# On the Rating of Dependent Risks

BAHRAM MIRZAI
Swiss Re New Markets, 8022 Zurich, Switzerland
Bahram_Mirzai@swissre.com

**Abstract**

Integrated covers, such as multi-lines, often assume independency among the covered risks. In practice, however, some of these risks may exhibit mutual dependency. The degree of dependency can, in turn, affect the expected loss burden of the integrated cover. In this paper, we introduce dependency through the frequency of covered lines and provide a rating approach that utilizes this to calculate the expected loss burden of a multi-line cover.

## 1 Introduction

Multi-line treaties have become an indispensable instrument in the risk underwriting of combined covers. The basic characteristic of a multi-line treaty is the aggregation of several lines of business within a single coverage with a common retention and limit. The traditional approach to the pricing of such covers assumes that the underlying lines of business are mutually independent. In practice, however, one may encounter risks that exhibit mutual dependency. In such an instance, the assumption of independence may serve as a first approximation to the true risk premium. We investigate here changes in premium that can arise from the interdependency of different lines of business.

This paper is organized as follows. In Sec. 2 we introduce the dependency of compound random variables through their frequency and calculate the implied correlation coefficient. Sec. 3 deals with the rating of dependent risks. Simulation results are provided in Sec. 4. We conclude with Sec. 5.

## 2 Dependent Risks

A common way to measure dependency of two random variables (RVs) $X$ and $Y$ is the notion of correlation coefficient $\rho_{xy}$, defined by

$$\rho_{xy} = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}}, \tag{1}$$

where $\mathrm{Cov}(X,Y)=\mathrm{E}[XY]-\mathrm{E}[X]\mathrm{E}[Y]$. The correlation coefficient is, in particular, a measure of linear dependency in the sense that for $Y=aX+b$ with $a,b\in\mathbb{R}$ it follows:

$$\rho_{xy}=\mathrm{sgn}(a)\,.$$

Furthermore, independent of the functional relationship of $X$ and $Y$, $\rho_{xy}$ satisfies the inequality:

$$|\rho_{xy}|\leq 1\,.$$

The RVs $X$ and $Y$ are said to be positively or negatively correlated if $\rho_{xy}>0$ or $\rho_{xy}<0$, respectively. Similarly, uncorrelated and fully correlated RVs are defined by equations $\rho_{xy}=0$ and $|\rho|=1$, respectively. Although the correlation coefficient may not always be an appropriate measure to fully capture the dependency relationship between two RVs [1], in the following, our analysis will be based on this measure.

To proceed, let $X$ be the RV of loss amount corresponding to an insured risk. Furthermore, let $N_x$ be the corresponding RV of annual claim count, henceforth, frequency. The aggregate or compound process $S_x$ is then defined by [2]:

$$S_x=\sum_{n=1}^{N_x}X_n\,.$$

In the following, we seek to calculate the correlation coefficient $\rho_{S_xS_y}$ of the compound processes $S_x$ and $S_y$, corresponding to the RVs $X$, $N_x$ and $Y$, $N_y$, respectively. Thereby, we make the assumption that the RVs $N_x$ and $N_y$ are Poisson distributed and that the only dependency between $S_x$ and $S_y$ arises from these RVs. Hence, the pairs $(X,Y)$, $(X,N_x)$ and $(Y,N_y)$ are assumed to be mutually independent RVs. From the dependency assumptions described, it follows:

$$\mathrm{Cov}(S_x,S_y)=\mathrm{E}[X]\mathrm{E}[Y]\mathrm{Cov}(N_x,N_y)\,. \tag{2}$$

Furthermore, by the Poisson assumption of $N_x$, we obtain [2]:

$$\mathrm{Var}(S_x)=\mathrm{E}[N_x]\mathrm{Var}(X)+\mathrm{E}[X]^2\mathrm{Var}(N_x)=\mathrm{E}[N_x]\mathrm{E}[X^2]\,. \tag{3}$$

A similar equation can be derived for $\mathrm{Var}(S_y)$. From (2-3), in conjunction with (1), it follows:

$$\rho_{S_xS_y}=\frac{\mathrm{E}[X]\mathrm{E}[Y]}{\sqrt{\mathrm{E}[X^2]\mathrm{E}[Y^2]}}\rho_{N_xN_y}\,. \tag{4}$$

## 3 Rating of Dependent Risks

In what follows, rating will refer to the pure risk premium calculation, i.e., to the expected loss burden and, hence, will exclude any additional loadings. The rating of insurance risks can be based on a frequency and severity analysis [2]. For given loss amount and frequency RVs $X$ and $N_x$, respectively, the expected aggregate loss is:

$$E[S_x] = E[\sum_{n=1}^{N_x} X_n] = E[N_x]E[X],$$

where the right equality will only hold if $X$ and $N_x$ are independent RVs. Analogously, in view of stop loss treaties, the expected loss in a layer with retention $R$ and limit $L$ can be written as[1]:

$$E[\mathcal{L}_{R,L}(S_x)] = E[\mathcal{L}_{R,L}(\sum_{n=1}^{N} X_n)],$$

where the layer operator $\mathcal{L}_{R,L}(\cdot)$ is defined by

$$\mathcal{L}_{R,L}(X) = \min(\max(X-R,0),L).$$

Generalization of this to the rating of multi-line covers can be done easily. For the sake of simplicity, we assume only two lines of business and denote the respective claim amount and frequency RVs by $X$, $Y$ and $N_x$, $N_y$. In the case of independent loss amounts $X$ and $Y$, and frequencies $N_x$ and $N_y$, the expected aggregate loss in the layer $[R, R+L]$ can be expressed as:

$$E[\mathcal{L}_{R,L}(S_{ind})] = E[\mathcal{L}_{R,L}(\sum_{n=0}^{N_x} X_n + \sum_{n=0}^{N_y} Y_n)]. \tag{5}$$

However, (5) needs to be modified, if applied to dependent risks. We introduce here dependency only through the frequency RVs. Although, this assumption may not always describe the actual nature of the dependency completely, due to its analytic tractability, we will restrict our analysis to it and will point out some of its shortcomings as we proceed. To be specific, we will do this in conjunction with an example.

Consider a combined cover of the two risks of earthquake and a tropical cyclone. In some regions, these risks may at times occur simultaneously and at other

[1] In this paper, the limit $L$ refers to the maximum covered amount in excess of $R$.

times independent of each other. To account for this, in such regions we assume a frequency split of the following form:

$$N_x = \hat{N}_x + N_{xy} \qquad N_y = \hat{N}_y + N_{xy}\,, \tag{6}$$

with mutually independent pairs ($\hat{N}_x$, $N_{xy}$), ($\hat{N}_y$, $N_{xy}$) and ($\hat{N}_x$, $\hat{N}_y$) of Poisson RVs. The designation "hat" shall refer to the intrinsic frequency, i.e., to that part of the frequency which is independent of presence or absence of any other risk. Furthermore, by $N_{xy}$, we denote the RV of claims that are common to both risks. For example, if earthquake and tropical cyclone happen with an average frequency of $\mathrm{E}[N_x] = 0.5$ and $\mathrm{E}[N_y] = 2$, respectively, and on an average every other four years the two risks occur together, then $\mathrm{E}[N_{xy}] = 0.25$, $\mathrm{E}[\hat{N}_x] = 0.25$ and $\mathrm{E}[\hat{N}_y] = 1.75$. The expected aggregate loss $S_{dep}$ corresponding to such a dependency for a layer with retention $R$ and limit $L$ is then given by:

$$\mathrm{E}[S_{dep}] = \mathrm{E}[\mathcal{L}_{R,L}\big(\sum_{n=1}^{\hat{N}_x} X_n + \sum_{n=1}^{\hat{N}_y} Y_n + \sum_{n=1}^{N_{xy}} (X_n + Y_n)\big)]\,. \tag{7}$$

In particular, in view of (4), from (6) we obtain:

$$\rho_{N_x N_y} = \frac{\mathrm{E}(N_{xy})}{\sqrt{\mathrm{E}(N_x)\mathrm{E}(N_y)}}\,. \tag{8}$$

A shortcoming that can arise from this approach is the application of the same loss RVs for the claims arising from $N_{xy}$ as well as for those arising from $\hat{N}_x$ and $\hat{N}_y$. In practice, one may expect a different loss magnitude than simply $X+Y$ whenever these two events happen simultaneously. From a computational point of view, this shortcoming can be removed by applying in (7) an appropriate RV $Z_n$ instead of $X_n + Y_n$. However, the correlation coefficient as given in (4) will no longer be valid.

To elaborate on a frequency dependency as in (6), we seek to compare the expected loss and the variance of the aggregate RVs $S_{ind}$ and $S_{dep}$ for $R = 0$ and $L = \infty$. In the case of independent frequency RVs we obtain:

$$\mathrm{E}[S_{ind}] = \mathrm{E}[N_x]\mathrm{E}[X] + \mathrm{E}[N_y]\mathrm{E}[Y]\,, \tag{9}$$

and

$$\mathrm{Var}[S_{ind}] = \mathrm{E}[N_x]\mathrm{E}[X^2] + \mathrm{E}[N_y]\mathrm{E}[Y^2]\,. \tag{10}$$

Similarly, for frequency dependent compound risks from (7) it follows:

$$\mathrm{E}[S_{dep}] = \mathrm{E}[\hat{N}_x]\mathrm{E}[X] + \mathrm{E}[\hat{N}_y]\mathrm{E}[Y] + \mathrm{E}[N_{xy}]\mathrm{E}[X+Y]\,. \tag{11}$$

Furthermore, we obtain

$$\begin{aligned}\mathrm{Var}[S_{dep}] =& \big(\mathrm{E}[\hat{N}_x]+\mathrm{E}[N_{xy}]\big)\,\mathrm{E}[X^2]+\big(\mathrm{E}[\hat{N}_y]+\mathrm{E}[N_{xy}]\big)\,\mathrm{E}[Y^2]+\\ & 2\,\mathrm{E}[N_{xy}]\,\mathrm{E}[X]\,\mathrm{E}[Y]\,. \end{aligned} \tag{12}$$

As seen from (9) and (11), clearly for independent RVs $X$ and $Y$ the introduction of the frequency dependency does not change the expected value of $S_{dep}$ and $S_{ind}$. However, in the case of frequency dependency, from (10) and (12) it follows that there is an increase in the variance given by:

$$\Delta\,\mathrm{Var} = \mathrm{Var}[S_{dep}] - \mathrm{Var}[S_{ind}] = 2\,\mathrm{E}[N_d]\,\mathrm{E}[X]\,\mathrm{E}[Y]\,.$$

The increase in variance in conjunction with the unchanged expected values, in particular, implies a heavier tailed distribution for $S_{dep}$ compared to $S_{ind}$. It also implies a lighter distribution in the lower layers for $S_{dep}$. This, in turn, implies that the expected value of two dependent compound risks in a higher layer will be larger than the respective value for the independent risks. A reverse situation, however, holds for lower layers.

A frequency dependency can also be introduced to remove coverage redundancies. Consider, e.g., the company ABC purchasing a stand-alone Directors and Officers (D&O) liability coverage. D&O policies usually provide a rather broad coverage including also employee related claims that otherwise would have been covered under Employment Practices Liability (EPL). As a result, in general a D&O severity and frequency distribution function will account for claims related to EPL. Assume now that the company ABC seeks to extend its coverage by adding EPL to it. Since the EPL severity and frequency distribution function already contains claims filed by employees against their directors and officers, there will be a double counting of these claims for the integrated cover. This double counting can be removed by introducing an appropriate frequency dependency. Let $X$ and $Y$ be the loss amount RVs of EPL and D&O, respectively. We introduce a frequency split as in (6), where $\hat{N}_x$ now corresponds to EPL claims that are not filed against directors and $\hat{N}_y$ to those D&O claims that can not be attributed to the employees. Furthermore, $N_{xy}$ describes the common claims, e.g., when an employee sues a director for wrongful termination. The expected aggregate loss in the layer $[R,\, R+L]$ can now be written as:

$$S = \mathrm{E}[\mathcal{L}_{R,L}\big(\sum_{n=1}^{\hat{N}_x} X_n + \sum_{n=1}^{\hat{N}_y} Y_n + \sum_{n=1}^{N_{xy}} X_n\big)]\,. \tag{13}$$

Note that the common claims are now taken into account only once through the EPL loss RV.

In practice, $N_{xy}$ may be specified by an equation of the form $\mathrm{E}[N_{xy}] = \alpha \mathrm{E}[N_x]$, where $\alpha$ is the percentage of the EPL claims in common with the D&O liability. In such a case, (8) can be expressed as:

$$\rho_{N_x N_y} = \frac{\alpha}{\sqrt{\frac{\mathrm{E}[N_y]}{\mathrm{E}[N_x]}}} .$$

Clearly, as $\alpha$ tends to zero the two RVs $N_x$ and $N_y$ become independent.

## 4 Simulation Results

Simulations are performed to compare the aggregate distribution functions of the dependent and independent compound risks. Dependency is introduced through the frequency, as discussed in the previous sections. The results below, in each case, assume 100000 independent random runs of simulation. Moreover, the expectation value is approximated by the sample mean [3]:

$$\mathrm{E}[X] = \frac{1}{N} \sum_{n=1}^{N} X_n .$$

For the distribution functions depicted below only the percentile values at $1, 2, \cdots, 99$ are taken. In depicted figures, solid lines correspond to the simulations assuming independent RVs, whereas dashed lines correspond to those assuming dependent RVs.

As a first simulation, we considered the three distributions given in Tab. 1, where Pareto($\alpha$, $\lambda$) denotes the Pareto distribution [4]

$$F_X(x) = 1 - \left(\frac{\lambda}{\lambda + x}\right)^{\alpha} .$$

Furthermore, Logn($\mu$, $\sigma$) denotes a lognormal distribution with parameters $\mu$ and $\sigma$, and Poi($\mu$) a Poisson distribution with mean $\mu$. The dependency was introduced by

| | Severity | Frequency |
|---|---|---|
| $X$ | Logn(8.8, 1.2) | Poi(6) |
| $Y$ | Logn(10.6, 0.8) | Poi(4) |
| $Z$ | Pareto(1.7, 10000) | Poi(4) |

Table 1: Severity and Frequency Distributions

$N_{xy} = 3$, $N_{xz} = 1$ and $N_{yz} = 1$, and, hence, $\hat{N}_x = 2$, $\hat{N}_y = 0$ and $\hat{N}_z = 2$. Simulations were performed in a layer with $R = 0$ and $L = 1M$. The resulting distribution

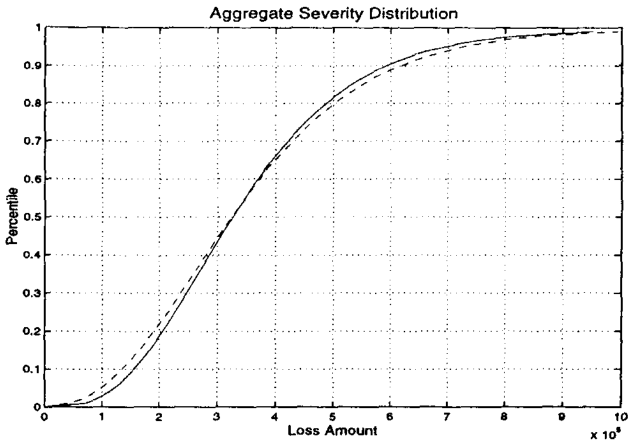


Figure 1: *Aggregate Distribution*

functions are depicted in Fig. 1. The depicted curves show the expected behavior, namely, an increase of variance in the case of dependent risks.

As a second simulation, we consider an example related to the removal of loss double counting as described in the previous section. The two risks are described by the two lognormal processes given in Tab. 2. As for dependency, we assumed $N_{xy} =$ Poi(0.5) and, hence, $\hat{N}_x =$ Poi(8.5) and $\hat{N}_x =$ Poi(1). Simulations were performed in a layer with $R = 0$ and $L = 100M$. After the removal of the redundancy the expected loss was reduced by an amount of 23%. The corresponding distribution functions of the independent case and dependent one and are depicted in Fig. 2.

| | Severity | Frequency |
|---|---|---|
| $X$ | Logn(10.866, 1.367) | Poi(9) |
| $Y$ | Logn(13.82, 2.174) | Poi(1.5) |

Table 2: Severity and Frequency Distributions

# 5 Conclusions

Introduction of dependencies among RVs allows one to better evaluate the underlying risk magnitude. In this paper, we introduced dependency of aggregate risks through their frequency RVs. This assumption, besides its analytic tractability, accounts for the fact that dependent compound risks can to a good approximation be described by a frequency dependency. The correlation coefficient was used as a

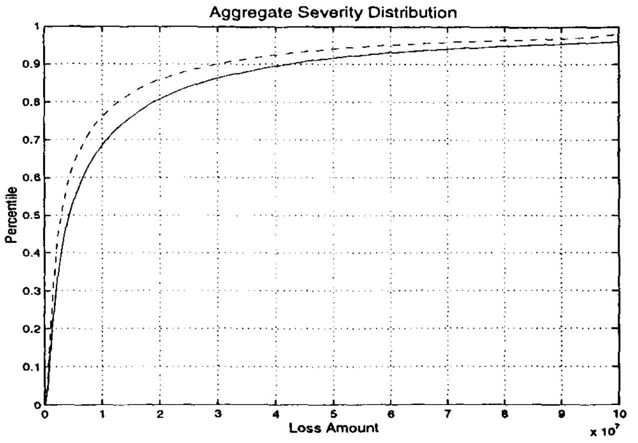


Figure 2: *Aggregate Distribution*

measure to quantify the degree of dependency of two compound risks. A frequency dependency as described in this paper, in particular, implies a heavier tailed distribution function for the dependent case.

## References

[1] P. Embrechts, A. Mcneil, and D. Steinmann, "Correlation and Dependency in Risk Management: Properties and Pitfalls," Preprint, Swiss Federal Institute of Technology, Zurich, 1998.

[2] C. D. Daykin, T. Pentikäinen, and M. Pesonen, *Practical Risk Theory for Actuaries*, Chapman & Hall, 1994.

[3] A. Law and D. Kelten, *Simulation Modeling & Analysis*, McGraw Hill, 1991.

[4] Robert V. Hogg and Stuart A. Klugman, *Loss Distributions*, John Wiley & Sons, 1984.