U) et F sa fonction de répartition.
Seule la loi de $X \mid X \geq U$ produit des observations. Cette loi conditionnelle est la loi de X tronquée à gauche de U.
Nous pouvons expliciter la fonction de répartition de la loi tronquée sachant que les observations sont au-delà du seuil U :

$$P(X < x \mid X \geq U) = \frac{P(U \leq X < x)}{P(X \geq U)} = \frac{F(x) - F(U)}{1 - F(U)} \quad \textit{pour tout } x > U$$

$$P(X \leq x \mid X \geq U) = 0 \quad \textit{pour tout } x \leq U$$

Le graphique ci-dessous montre la déformation d'une loi Log-normale $\ln(2\,;1)$ en fonction du niveau de seuil de collectes.



Remarque : certaines distributions sont stables par une troncature à gauche comme les lois Exponentielle et Pareto.

La méthode du maximum de vraisemblance nous permet de calibrer nos lois mais ne nous donne pas de solution analytique.
Soit $(x_1, x_2, \ldots, x_n)$ un échantillon de pertes dont le montant excède le seuil U, la vraisemblance conditionnelle au seuil U peut s'écrire :

$$\prod_{i=1}^{n} \frac{f(x_i)}{P(X_i \geq U)} = \prod_{i=1}^{n} \frac{f(x_i)}{1 - F(U)}$$

La log-vraisemblance peut alors s'écrire :

$$\sum_{i=1}^{n} \ln\left(\frac{f(x_i)}{1 - F(U)}\right) = \sum_{i=1}^{n} \ln(f(x_i)) - n \ln(1 - F(U))$$

Les paramètres sont estimés de manière classique en maximisant la log-vraisemblance.

Loi de la fréquence :

Lorsqu'il existe un seuil de collecte, la fréquence des pertes observées est sous-estimée et elle doit être réajustée pour tenir compte des pertes non déclarées. Pour chaque période d'observation i, le nombre de sinistres noté n_i est augmenté du nombre de sinistres estimés en-dessous du seuil noté m_i :

$$n_i^r = n_i + m_i$$

Le rapport entre le nombre de pertes en dessous du seuil et le nombre de pertes observées est le même que le rapport entre la partie en dessous de la courbe de la densité du coût individuel située à gauche du seuil, et celle située à droite du seuil :

$$\frac{m_i}{n_i} = \frac{\hat{F}(U)}{1 - \hat{F}(U)}$$

où $\hat{F}$ est la fonction de répartition tronquée de X, dont les paramètres ont été estimés avec la méthode précédente.

D'où :

$$n_i^r = n_i + m_i = n_i + \frac{n_i \times \hat{F}(U)}{1 - \hat{F}(U)} = \frac{n_i}{1 - \hat{F}(U)}$$

2. Modélisation des risques de gravité : méthode Bayésienne

L'approche Bayésienne consiste à réaliser une analyse qualitative des risques à dire d'experts et à la transformer en une analyse quantitative.

Un réseau Bayésien est un graphe causal probabiliste représentant la structure de la connaissance d'un certain domaine. Il est constitué de variables aléatoires discrètes reliées par des arcs orientés, ces variables étant appelées des nœuds. Une distribution est rattachée à chaque nœud. Les arcs sont des liens qui représentent une dépendance causale.

Nous avons étudié en particulier la méthode « XSG » (exposition, survenance, gravité) qui consiste à définir et modéliser les trois grandeurs caractéristiques du risque à savoir l'exposition, la survenance et la gravité. Ces trois grandeurs sont influencées par des variables appelées indicateurs de risque (ou KRI = Key Risk Indicator).

Nous allons présenter les méthodes pour évaluer les différents éléments du réseau Bayésien.

a) Evaluation de l'exposition

L'exposition est l'ensemble des éléments de l'entreprise qui sont exposés au risque. Elle doit être définie de sorte que le risque ne puisse survenir qu'une seule fois au plus dans l'année.

b) Evaluation de la survenance

L'objet exposé étant choisi de sorte qu'il ne puisse être frappé qu'au plus une seule fois, la survenance sera par construction une loi Binomiale B(n,p) où n est le nombre d'objets exposés et p la probabilité que nous aurons à estimer.

c) Evaluation de la gravité

Il faut se placer dans la situation où la survenance de la perte est avérée et identifier les variables quantifiables (KRI) intervenant dans le calcul de la gravité.

La structure du réseau Bayésien sera définie par les experts à travers des scénarios. Les paramètres du réseau Bayésien pourront être déterminés de façon empirique ou à dire d'experts.

Une fois le réseau Bayésien construit, il reste à définir l'algorithme de calcul.

Soient $(X_1, X_2, \ldots, X_n)$ les objets exposés au risque opérationnel étudié.

Soit $P_i = P(\text{Exposition} = X_i)$ la probabilité que l'exposition soit les objets X_i.

Soit $PS_i = P(\text{Survenance} = \text{«oui»} \mid \text{Exposition} = X_i)$ la probabilité que le risque survienne sachant que l'exposition est X_i.

Ces deux probabilités sont connues (elles ont été estimées comme vu précédemment).

Soit $PG_i = P(\text{Gravité} \mid \text{Survenance} = \text{«oui» et Exposition} = X_i)$ la distribution de la gravité sachant que le risque s'est produit sur les objets X_i.

L'algorithme consiste à réaliser successivement les étapes suivantes :

1) positionner l'exposition à X_i, la survenance à Oui dans le réseau bayésien et lire la distribution de la gravité $PG_i = P(\text{Gravité} \mid \text{Survenance} = \text{«oui» et Exposition} = X_i)$.

2) échantillonner le nombre de pertes F_i selon la loi binomiale $B(nb(X_i)\ ;\ PS_i)$.

3) pour chacun des incidents de 1 à F_i, échantillonner la gravité suivant la distribution PG_i.

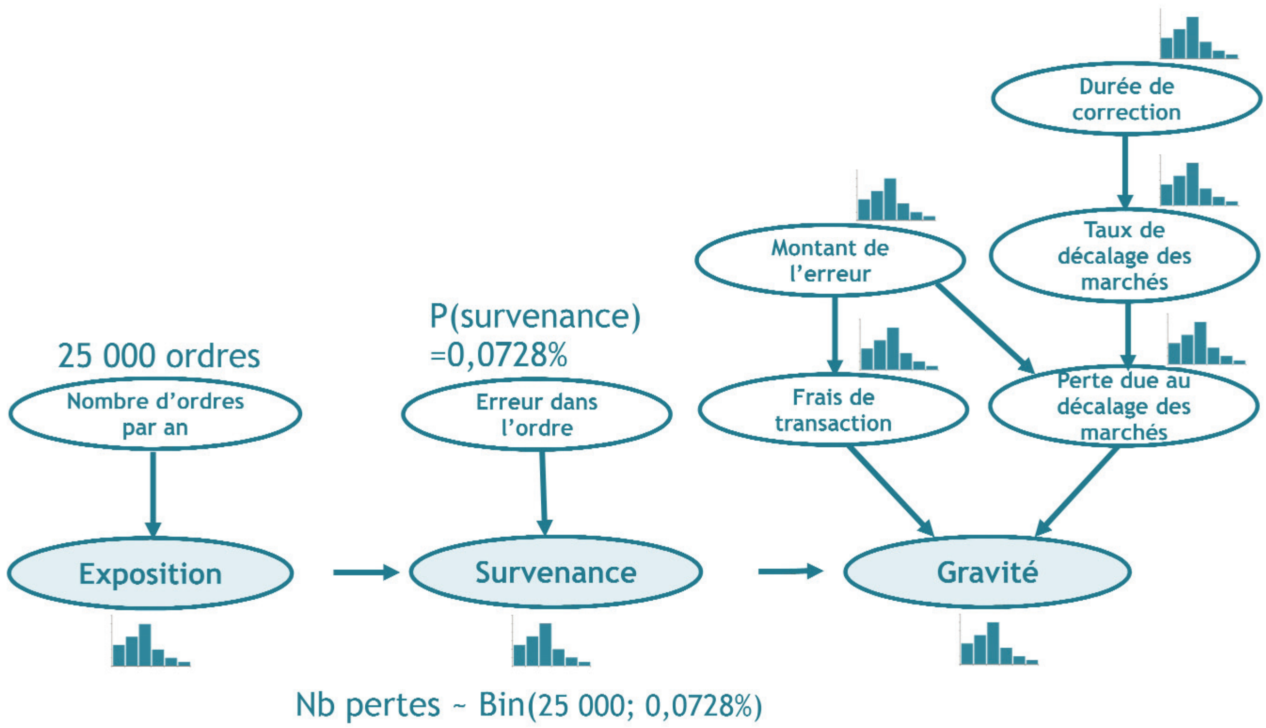
4) sommer les F_i gravités.

Répéter ces 4 étapes un grand nombre de fois en conservant les sommes des gravités à chaque fois. Nous obtenons ainsi une distribution de pertes totales.

3. Exemple d'application de la méthode Bayésienne

Nous appliquons à présent le modèle Bayésien au risque d'erreur dans le passage d'ordres sur le marché financier. Ce risque est traité de façon très simple car le but est de montrer comment s'applique l'approche Bayésienne.

La première étape dans l'approche Bayésienne consiste à créer le réseau à l'aide d'un graphe et à définir les distributions correspondant à chaque facteur. Ces paramètres pourront être revus dans une phase de « back-testing ».
Une analyse du risque nous a permis de construire le graphe suivant :



Rappelons que l'approche Bayésienne proposée pour le risque opérationnel consiste à définir trois objets : l'exposition, la survenance et la gravité.

L'exposition : l'exposition doit correspondre aux objets exposés de l'entreprise qui ne peuvent être touchés par le risque qu'une seule fois dans la période. Nous avons choisi les ordres. En effet un ordre ne peut être erroné qu'une seule fois. Nous n'anticipons pas d'augmentation dans le nombre d'ordres pour l'année à venir et le nombre d'ordres observés est en moyenne de 25 000 par an.

La survenance : la survenance est définie comme l'erreur sur un ordre. En moyenne, nous avons un nombre de pertes annuelles de 18,2 soit une probabilité d'erreur sur un ordre de 18,2/25 000 = 0,0728 %.

La gravité : la gravité est définie comme la somme de la perte due au décalage des marchés et des frais de transaction. Nous faisons les hypothèses suivantes sur les facteurs intervenant dans le calcul de la gravité :

• **le montant de l'erreur** (exprimé en millions d'euros) :

Montant de l'erreur (en millions d'euros)	
5	66%
15	18%
50	16%

• **la durée de correction de l'erreur :** c'est la période allant de la date de survenance de l'erreur à sa date de correction par le passage d'un nouvel ordre :

Durée de correction (en jours)	
0,125	66%
1	33%
90	1%

• **les frais de transaction :** nous supposons que les frais de transaction sont égaux à 0,25 % du montant de la transaction. Le montant d'erreur est passé une première fois puis une seconde fois lors de la correction. Les frais seront ici calculés en multipliant le montant de l'erreur par 0,5 %.

• **le taux de décalage des marchés :** il s'agit de la variation des taux observée pendant la durée de correction de l'erreur.

• **La perte due au décalage des marchés** : elle est calculée en multipliant le montant de l'erreur par le taux de décalage.

Taux de décalage des marchés	Durée de correction (en jours)		
	0,125	1	90
2%	66%	62%	60%
5%	34%	37,99%	28%
30%	0%	0,01%	12%

Lançons-nous dans l'algorithme décrit dans la partie théorique :

Ici, les objets exposés sont du même type, ce sont des ordres de même nature : $X_i = X$, ce qui implique que $P_i = P(\text{Exposition} = X_i) = 1$.

Nous en déduisons la probabilité que le risque survienne sachant que l'exposition est X_i :

$PS_i = P(\text{Survenance} = «\text{oui}» \mid \text{Exposition} = X_i) = P(\text{Survenance} = «\text{oui}») = 0{,}0728\ \%$

L'algorithme consiste à réaliser successivement les étapes suivantes :

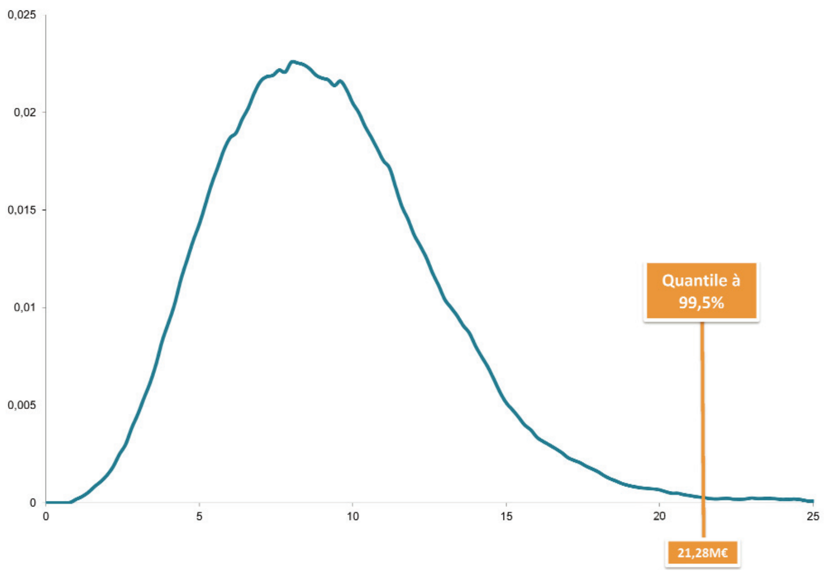
1) Calculer la distribution de la gravité $PG_i = P(\text{Gravité} \mid \text{Survenance} = «\text{oui}» \text{ et Exposition} = X_i)$.

2) Échantillonner le nombre de pertes F_i selon la loi Binomiale $B(nb(X_i);PS_i) = B(25\,000;0{,}0728\ \%)$. En effet, les ordres sont indépendants et chaque ordre suit une loi de Bernoulli.

3) Pour chacun des incidents de 1 à F_i, échantillonner la gravité suivant la distribution PG_i.

4) Sommer F_i les gravités.

Nous répétons ces 4 étapes 10 000 fois en conservant les sommes des gravités à chaque fois. Nous obtenons ainsi une distribution de pertes totales et nous pouvons en déduire la VaR à 99,5 %.



Nous avons réalisé des tests de sensibilité aux différents paramètres du réseau Bayésien : à chaque jeu de tests, nous faisons varier les paramètres d'un facteur en fixant les paramètres de tous les autres facteurs.

Le modèle Bayésien présente de nombreux avantages :
- il permet de prendre en compte à la fois des facteurs quantitatifs mais aussi des facteurs qualitatifs, ce que ne font pas la plupart des modèles ;
- il permet de visualiser les liaisons de causalité entre les variables : l'agrégation des risques est réalisée par la construction même des réseaux, ce qui évite l'estimation de corrélations;
- il permet de détecter des facteurs de réduction du risque grâce à l'inférence ;
- le Risk Management peut l'utiliser pour mettre en place des plans d'action et constater l'efficacité de ces derniers.

L'inconvénient majeur des réseaux Bayésiens est qu'ils sont longs à mettre en place car ils nécessitent une analyse détaillée de chaque risque.

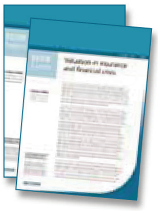
4. Bibliographie

Condamin, L. ; Louisot, J.-P. ; Naïm, P. (février 2007), Risk Quantification, Wiley, 286p.

Partrat Christian ; Besson Jean-Luc (décembre 2004), Assurance non-vie, Modélisation, Simulation, Economica, Paris, 820 pages.

Naïm P. ; Wuillemin P.-H. ; Leray P., Pourret O. ; Becker A. (novembre 2007), Réseaux Bayésiens, Eyrolles, collection Algorithmes, Paris, 424 pages.

Précédents SCOR Papers



SCOR Paper N°1 - Septembre 2008
Using Capital Allocation to Steer the Portfolio towards Profitability

SCOR Paper N°2 - Août 2008
La bancassurance : généralisation ou déclin du modèle ?

SCOR Paper N°3 - Décembre 2008
Valuation in insurance and financial crisis

SCOR Paper N°4 - Mars 2009
Modern Companies and Extreme Risks

SCOR Paper N°5 - Juillet 2009
Securitization, Insurance and Reinsurance

SCOR Paper N°6 - Janvier 2010
De la nécessité d'adapter la réglementation sur la solvabilité en période de crise et d'accepter les risques inhérents à la situation

SCOR Paper N°7 - Janvier 2010
The Influence of Risk Measures and Tail Dependencies on Capital Allocation

SCOR Paper N°8 - Mars 2010
Principle-based Solvency: A Comparison between Solvency II and the Swiss Solvency Test

SCOR Paper N°9 - Décembre 2010
Le vieillissement : un phénomène mondial

SCOR Paper N°10 - Décembre 2010
PrObEx: a New Method for the Calibration of Copula Parameters from Prior Information, Observations and Expert Opinions

SCOR Paper N°11 - Décembre 2010
Le risque de développement

SCOR Paper N°12 - Décembre 2010
Étude de l'impact de l'inflation et de la croissance du PIB sur l'évolution des primes en assurance IARD et en assurance Vie

SCOR Paper N°13 - Mars 2011
Preparing for Solvency II: Points of debate in the Standard Formula

Chaque année, SCOR récompense à travers le Prix de l'actuariat les meilleurs travaux de recherche en sciences actuarielles. Ce prix est destiné à promouvoir et à encourager la recherche en sciences actuarielles, et à contribuer à l'amélioration des connaissances et de la gestion des risques. Les jurys sont composés d'universitaires et de professionnels de l'assurance, de la réassurance et de la finance.

Prix de l'actuariat

SCOR
Global Risk Center

Les SCOR Papers, édités par SCOR, sont l'une des composantes du SCOR Global Risk Center.
Le SCOR Global Risk Center rassemble et analyse les ressources les plus pertinentes sur le risque. Il comporte à la fois des documents internes et des ressources externes sélectionnées par SCOR.
Pour l'accès au SCOR Global Risk Center : www.scor.com ou www.scorglobalriskcenter.com