MYLENE BEDARD

DEPENDENCE STRUCTURES IN RISK THEORY

Research project
presented
to Professor Thierry Duchesne

Statistics Department
UNIVERSITY OF TORONTO
TORONTO

APRIL 2002

Introduction

Risk theory is the branch of actuarial science that evaluates the global risk of a portfolio subscribed by an insurance company, for example, in home, auto, or group insurance. The individual and collective risk models are two approaches to evaluate the global risk of a portfolio. Traditionally, in the study of these two models, one makes the hypothesis that the random variables are independent. This assumption allows simplifying the calculations. Unfortunately, in many situations, this hypothesis is not representative of the reality and some situations may arise where it is impossible to deny the existence of dependence between the risks. We can think, for example, of the occurrence of a natural disaster which generates numbers of claims from the policyholders living in the aœected area. Also, consider a group life insurance or a group health insurance contract issued to a company for a section of its employees working in a mine, on a steel plant, in a paper mill, etc. In these cases, a single event (e.g. explosion, breakdown) inÆuences the risks of the entire portfolio. The impact of the hypothesis of independence is nevertheless very important because ignoring the dependence between the risks may bring to an underestimation of the global risk of the portfolio. For this reason, it is important to consider some models introducing dependence between the risks.

The ørst part of this report will present the independent models. The individual and collective risk models will be discussed. An extension introducing many classes in the collective risk model will also be presented. The dependence models will then be introduced. The common mixture model will ørst be treated, following by the components models, where we discuss essentially of the common shock model. The distortion method will then be presented, and we will close the section with the copulas, where we also present some measures of correlation useful for their analysis.

Independent Models

1. Individual Risk Model

The individual risk model is usually applied in life insurance, in private health insurance, in car insurance and in other lines of non-life insurance. We assume that a portfolio is constituted of a ønite number (n) of individual contracts. This model is characterized by the fact that only a global claim amount can be associated with each risk of the portfolio, and this global amount comes from many reclamations made by the same policyholder.

We deøne S as a random variable (rv) representing the aggregate loss amount of a portfolio during a øxed period

$$S = X_1 + ... + X_n,$$

where the random variables X_i $(i = 1, ..., n)$ are the claim amount of the i^{th} contract for the given period. F_{X_i} is the cumulative distribution function (cdf) of the *rv* X_i $(i = 1, ..., n)$. It is important to note that the X_i's are independent but not necessarily identically distributed.

A more general model is to split the X_i's in two components, the occurence of at least one claim and the amount of the claims for a given period. As shown in Bowers et al. (1997), the rv of the claim amount X_i $(i = 1, ..., n)$ is expressed as

$$X_i = I_i B_i \ ,$$

where I_i is the indicator for the event that at least one claim occurs and B_i is a rv representing the total claim amount incurred during the period. The rv's I_i and B_i $(i = 1, ..., n)$ are independent (for a øxed i). The random variable I_i is a Bernoulli rv with

$$P[I_i = 1] = q_i \text{ and } P[I_i = 0] = 1 - q_i = p_i,$$

representing the probability that at least one claim occurs or not. We are assuming that the I_i's are independent. The rv's B_i $(i = 1, ..., n)$ are mutually independent, but are not necessarily identically distributed. One designates F_{B_i} as the cdf of B_i $(i = 1, ..., n)$.

We are interested in the cumulative distribution function of S, which is very useful to determine the diœerent characteristics of a portfolio. It is possible to ønd F_S, the cdf of S, through the convolution of the cdf 's of $X_1, ..., X_n$

$$F_S(s) = F_{X_1} * F_{X_2} * ... * F_{X_n}(s),$$

where

$$F_{X_1 * X_2}(s) = \int_0^s F_{X_1}(s-y) f_{X_2}(y) \, dy,$$

and where f_{X_i} is the probability density function (pdf) of $X_i \ (i = 1, ..., n)$.

Convoluting the X_i's is often tedious, and in many cases, is it easier to use the moment generating function (mgf) of S to determine F_S.

The mgf is deøned as $M_X(t) = E\left[e^{tX}\right]$, and can be expressed with the probability generating function (pgf), which is deøned as $P_X(t) = E\left[t^X\right]$. Hence, it is easy to see that

$$M_X(t) = P_X\left(e^t\right),$$

and inversely

$$P_X(t) = M_X(\ln t).$$

Now, for the mgf of S, we have

$$\begin{aligned} M_S(t) &= E\left[e^{tS}\right] \\ &= E\left[e^{t(X_1 + ... + X_n)}\right] \\ &= E\left[e^{tX_1 + ... + tX_n}\right] \end{aligned}$$

and by independence between the X_i's, it follows that

$$\begin{aligned} M_S(t) &= E\left[e^{tX_1}\right] ... E\left[e^{tX_n}\right] \\ &= M_{X_1}(t) ... M_{X_n}(t). \end{aligned} \quad (1)$$

Since an insurance company groups its risks in homogeneous classes, it is realist to suppose that the X_i's are independent and identically distributed (iid). It then follows from (1) that

$$M_S(t) = (M_X(t))^n.$$

After having obtained an explicit form for $M_S(t)$, we can inverse it by using numerical methods to approximate the *cdf* of S, F_S.

Sometimes, the mgf of the X_i's does not exists. A good way to overcome this problem is to use the Fourier transform, or characteristic function, deøned as $\phi_X(t) = E\left[e^{itX}\right]$. Note that i is deøned as $\sqrt{-1}$.

We develop the characteristic function of S. By the same argument as (1), we have

$$\phi_S(t) = \phi_{X_1}(t) \ldots \phi_{X_n}(t).$$

It is then possible to compute the *cdf* of S using the concept of the Fast Fourier Transform (FFT).

We are interested in the moments of the aggregate loss amount S. It is important to note that the variance is a good indicator of the risk of a portfolio. Let $E[B_i] = \mu_i$ and $Var(B_i) = \sigma_i^2$; since the expectation is a linear operator, we have

$$\begin{aligned} E[S] &= E\left[\sum_{i=1}^{n} X_i\right] \\ &= \sum_{i=1}^{n} E[X_i] \ , \end{aligned}$$

where

$$\begin{aligned} E[X_i] &= E_{I_i}[E[I_i B_i | I_i]] \\ &= E[\mu_i I_i] \\ &= \mu_i E[I_i] \\ &= \mu_i q_i. \end{aligned} \qquad (2)$$

For the variance of S, we develop

$$\begin{aligned} Var(S) &= Var\left[\sum_{i=1}^{n} X_i\right] \\ &= \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n} \sum_{j=1, j\neq i}^{n} Cov(X_i, X_j). \end{aligned}$$

Since the rv X_i $(i = 1, ..., n)$ are independent, we have

$$Var(S) = \sum_{i=1}^{n} Var(X_i),$$

where

$$Var(X_i) = E_{I_i}\left[E\left[(I_i B_i)^2 | I_i\right]\right] - E[I_i B_i]^2.$$

From (2), we have

$$\begin{aligned} Var(X_i) &= q_i E\left[B_i^2\right] - (\mu_i q_i)^2 \\ &= q_i\left(\sigma_i^2 + \mu_i^2(1 - q_i)\right) \\ &= \mu_i^2 q_i (1 - q_i) + \sigma_i^2 q_i. \end{aligned} \qquad (3)$$

2. Collective Risk Model

We now present the collective risk model, which is diœerent from the individual risk model by the fact that the number of claims is now represented by a random variable. We ørst present the classical model, and we then present an m-class extension of this model. The characteristic of this model is that the portfolio is considered as an whole rather than in terms of the individual policies.

2.1. Classical Collective Model.

We assume a homogeneous portfolio with a deøned number of contracts. N is a discret rv representing the number of claims during a øxed period, for example a year or a month. The amount of the i^{th} claim is represented by the rv X_i $(i = 1, 2, ...)$. We suppose that the X_i's are independent and identically distributed (iid). Furthermore, we assume that the rv's X_i $(i = 1, 2, ...)$ and the rv N are independent. F_X is deøned as the cdf of the X_i's.

S, the random variable representing the aggregate claims for the whole portfolio for a øxed and deøned period, is deøned as

$$S = X_1 + ... + X_N = \sum_{i=1}^{N} X_i.$$

One designates the cdf by F_S, which is of the following form

$$\begin{aligned} F_S(s) &= P\left[\sum_{i=1}^{N} X_i \leq s\right] \\ &= \sum_{n=0}^{\infty} P[N = n] P\left[\sum_{i=1}^{n} X_i \leq s \,| N = n\right]. \end{aligned}$$

Because of the hypothesis of independence between the rv N and the rv's X_i $(i = 1, 2, ...)$, we have

$$\begin{aligned} F_S(s) &= \sum_{n=0}^{\infty} P[N = n] P\left[\sum_{i=1}^{n} X_i \leq s\right] \\ &= \sum_{n=0}^{\infty} P[N = n] F_X^{*n}(s), \end{aligned} \tag{4}$$

where F_X^{*n} represents the n-fold convolution of the cumulative distribution function of X.

Convoluting the X_i's is often tedious, as mentioned in the previous section, and it is often easier to use the moment generating function (mgf) or the probability generating function (pgf) of S to determine F_S

$$\begin{aligned} M_S(t) &= E\left[e^{tS}\right] \\ &= E\left[e^{t(X_1+\ldots+X_N)}\right] \\ &= E_N\left[E\left[e^{tX_1+\ldots+tX_N}\,|N\right]\right], \end{aligned}$$

and since the X_i's are iid and independent of N, we have

$$\begin{aligned} M_S(t) &= E_N\left[E\left[e^{NtX_1}\,|N\right]\right] \\ &= E_N\left[E\left[e^{tX_1}\right]^N\right] \\ &= E_N\left[\left(M_{X_1}(t)\right)^N\right] \\ &= P_N\left(M_{X_1}(t)\right). \end{aligned} \tag{5}$$

After having obtained an explicit form for $P_N\left(M_{X_1}(t)\right)$, we can inverse $M_S(t)$ by using numerical methods to approximate the cdf of S, F_S.

Sometimes, the mgf is undeøned for the cdf of the X_i's and then, we shall use the characteristic function since it always exists. Note that we can express it with the mgf and the pgf, as

$$\phi_X(t) = E\left[e^{itX}\right] = P_X\left(e^{it}\right) = M_X(it), \tag{6}$$

where $i = \sqrt{-1}$.

From (5) and (6), we have

$$\begin{aligned} \phi_S(t) &= E\left[e^{itS}\right] \\ &= M_S(it) \\ &= P_N\left(M_{X_1}(it)\right) \\ &= P_N\left(\phi_{X_1}(t)\right). \end{aligned} \tag{7}$$

It is then possible to compute the *cdf* of S using the concept of the Fast Fourier Transform (FFT). The preceding results are used in an example presented at the end of this section, along with Splus algorithms for the discretization of a continuous distribution and for the FFT.

As mentioned in the previous section, we shall deøne the expectation and the variance of S. We determine the expectation of S as

$$\begin{aligned} E[S] &= E[X_1 + \ldots + X_N] \\ &= E\left[\sum_{i=1}^{N} X_i\right], \end{aligned}$$

conditioning on N, we get

$$\begin{aligned} E[S] &= E_N\left[E\left[\sum_{i=1}^{N} X_i \,|N\right]\right] \\ &= E_N\left[\sum_{i=1}^{N} E[X_i\,|N]\right], \end{aligned}$$

and because the X_i's are iid and independent of N, it follows that

$$\begin{aligned} E[S] &= E_N[NE[X]] \\ &= E[N]\,E[X]. \end{aligned} \tag{8}$$

We then develop an expression for the variance of S

$$Var(S) = E\left[S^2\right] - E^2[S]. \tag{9}$$

For $E[S^2]$, we have

$$\begin{aligned} E\left[S^2\right] &= E\left[\left(\sum_{i=1}^{N} X_i\right)^2\right] \\ &= E_N\left[E\left[\sum_{i=1}^{N} X_i^2 + \sum_{i=1}^{N}\sum_{j=1, j\neq i}^{N} X_iX_j\,|N\right]\right], \end{aligned}$$

and again by the independence between N and the iid X_i's, we have

$$\begin{aligned} E\left[S^2\right] &= E_N\left[NE\left[X^2\right] + N(N-1)\,E^2[X]\right] \\ &= E[N]\,E\left[X^2\right] + E\left[N^2\right]E^2[X] - E[N]\,E^2[X]. \end{aligned} \tag{10}$$

Replacing (8) and (10) in (9), we obtain

$$\begin{aligned} Var[S] &= E[N]\,E\left[X^2\right] + E\left[N^2\right]E^2[X] \\ &\quad -E[N]\,E^2[X] - E^2[N]\,E^2[X] \\ &= E[N]\,Var[X] + Var[N]\,E^2[X]. \end{aligned} \tag{11}$$

Example 2.1. Suppose that N is Poisson(100) and that the X_i's are iid Lognormal$(2, 1)$. The pgf of a Poisson(λ) is

$$\begin{aligned} P_N(t) &= E\left[t^N\right] \\ &= \sum_{i=0}^{\infty} t^n \frac{\lambda^n e^{-\lambda}}{n!} \\ &= e^{-\lambda}\sum_{n=0}^{\infty}\frac{(t\lambda)^n}{n!} \end{aligned}$$

and by the following result from the Taylor expansion

$$e^x = \sum_{i=0}^{\infty} \frac{x^i}{i!},$$

we get

$$\begin{aligned} P_N(t) &= e^{-\lambda} e^{t\lambda} \\ &= e^{\lambda(t-1)}. \end{aligned} \qquad (12)$$

Now, we can obtain the characteristic function of S from (7)

$$\begin{aligned} \phi_S(t) &= P_N\left(\phi_{X_1}(t)\right) \\ &= e^{\lambda\left(\phi_{X_1}(t)-1\right)}, \end{aligned}$$

where $\phi_{X_1}(t)$ is the characteristic function of X_1 (same for all X_i's, since they are iid). In our case, we have

$$\phi_S(t) = e^{100\left(\phi_{X_1}(t)-1\right)}.$$

We can use this expression along with the FFT algorithm presented in the appendix to plot diœerent functions of S. The ørst graph (Figure 1) shows the cdf of S. The second graph (Figure 2) shows the stop-loss premium, which we designate as the expectation of the rv I_d. We deøne I_d as

$$I_d(S) = \begin{cases} 0, & S < d \\ S - d, & S \geq d \end{cases},$$

where d is the deductible. We also denote the stop-loss premium as $\pi(d) = E\left[(S-d)_+\right]$. Note that the stop-loss premium allows to quantify the risk related to a portfolio. Of course, as we increase the deductible d, the stop-loss premium decreases. Also, for $d = 0$, we have $\pi(d) = E[S]$ and from (8), $\pi(d) = 100\left(e^{2+1/2}\right) = 1218.25$.

2.2. Collective Model with m-class extension. We now consider an adaptation of the preceding collective risk model. This extension of the classical model will allow to include dependence between the diœerent classes of a portfolio. We suppose a portfolio including m classes of risks. For the i^{th} class, we deøne the random variables

- $N^{(i)} =$ number of claims in the i^{th} class of policyholders $(i = 1, ..., m)$;
- $X_j^{(i)} = j^{\text{th}}$ $(j = 1, 2, ...)$ claim in the i^{th} class of policyholders $(i = 1, ..., m)$;

Figure 1. Graph of F_S where N ∼Poisson(100) and the X_i's are iid Lognormal$(2, 1)$

- $S^{(i)}$ = total amount of claims in the i^{th} class of policyholders $(i = 1, ..., m)$ in a øxed period

$$S^{(i)} = \sum_{j=1}^{N^{(i)}} X_j^{(i)}; \tag{13}$$

- S = total amount of claims for the portfolio in the øxed period

$$S = S^{(1)} + ... + S^{(m)}. \tag{14}$$

In the i^{th} class $(i = 1, ..., m)$, we suppose that the random variables $X_j^{(i)}$ $(j = 1, 2, ...)$ are iid. We also suppose that the rv's $X_j^{(i)}$ $(i = 1, ..., m$ and $j = 1, 2, ...)$ and the rv's $N^{(i)}$ $(i = 1, ..., m)$ are independent. We designate by $F_{X^{(i)}}$ the common cdf of $X_j^{(i)}$, for each j. We also make the additional hypothesis that the random vectors $\left(X_1^{(i)}, X_2^{(i)}, ...\right)$ and $\left(X_1^{(i')}, X_2^{(i')}, ...\right)$ are independent for two diœerent classes of policyholders $(i \neq i')$.

Figure 2. Graph of $E\left[(S-d)_{+}\right]$, the stop-loss premium with limit d, versus d

It follows from (4) that the cdf of $S^{(i)}$, $F_{S^{(i)}}$, is of the following form

$$F_{S^{(i)}}(s)=\sum_{n=0}^{\infty} P\left(N^{(i)}=n\right) F_{X^{(i)}}^{*n}(s).$$

Then, the cdf of S is expressed as

$$F_{S}(s)=F_{S^{(1)}} * \ldots * F_{S^{(n)}}(s).$$

From the preceding development of the mgf of S for the classical collective risk model, we know that each class taken individually is expressed, from (5), as

$$\begin{aligned} M_{S^{(i)}}(t) &= E\left[e^{tS^{(i)}}\right] \\ &= P_{N^{(i)}}\left(M_{X^{(i)}}(t)\right). \end{aligned} \tag{15}$$

Hence, we the mgf of S for a model with m classes is

$$\begin{aligned} M_S(t) &= E\left[e^{tS}\right] \\ &= E\left[e^{t\left(S^{(1)}+...+S^{(m)}\right)}\right] \\ &= E\left[e^{tS^{(1)}+...+tS^{(m)}}\right], \end{aligned}$$

and by independence between the classes and from (15), we obtain

$$\begin{aligned} M_S(t) &= E\left[e^{tS^{(1)}}\right]...E\left[e^{tS^{(m)}}\right] \\ &= \prod_{i=1}^{m} P_{N^{(i)}}\left(M_{X^{(i)}}(t)\right). \end{aligned} \tag{16}$$

Now that we identiøed the mgf of S for a model with m classes, we can inverse it and ønd the cdf of S.

We now develop the expectation and the variance of S for this model. For the expectation of S, we have

$$\begin{aligned} E[S] &= E\left[S^{(1)}+...+S^{(m)}\right] \\ &= E\left[S^{(1)}\right]+...+E\left[S^{(m)}\right]. \end{aligned}$$

Using (8), we obtain

$$\begin{aligned} E[S] &= E\left[N^{(1)}\right]E\left[X^{(1)}\right]+...+E\left[N^{(m)}\right]E\left[X^{(m)}\right] \\ &= \sum_{i=1}^{m} E\left[N^{(i)}\right]E\left[X^{(i)}\right]. \end{aligned}$$

The variance of S is

$$\begin{aligned} Var[S] &= Var\left[S^{(1)}+...+S^{(m)}\right] \\ &= \sum_{i=1}^{n} Var(X_i) + \sum_{i=1}^{n}\sum_{j=1,j\neq i}^{n} Cov(X_i, X_j). \end{aligned} \tag{17}$$

By independence between the random variables $S^{(i)}$ $(i = 1, ..., m)$, (17) becomes

$$Var[S] = \sum_{i=1}^{m} Var\left[S^{(i)}\right],$$

and from (11) we can deduct

$$Var[S] = \sum_{i=1}^{m}\left(E\left[N^{(i)}\right]Var\left[X^{(i)}\right] + Var\left[N^{(i)}\right]E^2\left[X^{(i)}\right]\right).$$

Example 2.2. We are interested in ønding the distribution of S when $S^{(1)}, ..., S^{(m)}$ are distributed as independent Compound-Poisson (CP) distributions with parameters λ_i and $F_{X^{(i)}}(x)$ (cdf of the rv of the claim amount). We designate the distribution of the random variable $S^{(i)}$ $(i = 1, ..., m)$ by

$$S^{(i)} \sim CP\left(\lambda_i, F_{X^{(i)}}(x)\right).$$

We use the hypotheses of the collective model with a portfolio of m independent classes. From (16), we have for the mgf of S

$$M_S(t) = \prod_{i=1}^{m} P_{N^{(i)}}\left(M_{X^{(i)}}(t)\right). \tag{18}$$

Since the random variables $N^{(i)}$ $(i = 1, ..., m)$ are distributed as Poisson distributions with parameter λ_i, we have from (12)

$$\begin{aligned} M_S(t) &= \prod_{i=1}^{m} e^{\lambda_i\left(M_{X^{(i)}}(t)-1\right)} \\ &= e^{\lambda \sum_{i=1}^{m} \frac{\lambda_i}{\lambda}\left(M_{X^{(i)}}(t)-1\right)} \\ &= e^{\lambda\left(\sum_{i=1}^{m} \frac{\lambda_i}{\lambda} M_{X^{(i)}}(t)-1\right)}, \end{aligned} \tag{19}$$

which represents the mgf of a Compound-Poisson with parameters $\lambda = \sum_{i=1}^{m} \lambda_i$ and $F_Y(y) = \sum_{i=1}^{m} \frac{\lambda_i}{\lambda} F_{X^{(i)}}(y)$ (i.e. a convex combination of the cdf's of the rv's $X^{(i)}$ $(i = 1, ..., m)$).

Thus,

$$S \sim \text{CP}\left(\lambda, F_Y(y) = \sum_{i=1}^{m} \frac{\lambda_i}{\lambda} F_{X^{(i)}}(y)\right),$$

where $S = S^{(1)} + ... + S^{(m)}$ and $\lambda = \sum_{i=1}^{m} \lambda_i$.

Dependent Models

We now introduce some models allowing the inclusion of relations of dependence. The models presented in this section are essentially based on the paper of Wang (1998).

3. Common Mixture Models

In many situations, individual risks are correlated since they are inÆuenced by changes in their common environment, either in their economic, climatic or any other environment. For instance, consider many cultivators of cereals in Saskatchewan. Their crops all depend, among other things, on the climatic conditions. Hence, they cannot be considered as independent risks since they are related by the same factor, the weather. Also, in property insurance, portfolios subscribing risks in the same geographic area are correlated, since they are contingent upon the occurrence of a natural disaster. We can think of a particularly rigorous winter, a few years ago in Quebec, which brought an incredible amount of snow and resulted in damages of roofs and swimming pools. In liability insurance, the eœect of inÆation may set new trends that aœect the settlement of all liability claims for one line of business.

Individual risks $\{X_1, X_2, ...\}$ subject to the same factor may be modeled by using a secondary mixing distribution. This factor's uncertainty is described by θ, a realization of the random variable Θ, allowing to represent the distribution of each individual risk. This means that there exists a distribution function of the random variable Θ representing the common factor, F_Θ, upon which each individual risk depend. Thus, it is easy to see that these risks are not independent, since they all depend on the same factor Θ. However, when we know the realization of this random variable, θ, then we can consider the individual risks as independent, since they are not aœected anymore by the distribution function F_Θ.

The aggregate loss of the portfolio can then be determined in two steps. First the parameter $\Theta = \theta$ is drawn from F_Θ, the cdf of Θ. This means that we know the conditional distribution functions of $X_i\,|\Theta$,

$F_{X_i|\Theta}$, $(i = 1, 2, ...)$, which are independent. Next, we obtain the claim frequency (or severity) of each individual risk, which is a realization from the conditional distribution function of $X_i|\Theta$, $F_{X_i|\Theta}$, $(i = 1, 2, ...)$.

In the previous example, we considered that the individual risks depended only on one extern mechanism. However, in many situations, the risks are subject to more than one factor and it is why we sometimes need to extend this method to more than one mixing distribution. In the case where the mixing distributions are independent, the situation is not more complicated than that with only one mixing distribution. The individual risks depend on the two independent distributions of Θ and Λ, F_Θ and F_Λ, respectively. Hence, the aggregate loss of the risk portfolio can be determined by ørst drawing the parameters $\Theta = \theta$ and $\Lambda = \lambda$ from F_Θ and F_Λ. Then, knowing the conditional distribution functions of $X_i|\Theta, \Lambda$, $F_{X_i|\Theta,\Lambda}$, $(i = 1, 2, ...)$, which are independent, we can obtain the claim frequency (or severity) as a realization of $F_{X_i|\Theta,\Lambda}$.

This method of including dependence may reveal itself very useful, as we just seen. However, it has also some disadvantages. For instance, consider the case where the individual risks depend on more than one external mechanism, as we just seen, but this time the external mechanisms are dependent. In this kind of situation, we would need either the joint cdf of all the factors, or another method to calculate the dependence between them, which bring us to our initial problem. Since the joint cdf's are often complicated, this kind of situation is in general more diŒcult to manage and in such cases, we may be better to use another method of including dependence. Also, note that in using some distributions as the Negative Binomial, some restrictions in the choice of the parameters arise and we cannot be as Æexible as we may wish.

Note that the mixing distribution can be either discrete or continuous. However, we must specify two important points about ønite mixtures. First, such models are often oversimpliøed, since the risks are generally more likely to generate numbers of risk levels (maybe a continuum of risk levels). Also, the number of parameters to estimate in a ønite mixture is very high. If we have r classes, then we will have $r-1$ mixing parameters in addition to the total number of parameters in the r component distributions. These reasons justify the fact that continuous mixtures are frequently preferred to ønite ones. We now present an example of a continuous mixture.

Example 3.1. Suppose the parameter q of a binomial follows a beta distribution with parameters α and β. The beta distribution has *pdf*

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

where $0 < q < 1$, $\alpha > 0$, $\beta > 0$.

Then, the mixed distribution has probabilities

$$\begin{aligned} p_k &= \int_0^1 \binom{m}{k} q^k (1-q)^{m-k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} q^{\alpha-1}(1-q)^{\beta-1} dq \\ &= \binom{m}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 q^{\alpha+k-1}(1-q)^{\beta+m-k-1} dq \\ &= \binom{m}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+k)\Gamma(\beta+m-k)}{\Gamma(\alpha+k+\beta+m-k)} \\ & \int_0^1 \frac{\Gamma(\alpha+k+\beta+m-k)}{\Gamma(\alpha+k)\Gamma(\beta+m-k)} q^{\alpha+k-1}(1-q)^{\beta+m-k-1} dq. \end{aligned}$$

Since we have the *pdf* of a beta with parameters $\alpha^* = \alpha + k$ and $\beta^* = \beta + m - k$ integrated over its whole range $(0, 1)$, then the integral equals 1, which gives

$$\begin{aligned} p_k &= \binom{m}{k} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+k)\Gamma(\beta+m-k)}{\Gamma(\alpha+\beta+m)} \\ &= \frac{\Gamma(m+1)\Gamma(\alpha+\beta)\Gamma(\alpha+k)\Gamma(\beta+m-k)}{\Gamma(k+1)\Gamma(m-k+1)\Gamma(\alpha)\Gamma(\beta)\Gamma(\alpha+\beta+m)}, \quad k = 0, 1, 2, ... \end{aligned}$$

This distribution is called Binomial-Beta, Negative Hypergeometric, or Polyo-Eggenberger.

4. Components Models

If we consider the aggregation of diœerent lines of business, we can realize that each of them may be diœerent from one region to another. Therefore, dividing the risks into components and model each one separately may be a more appropriate way to perform the calculations. For instance, the amount of risk may diœer if some lines of business are located in a high catastrophe risk region, while others are located in a safer one. In this case, the use of a common mixture model or of a common shock may be required in order to include the high risk factor in the ørst lines of business.

When using the components models, the distributions inønitely divisible are very useful. Many families of frequency or severity distributions have this property. A distribution is said to be inønitely divisible if it can be obtained by a sum of independent distributions in the same family. It is very useful because they allow dividing risks into independent components. Among the most famous inønitely divisible distributions, we can state the Poisson, the Negative Binomial and the Gamma

$$\begin{aligned}
\text{Poisson}(\lambda_1) \oplus \text{Poisson}(\lambda_2) &= \text{Poisson}(\lambda_1+\lambda_2) \\
\text{NB}(\alpha_1,\beta) \oplus \text{NB}(\alpha_2,\beta) &= \text{NB}(\alpha_1+\alpha_2,\beta) \\
\text{Gamma}(\alpha_1,\beta) \oplus \text{Gamma}(\alpha_2,\beta) &= \text{Gamma}(\alpha_1+\alpha_2,\beta),
\end{aligned}$$

where $\oplus$ is used to represent the sum of two independent random variables.

4.1. Common Shock Models. We shall begin by consider the case where the risks can be aœected by only one common shock. Let $X_j = X_{ja} \oplus X_{jb}$, $j = 1, 2, ...$, be a decomposition into two independent components. If we assume that $X_{1a} = X_{2a} = ... = X_0$, then we obtain $X_j = X_0 \oplus X_{jb}$, $j = 1, 2, ...$. This means that each risk is aœected by the occurrence of a same event (the common shock X_0), or by the occurrence of individual events (X_{jb}). Hence, the only relation of dependence between the diœerent risks comes from the common shock X_0. This can be seen by calculating the covariance between X_i and X_j

$$\begin{aligned}
\text{Cov}[X_i,X_j] &= \text{Cov}[X_0+X_{ib},X_0+X_{jb}] \\
&= \text{Cov}[X_0,X_0]+\text{Cov}[X_0,X_{jb}] \\
&\quad +\text{Cov}[X_{ib},X_0]+\text{Cov}[X_{ib},X_{jb}].
\end{aligned}$$

Since the diœerent components are independent, we have

$$Cov[X_i,X_j] = Var[X_0].$$

This common shock model may be easily extended to higher dimension. For instance, if we consider three variables, we get

$$X_j = X_0 \oplus X_{ij} \oplus X_{kj} \oplus X_{jb},$$

where X_0 represents the common shock among all three variables, X_{ij} represents the common shock between i and j, X_{kj} represents the common shock between k and j, and X_{jb} represents the individual risk. Also note that X_0, X_{ij}, X_{kj}, and X_{jb} are independent. Hence, the correlation will be included through

$$\begin{aligned}
\text{Cov}[X_i,X_j] &= \text{Cov}[X_0+X_{ij}+X_{ik}+X_{ib},X_0+X_{ij}+X_{kj}+X_{jb}] \\
&= \text{Cov}[X_0,X_0]+\text{Cov}[X_0,X_{ij}]+...+\text{Cov}[X_{ib},X_{jb}].
\end{aligned}$$

Since the diœerent components are independent, we have

$$\text{Cov}[X_i,X_j] = \text{Var}[X_0]+\text{Var}[X_{ij}],$$

which shows that i and j are correlated under the common shock among all three i, j, and k, and under the extra common shock between i and j.

Note that as we increase the dimension, i.e. as we increase the number of independent components for a risk, the calculation becomes tedious since we have to consider new sources of correlation, and this number grows rather fast. Thus, for situations where the variables are subject to numbers sources of dependence, some other models may be more appropriate or, at least, easier to compute.

Example 4.1. Consider a portfolio with two risks. Let

$$X_1 = X_0 \oplus X_{1b} \text{ and } X_2 = X_0 \oplus X_{2b},$$

where X_{1b} ~Gamma(α_1, λ) , X_{2b} ~Gamma(α_2, λ) , X_0 ~Gamma(α_0, λ) , and where X_{1b}, X_{2b}, X_0 are independent. We are interested in the distribution of $S = X_1 + X_2$. This results by ønding the *mgf* of S, $M_S(t)$

$$\begin{aligned} M_S(t) &= E\left[e^{tS}\right] \\ &= E\left[e^{t(X_1+X_2)}\right] \\ &= E\left[e^{t(X_0+X_{1b}+X_0+X_{2b})}\right]. \end{aligned}$$

By independence between X_0, X_{1b}, and X_{2b}, we have

$$\begin{aligned} M_S(t) &= E\left[e^{2tX_0}\right] E\left[e^{tX_{1b}}\right] E\left[e^{tX_{2b}}\right] \\ &= M_{X_0}(2t)\, M_{X_{1b}}(t)\, M_{X_{2b}}(t). \end{aligned}$$

We can then inverse $M_S(t)$ by numerical methods, as stated previously.

4.2. Peeling Method. We saw in the section treating the common mixtures that this method presents some restrictions in the choice of the parameters. The Peeling method presents a way to overcome this limitation, oœering more Æexibility in the choice of the parameters of our distributions.

As an example, if $N_j\,|\Theta$ ~Poisson$(\theta\lambda_j)$ and Θ ~Gamma$(\alpha, 1)$, then the marginal distribution of N_j is a NB(α, λ_j)(see Wang (1998), p.892). We can see that the same parameter α is required in the marginal Negative Binomial distributions. However, the components method allows to construct correlated multivariate Negative Binomial distributions with arbitrary parameters (α_j, λ_j) through two methods.

The ørst method shows that this can be done by separating each marginal N_j $(j = 1, ..., k)$ of a set of k marginal Negative Binomial distributions into the following decomposition

$$N_j = N_{1j} \oplus N_{2j},$$

where

$$N_{1j} \sim \mathrm{NB}(\alpha_0, \lambda_j), \text{ and } N_{2j} \sim \mathrm{NB}\left(\alpha_j - \alpha_0, \lambda_j\right),$$

and where NB represents a Negative Binomial distribution and α_0 is such that $\alpha_0 \leq \min\{\alpha_1, ..., \alpha_k\}$.

The second method assumes that the α_j's are in ascending order $\alpha_1 \leq ... \leq \alpha_k$. We can then write each marginal distribution in the following decomposition

$$\mathrm{NB}\left(\alpha_j, \lambda_j\right) = \mathrm{NB}\left(\alpha_1, \lambda_j\right) \oplus \mathrm{NB}\left(\alpha_2 - \alpha_1, \lambda_j\right) \oplus ... \oplus \mathrm{NB}\left(\alpha_j - \alpha_{j-1}, \lambda_j\right).$$

4.3. Mixed Correlation Models. This method may be useful in testing a set of possible scenarios. It consists in mixing joint probability generating functions having the same set of marginal probability generating functions, and this results obviously in a mixed joint probability generating function still having the same set of marginal probability generating functions. Hence, assembling each scenario with its probability of occurrence, it may be a good way to compute the overall probability of an event.

5. Distortion Method

This method consists to introduce a correlation structure in a joint probability generating function through a function, say g.

Let $X_1, ..., X_n$ be n random variables with probability generating functions $P_{X_1}(t_1), ..., P_{X_n}(t_n)$, respectively. Assuming that the X_j's are mutually independent, we can write

$$P_{X_1,...,X_n}(t_1, ..., t_n) = \prod_{j=1}^{n} P_{X_j}(t_j).$$

Also, let g be a strictly increasing function over $[0,1]$, with $g(1) = 1$ and whose inverse function is g^{-1}. We assume that $g \circ P_{X_1,...,X_n}$ consists in a joint *pgf* with marginal *pgf*'s $g \circ P_{X_j}$, $(j = 1, ..., n)$. If we assume that these marginal *pgf*'s are non-correlated, we then have

$$g \circ P_{X_1,...,X_n}(t_1, ..., t_n) = \prod_{j=1}^{n} g \circ P_{X_j}(t_j).$$

A correlation structure is now introduced to the original joint probability generating function:

$$P_{X_1,...,X_n}(t_1, ..., t_n) = g^{-1}\left\{\prod_{j=1}^{n} g \circ P_{X_j}(t_j)\right\}.$$

Note that for mathematical convenience, the function $h(x) = \ln g(x)$, which is a strictly increasing function over $[0,1]$ with $h(1) = 0$, is often used instead of $g(x)$. Hence, the joint *pgf* is expressed as

$$P_{X_1,...,X_n}(t_1,...,t_n) = h^{-1}\left\{\prod_{j=1}^{n} h \circ P_{X_j}(t_j)\right\}. \qquad (20)$$

However, we should note that since the only constraint on the joint probability function is that it sums to one, we might not obtain a proper multivariate distribution. It deønes a proper multivariate distribution only if the joint probability function $f_{X_1,...,X_n}$, is non-negative everywhere. Hence, $P_{X_1,...,X_n}$ deønes a proper joint probability generating function if, and only if, its partial derivatives at $t_1 = ... = t_n = 0$ are all non-negative.

Theorem 5.1. Suppose that Equation (20) deønes a joint probability generating function; we have

$$Cov[X_i, X_j] = \left\{\frac{h''(1)}{h'(1)} + 1\right\} E[X_i]\, E[X_j].$$

Proof. This can be shown be taking the second order partial derivative, $\partial^2/\partial t_i \partial t_j$ $(i \neq j)$, on both sides of the equation

$$h \circ P_{X_1,...,X_n}(t_1,...,t_n) = \prod_{j=1}^{n} h \circ P_{X_j}(t_j).$$

For a complete proof, see Wang (1998). ∎

Note that the distortion method also oœers the possibility to include dependence in a set of Negative Binomial distributions having diœerent parameters (i.e. no restriction on the α_i's).

Example 5.1. Consider the following common Poisson mixture model: $N_j | \Theta \sim$Poisson$(\theta\lambda_j)$, and Θ has a chi-squared distribution with $1/p$ degrees of freedom. According to Wang (1998), if we deøne $h(y) = M_\Theta^{-1}(y)$, then the joint pgf for this model satisøes (20).

Since the mgf of Θ is $M_\Theta(z) = (1-2z)^{-1/2p}$, deøne

$$h(y) = (1 - y^{-2p})/2.$$

Using (20), we get the following joint pgf:

$$\begin{aligned}
P_{N_1,...,N_n}(t_1,...,t_n) &= h^{-1}\left\{\sum_{j=1}^{n} h\circ P_{N_j}(t_j)\right\} \\
&= h^{-1}\left\{\sum_{j=1}^{n}\left(\frac{1-P_{N_j}(t_j)^{-2p}}{2}\right)\right\} \\
&= h^{-1}\left\{\frac{1}{2}\left(n-\sum_{j=1}^{n} P_{N_j}(t_j)^{-2p}\right)\right\} \\
&= \left\{1-2\left[\frac{1}{2}\left(n-\sum_{j=1}^{n} P_{N_j}(t_j)^{-2p}\right)\right]\right\}^{-1/2p} \\
&= \left(\sum_{j=1}^{n} P_{N_j}(t_j)^{-2p}-n+1\right)^{-1/2p},
\end{aligned}$$

where $p \neq 0$. By the preceding theorem on the covariance, we have

$$Cov\left[N_i, N_j\right] = -2pE\left[N_i\right]E\left[N_j\right],$$

since

$$h'(y) = py^{-(2p+1)} \text{ and } h''(y) = -p(2p+1)y^{-(2p+2)},$$

which implies that

$$\frac{h''(1)}{h'(1)} + 1 = -2p.$$

Finally,

$$\lim_{p\to 0} P_{N_1,...,N_n} = P_{N_1}(t_1)...P_{N_n}(t_n),$$

since the covariance goes to 0.

We will see in the next section, that the distortion method has an application in the construction of some copulas.

6. Copulas

The popularity of this concept is due, among other things, to the fact that it is often very diŒcult to work with multivariate distributions. They are diŒcult to invert, diŒcult to simulate, etc.. The copulas allow, in comparison with the preceding methods presented, a dependence structure more Æexible. Wang (1998) provides a good introduction to this concept.

Definition 6.1. A copula is deøned as the joint cumulative distribution function of k uniform random variables

$$C(u_1, ..., u_k) = P[U_1 \leq u_1, ..., U_k \leq u_k].$$

Let $Y_1, ..., Y_k$ be a sequence of random variables with cdf 's $F_{Y_1}, ..., F_{Y_k}$ respectively. We can then express the multivariate cdf $(F_{Y_1,...,Y_k})$ with a copula C

$$F_{Y_1,...,Y_k}(y_1, ..., y_k) = C(F_{Y_1}(y_1), ..., F_{Y_k}(y_k)).$$

We can also express the joint survivor function

$$S_{Y_1,...,Y_k}(y_1, ..., y_k) = C(S_{Y_1}(y_1), ..., S_{Y_k}(y_k)).$$

A copula C is a cdf of a random vector $(U_1, ..., U_k)$, where U_i is a uniform variable on the interval $(0, 1)$. This multivariate function $C : [0, 1]^k \to [0, 1]$ has, among others, the following properties:

- $C(0, ..., 0) = 0$;
- $C(1, ..., 1) = 1$;
- $C(1, ..., u_i, ..., 1) = u_i$, for $i = 1, ..., k$ and $u_i \in [0, 1]$;
- $C(u_1, ..., u_k)$ is increasing in each component u_i;
- For each $(a_1, ..., a_k), (b_1, ..., b_k) \in [0, 1]^k$ with $a_i \leq b_i$, we have

$$\sum_{i_1=1}^{2} \cdots \sum_{i_k=1}^{2} (-1)^{i_1+...+i_k} C\left(u_{i_1}^{(1)}, ..., u_{i_k}^{(k)}\right) \geq 0,$$

where $u_1^{(i)} = a_i$ and $u_2^{(i)} = b_i$ for $i = 1, ..., k$.

The last property state that there is no negative weight on $[a_1, b_1] \times ... \times [a_k, b_k]$. These properties are required for C to be a cumulative distribution function.

With the copulas, with can distinguish the choice of the marginal distributions (F_{Y_i}, S_{Y_i}) from the dependence structure. Also, if the marginal distributions are continuous, there exists a unique copula to represent the multivariate cdf, and every multivariate *cdf* can be expressed as a copula. Unfortunately, some problems are due to the fact that we cannot always identify this copula. However, it is important to notice that even if the copula for the joint *cdf* and that for the joint survivor function have the same set of Kendall's tau and of rank correlation coeŒcients, they are usually diœerent. The copulas have the nice property to be easily simulated. We will show an example of this application at the end of the section.

6.1. Measures of correlation. We now present some measures of correlation, concepts that will be very useful for a better understanding of the copulas.

There exists diœerent methods of measuring the correlation between two random variables. The most famous is probably the Pearson correlation coeŒcient, which is deøned as

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

and always lies in the range $[-1,1]$. A linear relationship between X and Y ($X = aY + b$, for some constants $a > 0$ and b) is traduced by $\rho(X,Y) = 1$, a linear relationship between X and $-Y$ has $\rho(X,Y) = -1$, while $\rho(X,Y) = 0$ when X and Y are independent.

The covariance coeŒcient is deøned as

$$\begin{aligned} \omega(X,Y) &= \frac{Cov[X,Y]}{E[X]E[Y]} \\ &= \rho(X,Y)\frac{\sigma[X]\sigma[Y]}{E[X]E[Y]} \\ &= \rho(X,Y)\,\text{CV}(X)\,\text{CV}(Y), \end{aligned}$$

where CV refers to the coeŒcient of variation. This time, the range of $\omega(X,Y)$ depends on the shape of the marginal distributions.

The Spearman's rank correlation coeŒcient is

$$\text{RankCorr}(X,Y) = 12E[(F_X(x) - 0.5)(F_Y(y) - 0.5)].$$

The Kendall's tau is the measure of correlation the most used with the concept of copulas. It is deøned as

$$\begin{aligned} \tau &= \tau(X,Y) \\ &= P[(X_2 - X_1)(Y_2 - Y_1) \geq 0] - P[(X_2 - X_1)(Y_2 - Y_1) < 0] \\ &= 4\int_0^1\int_0^1 F_{X,Y}(x,y)\,d^2F_{X,Y}(x,y) - 1, \end{aligned}$$

where (X_1, Y_1) and (X_2, Y_2) are two independent realizations of a joint distribution.

Both Kendall's tau and Spearman's rank correlation coeŒcient satisfy the following properties:

- $-1 \leq \tau \leq 1; -1 \leq \text{RankCorr} \leq 1$;
- if X and Y are comonotonic, then $\tau = 1$ and $\text{RankCorr} = 1$;
- if X and $-Y$ are comonotonic, then $\tau = -1$ and $\text{RankCorr} = -1$;
- if X and Y are independent, then $\tau = 0$ and $\text{RankCorr} = 0$;

- τ and RankCorr are invariant under monotone transforms, i.e.

$$\tau\left(f\left(X\right),g\left(Y\right)\right)=\tau\left(X,Y\right)$$

  and

$$\text{RankCorr}\left(f\left(X\right),g\left(Y\right)\right)=\text{RankCorr}\left(X,Y\right).$$

Note that a non-parametric estimate of Kendall's tau is given by

$$\widehat{\tau}\left(X,Y\right)=\frac{2}{k\left(k-1\right)}\sum_{i<j}sign\left[\left(X_i-X_j\right)\left(Y_i-Y_j\right)\right].$$

A Splus function allowing to estimate the Kendall's tau for a random sample of bivariate observations, $\left(X_i,Y_i\right),i=1,...,k$, is shown in the appendix.

We now present further important concepts concerning bivariate random variables, say $\left(X,Y\right)$. Let

$$F_{X,Y}\left(x,y\right)=P\left[X\leq x,Y\leq y\right]$$

be the joint cdf and

$$S_{X,Y}\left(x,y\right)=P\left[X>x,Y>y\right]$$

be the joint survivor function of $\left(X,Y\right)$. Note that

$$\begin{aligned}S_{X,Y}\left(x,y\right)&=1-F_X\left(x\right)-F_Y\left(y\right)+F_{X,Y}\left(x,y\right)\\&\neq 1-F_{X,Y}\left(x,y\right).\end{aligned}$$

Lemma 6.1. For any bivariate cumulative distribution function $F_{X,Y}$ with given marginal distributions F_X and F_Y, we have

$$\max\left[F_X\left(x\right)+F_Y\left(y\right)-1,0\right]\leq F_{X,Y}\left(x,y\right)\leq\min\left[F_X\left(x\right),F_Y\left(y\right)\right],$$

where $\max\left[F_X\left(x\right)+F_Y\left(y\right)-1,0\right]$ and $\min\left[F_X\left(x\right),F_Y\left(y\right)\right]$ are called the Frechet bounds.

Proof. We will prove each inequality separately. For the lower bound, we have

$$F_{X,Y}\left(x,y\right)=F_X\left(x\right)+F_Y\left(y\right)-S_{X,Y}\left(x,y\right).$$

Since we know that $S_{X,Y}\left(x,y\right)\leq 1$, we then have

$$F_{X,Y}\left(x,y\right)\geq F_X\left(x\right)+F_Y\left(y\right)-1.$$

Also, since $F_{X,Y}\left(x,y\right)\geq 0$, we can write

$$F_{X,Y}\left(x,y\right)\geq\max\left[F_X\left(x\right)+F_Y\left(y\right)-1,0\right].$$

Now, for the upper bound, we have

$$F_{X,Y}\left(x,y\right)=P\left[X\leq x,Y\leq y\right]=P\left[X\leq x\cap Y\leq y\right].$$

Since $P[A\cap B]\leq P[A]$ and $P[A\cap B]\leq P[B]$, we have

$$\begin{aligned} F_{X,Y}(x,y) &\leq \min[P[X\leq x], P[Y\leq y]] \\ &= \min[F_X(x), F_Y(y)], \end{aligned}$$

which complete the proof. ∎

The concept of comonotonicity is closely related with Frechet bounds.

Definition 6.2. Two random variables X and Y are comonotonic if there exists a random variable Z such that

$$X=u(Z),\ Y=v(Z),\ \text{with probability one},$$

where the functions u and v are non-decreasing.

We can see this concept as an extension of the perfect correlation. For example, we can think of the variable Z as the salary of an individual, the variable X as the amount of taxes and the variable Y as the percentage of salary deducted for the pension of this individual in a year.

$$X=\begin{cases} \alpha_1 Z, & 0\leq Z\leq d_1 \\ \alpha_2(Z-d_1), & d_1<Z\leq d_2 \\ \alpha_3(Z-d_2), & Z>d_2 \end{cases}$$

where $0\leq\alpha_1<\alpha_2<\alpha_3\leq 1$, $d_1<d_2$, and

$$Y=\beta Z,\ Z\geq 0,$$

where $0<\beta<1$. Since the proportion of taxes paid is a step function and the deduction for the pension is a constant percentage, then we cannot express X as a linear function of Y, or vice versa. However, X and Y are non-decreasing functions depending on the variable Z, and then are comonotonic. They depend on the same variable, and they do not hedge against each other.

The concept of comonotonicity has a property related to the Frechet bounds.

Proposition 6.2. If X and Y are comonotonic, then the upper Frechet bound is reached and if X and $-Y$ are comonotonic, then the lower Frechet bound is reached.

Proof. To prove the ørst case, let $X=u(Z)$ and $Y=v(Z)$ for some random variable Z and non-decreasing functions u and v. Deøne also the inverse functions of u and v as

$$u^{-1}(x)=\sup\{z: u(z)\leq x\} \qquad \text{and} \qquad v^{-1}(y)=\sup\{z: v(z)\leq y\}.$$

Hence, we have

$$\begin{aligned}F_{X,Y}(x,y) &= P[X \leq x, Y \leq y] \\ &= P[u(Z) \leq x, v(Z) \leq y] \\ &= P\left[Z \leq u^{-1}(x), Z \leq v^{-1}(y)\right] \\ &= P\left[Z \leq \min\left(u^{-1}(x), v^{-1}(y)\right)\right].\end{aligned}$$

This leads to

$$\begin{aligned}F_{X,Y}(x,y) &= \min\left(P\left[Z \leq u^{-1}(x)\right], P\left[Z \leq v^{-1}(y)\right]\right) \\ &= \min\left(P[u(Z) \leq x], P[v(Z) \leq y]\right) \\ &= \min\left(P[X \leq x], P[Y \leq y]\right) \\ &= \min\left(F_X(x), F_Y(y)\right),\end{aligned}$$

and the ørst case is proved.

To prove the case where X and $-Y$ are comonotonic, let $X = u(Z)$ and $-Y = -v(Z)$ for some random variable Z, and non-decreasing functions u and v. Hence, $-v$ is a non-increasing function. Deøne also the inverse functions of u and $-v$ as

$$u^{-1}(x) = \sup\{z : u(z) \leq x\} \qquad \text{and} \qquad -v^{-1}(y) = \inf\{z : -v(z) \geq y\}.$$

We then have

$$\begin{aligned}F_{X,-Y}(x,y) &= P[X \leq x, -Y \leq y] \\ &= P[u(Z) \leq x, -v(Z) \leq y] \\ &= P\left[Z \leq u^{-1}(x), Z \geq v^{-1}(-y)\right] \\ &= P\left[Z \leq u^{-1}(x)\right] - P\left[Z \leq v^{-1}(-y)\right],\end{aligned}$$

which leads to

$$\begin{aligned}F_{X,-Y}(x,y) &= P[u(Z) \leq x] - P[-v(Z) \geq y] \\ &= P[u(Z) \leq x] - (1 - P[-v(Z) \leq y]) \\ &= P[X \leq x] + P[-Y \leq y] - 1 \\ &= F_X(x) + F_{-Y}(y) - 1.\end{aligned}$$

Since we know that $F_{X,-Y}(x,y) \geq 0$, we must write

$$F_{X,-Y}(x,y) = \max\left(F_X(x) + F_{-Y}(y) - 1, 0\right),$$

which completes the proof. ∎

The concept of comonotonicity is also related to the Kendall's tau. When two random variables X and Y are comonotonic, Kendall's tau, $\tau(X, Y)$, equals one. Suppose that for a realization of the random variable Z, we get the bivariate sample (X_1, Y_1), and for a second realization of Z, we get (X_2, Y_2). Hence, we can say that $X_1 \leq X_2 \Longleftrightarrow Y_1 \leq Y_2$, and $X_2 \leq X_1 \Longleftrightarrow Y_2 \leq Y_1$, since $X = u(Z)$ and $Y = v(Z)$ bet on the same event Z, and since u and v are both non-decreasing functions. Thus, the expression $(X_2 - X_1)(Y_2 - Y_1)$ must be greater or equal to zero (≥ 0), since it represents a multiplication of two numbers either both positive or both negative. We can then write for the Kendall's tau τ

$$\begin{aligned}\tau &= \tau(X, Y) \\ &= P[(X_2 - X_1)(Y_2 - Y_1) \geq 0] - P[(X_2 - X_1)(Y_2 - Y_1) < 0] \\ &= 1,\end{aligned}$$

since $P[(X_2 - X_1)(Y_2 - Y_1) \geq 0] = 1$ and $P[(X_2 - X_1)(Y_2 - Y_1) < 0] = 0$ from the preceding explanation.

We can also show that when X and Y are independent, Kendall's tau equals 0:

$$\begin{aligned}\tau &= 4\int_0^1 \int_0^1 F_{X,Y}(x, y)\, d^2 F_{X,Y}(x, y) - 1 \\ &= 4\int_0^1 \int_0^1 F_X(x)\, F_Y(y)\, dF_X(x)\, dF_Y(y) - 1,\end{aligned}$$

by independence between X and Y. Since $f_X(x)$ and $f_Y(y)$ are the derivatives of $F_X(x)$ and $F_Y(y)$, respectively, we have

$$\tau = 4\int_0^1 F_X(x)\, f_X(x)\, dx \int_0^1 F_Y(y)\, f_Y(y)\, dy - 1,$$

and by letting $s = F_X(x)$, $ds = f_X(x)\, dx$ and $t = F_Y(y)$, $dt = f_Y(y)\, dy$, we ønd

$$\begin{aligned}\tau &= 4\int_0^1 s ds \int_0^1 t dt - 1 \\ &= 4\left(\left.\frac{s^2}{2}\right|_{s=0}^{1}\right)\left(\left.\frac{t^2}{2}\right|_{t=0}^{1}\right) - 1 \\ &= 4\left(\frac{1}{2}\right)\left(\frac{1}{2}\right) - 1 \\ &= 0.\end{aligned}$$

Now that we know the principal measures of correlation, we will present some particular copulas.

6.2. Cook-Johnson Copula. There exists a lot a copulas, which are often classiøed in families. The Archimedean family of copulas is one of them, and the copulas belonging to this family are constructed with the distortion method. The Cook-Johnson copula belongs to the Archimedean family.

$(U_1, ..., U_k)$ is a k-dimensional uniform distribution with support on the hypercube $(0,1)^k$ and having the joint *cdf*

$$F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) = \left\{ \sum_{j=1}^{k} u_j^{-1/\alpha} - k + 1 \right\}^{-\alpha}, \tag{21}$$

where $u_j \in (0,1)$, $j = 1, ..., k$ and $\alpha > 0$.

The Cook-Johnson copula can be simulated by the following algorithm:

1. Let $Y_1, ..., Y_k$ be k *iid* Exponential(1) random variables;
2. Let Z be a Gamma$(\alpha, 1)$ random variable independent of the Y_i's;
3. Compute the variables

$$U_j = \left(1 + \frac{Y_j}{Z}\right)^{-\alpha}, \ j = 1, ..., k.$$

Then, the U_j's have a joint cdf given by (21). As an exercise, we will show this result. We have

$$\begin{aligned} F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) &= P(U_1 \le u_1, ..., U_k \le u_k) \\ &= P\left(\left(1 + \frac{Y_1}{Z}\right)^{-\alpha} \le u_1, ..., \left(1 + \frac{Y_k}{Z}\right)^{-\alpha} \le u_k\right). \end{aligned}$$

By the law of total probabilities, we have

$$\begin{aligned} F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) &= \int_0^\infty P\left(\left(1 + \frac{Y_1}{Z}\right)^{-\alpha} \le u_1, ..., \right. \\ &\qquad \left. \left(1 + \frac{Y_k}{Z}\right)^{-\alpha} \le u_k \,|Z = z\right) f_Z(z)\, dz \\ &= \int_0^\infty P\left(Y_1 \ge \left(u_1^{-1/\alpha} - 1\right) z, ..., \right. \\ &\qquad \left. Y_k \ge \left(u_k^{-1/\alpha} - 1\right) z \,|Z = z\right) f_Z(z)\, dz. \end{aligned}$$

Since $Y_1, ..., Y_k$ are *iid* Exponential(1) and Z is Gamma$(\alpha, 1)$, it follows that

$$\begin{aligned} F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) &= \int_0^\infty e^{-\left(u_1^{-1/\alpha}-1\right)z}...e^{-\left(u_k^{-1/\alpha}-1\right)z}\frac{1}{\Gamma(\alpha)}z^{\alpha-1}e^{-z}dz \\ &= \int_0^\infty \frac{1}{\Gamma(\alpha)}z^{\alpha-1}e^{-z\left(\sum_{j=1}^k\left(u_j^{-1/\alpha}-1\right)+1\right)}dz \\ &= \frac{1}{\left(\sum_{j=1}^k u_j^{-1/\alpha}-k+1\right)^\alpha} \\ &\quad \int_0^\infty \frac{\left(\sum_{j=1}^k u_j^{-1/\alpha}-k+1\right)^\alpha}{\Gamma(\alpha)}z^{\alpha-1}e^{-z\left(\sum_{j=1}^k u_j^{-1/\alpha}-k+1\right)}dz \end{aligned}$$

We now integrate a Gamma density function with parameters $\alpha^* = \alpha$ and $\lambda^* = \left(\sum_{j=1}^k u_j^{-1/\alpha} - k + 1\right)$ over its whole range $(0, \infty)$. Thus, we find

$$F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) = \left\{\sum_{j=1}^k u_j^{-1/\alpha} - k + 1\right\}^{-\alpha},$$

where $u_j \in [0, 1]$ and $\alpha > 0$.

It is possible to show that this multivariate uniform distribution has a Kendall's tau:

$$\tau(X_i, X_j) = \tau(U_i, U_j) = \frac{1}{1+2\alpha}.$$

Then, when α decreases to 0, τ goes to one, i.e. the correlation approaches its maximum and by the preceding results on the Frechet upper bound, we have

$$\lim_{\alpha\to 0} F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) = \min[u_1, ..., u_k].$$

Also, when α increases to infinity, τ goes to 0, as well as the correlation, which gives

$$\lim_{\alpha\to\infty} F^{(\alpha)}_{U_1,...,U_k}(u_1, ..., u_k) = \prod_{j=1}^k u_j.$$

An example on the simulation of the Cook-Johnson copula will be presented at the end of this section. The Splus program used is shown in the appendix.

For a set of arbitrary marginal distributions, $F_{X_1}, ..., F_{X_k}$, we can deøne a joint *cdf* by

$$F_{X_1,...,X_k}(x_1, ..., x_k) = \left(\sum_{j=1}^{k} F_{X_j}(x_j)^{-1/\alpha} - k + 1 \right)^{-\alpha},$$

and we can also deøne a joint survivor function by

$$S_{X_1,...,X_k}(x_1, ..., x_k) = \left(\sum_{j=1}^{k} S_{X_j}(x_j)^{-1/\alpha} - k + 1 \right)^{-\alpha}.$$

We will now show that the Cook-Johnson copula is constructed through the distortion function $g(t) = \exp\left\{1 - t^{-1/\alpha}\right\}$, $\alpha > 0$. We have

$$g\left[S_{X_1,...,X_k}(x_1, ..., x_k)\right] = \prod_{j=1}^{k} g\left[S_{X_j}(x_j)\right],$$

which leads to

$$S_{X_1,...,X_k}(x_1, ..., x_k) = g^{-1}\left[\prod_{j=1}^{k} g\left[S_{X_j}(x_j)\right]\right].$$

If we let $h(t) = -\log g(t) = t^{-1/\alpha} - 1$, we get

$$\begin{aligned} S_{X_1,...,X_k}(x_1, ..., x_k) &= h^{-1}\left[\sum_{j=1}^{k} h\left[S_{X_j}(x_j)\right]\right] \\ &= \left(\sum_{j=1}^{k} \left[S_{X_j}(x_j)^{-1/\alpha} - 1\right] + 1 \right)^{-\alpha} \\ &= \left(\sum_{j=1}^{k} S_{X_j}(x_j)^{-1/\alpha} - k + 1 \right)^{-\alpha}, \end{aligned}$$

which is the same joint survivor function as previously.

In this dependency model no restriction is imposed on the marginal distributions F_{X_j} or S_{X_j}, $j = 1, ..., k$. However, we are restricted at the correlation parameters level, since this model requires to have the same set of Kendall's tau between any pair of risks.

Example 6.1. Let $S = X_1 + ... + X_{20}$, where the X_i's all have a Gamma$(\alpha = 2, \lambda = 2)$ distribution. If the X_i's are independent, we want to ønd the distribution of S. We ørst determine the moment generating function of X_i, $i = 1, ..., 20$

$$\begin{aligned} M_X(t) &= E\left[e^{tX}\right] \\ &= \int_0^\infty e^{tx} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{(\lambda - t)^\alpha} \int_0^\infty \frac{(\lambda - t)^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x(\lambda - t)} dx. \end{aligned}$$

Since we integrate the *pdf* of a Gamma with parameters $\alpha^* = \alpha$ and $\lambda^* = \lambda - t$ over its whole range $(0, \infty)$, it equals 1 and

$$M_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^\alpha.$$

In our case the *mgf* is

$$M_X(t) = \left(\frac{2}{2 - t}\right)^2.$$

We want the *mgf* of S, $M_S(t)$

$$\begin{aligned} M_S(t) &= E\left[e^{tS}\right] \\ &= E\left[e^{t(X_1 + ... + X_{20})}\right] \\ &= E\left[e^{tX_1 + ... + tX_{20}}\right] \\ &= E\left[e^{tX_1}\right] ... E\left[e^{tX_{20}}\right], \end{aligned}$$

since the X_i's are independent. Hence, we have

$$\begin{aligned} M_S(t) &= \prod_{j=1}^{20} E\left[e^{tX_j}\right] \\ &= \left(E\left[e^{tX}\right]\right)^{20} \\ &= \left(\frac{2}{2 + t}\right)^{40}, \end{aligned}$$

since the X_i's are identically distributed. It follows that the distribution of S is Gamma$(\alpha = 40, \lambda = 2)$.

Now, we assume that the X_i's are correlated, with $\tau(X_i, X_j) = 2/3$. We also assume a Cook-Johnson correlation structure, and we simulate 1000 samples of $X_1, ..., X_{20}$, i.e. 1000 values of S. The following plot (Figure 3) shows the *cdf* of S under the hypothesis of independence, and also the empirical distribution of S under the introduction of the correlation structure. The variance of F_S under Cook-Johnson is bigger than the variance of F_S under independence. Also, we should notice that the introduction of dependence allows extreme values, comparatively with the F_S under independence. The maximum claim amount

Figure 3. Graph of F_S for a Gamma$(40, 2)$ (independence hypothesis) and under a Cook-Johnson correlation with $\tau = 2/3$ (dependence hypothesis)

for S under the hypothesis of independence is around 25, while that for S with a correlation structure is around 50. The two distributions cross about in the center, around $(20, 0.5)$. Note that the expectation of a Gamma$(40, 2)$ is 20.

6.3. Normal Copula. As the Cook-Johnson copula, the normal copula does not impose any restriction on the choice of the marginal distributions. This copula also oœers more Æexibility, as it allows complete freedom in selecting Kendall's tau between any pair of risks. This copula also has the property to be easily implemented as a computational algorithm.

Theorem 6.3. Assume that $(Z_1, ..., Z_k)$ have a multivariate normal joint probability density function given by

$$f(z_1, ..., z_k) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left\{-\frac{1}{2}\underline{z}'\Sigma^{-1}\underline{z}\right\},$$

where $\underline{z} = (z_1, ..., z_k)$ with correlation coeŒcient $\rho_{ij} = \rho(Z_i, Z_j)$. Let $H(z_1, ..., z_k)$ be their joint cumulative distribution function. Then

$$C(u_1, ..., u_k) = H\left(\Phi^{-1}(u_1), ..., \Phi^{-1}(u_k)\right)$$

deønes a multivariate uniform cumulative distribution function called the normal copula.

For any set of given marginal *cdf*'s $F_1, ..., F_k$, the variables

$$X_1 = F_1^{-1}\left(\Phi\left(Z_1\right)\right), ..., X_k = F_k^{-1}\left(\Phi\left(Z_k\right)\right)$$

have a joint *cdf*

$$F_{X_1,...,X_k}\left(x_1, ..., x_k\right) = H\left(\Phi^{-1}\left(F_1\left(x_1\right)\right), ..., \Phi^{-1}\left(F_k\left(x_k\right)\right)\right)$$

with marginal *cdf*'s $F_1, ..., F_k$. The multivariate variables $\left(X_1, ..., X_k\right)$ have Kendall's tau

$$\tau\left(X_i, X_j\right) = \tau\left(Z_i, Z_j\right) = \frac{2}{\pi}\arcsin\left(\rho_{ij}\right)$$

and Spearman's rank correlation coeŒcients

$$\text{RankCorr}\left(X_i, X_j\right) = \text{RankCorr}\left(Z_i, Z_j\right) = \frac{6}{\pi}\arcsin\left(\frac{\rho_{ij}}{2}\right).$$

The analytical form of the normal copula is not very simple, but it makes possible to implement a very simple Monte Carlo simulation algorithm. In practice, it is common to have only some information about the correlation parameters, without necessarily knowing the exact multivariate distribution. In these cases, the normal copula allows to simulate the correlated variables in a simple way.

We presented only two copulas, but we have to notice that there exists an inønity of them. The Cook-Johnson and the normal copulas are among the most popular.

Conclusion

We have presented a set of tools for modeling and combining correlated risks. We discussed some common mixtures, components and distortion models, as well as copulas. We also presented some measures of correlation, as well as concepts relating to that. Using these methods along with some algorithms and Monte Carlo simulation methods may reveal to be very useful in modeling dependency.

Bibliography

[1] Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and C.J. Nesbitt (1997). Actuarial Mathematics. Schaumburg, Ill.: Society of Actuaries.

[2] Gaillardetz, P. (2000). Structures de dpendance en thorie du risque et pour les obligations avec risque de dfaut. Universit Laval, Qubec.

[3] Klugman, S.A., Panjer, H. H. and G. E. Willmot, (1998). Loss Models: From Data to Decisions, John Wiley, New York.

[4] Marceau, E. et H. Cossette (1999). The Discrete-time Risk Model with Correlated Classes of Business. Universit Laval, Qubec.

[5] Marceau, E., Cossette, H., Gaillardetz, P. et J. Rioux (1999).Dependance in the Individual Risk Model. Universit Laval, Qubec.

[6] Wang, S. (1998). Aggregation of Correlated Risk Portfolios: Models and Algorithms. CAS Report to the risk theory commitee.

Appendix

The Splus functions used are presented in this section.

```
stoploss <- function(fs, d1, pas = 1)–
#This function calculates the stop-loss(d) premiums
psld <- c()
s <- c(1:length(fs) - 1) * pas
for(i in 1:length(d1)) –
     s1 <- ((s > d1[i]) * (d1[i] >= 0)) * (s - d1[i])
     psld <- c(psld, sum(s1 * fs))
”
return(psld)
”
```

```
MC.poisson <- function(pas = 0.25, lambda = 100, mu = 2, sigma
= 1, v1 = 0:120 * 16)–
#This function calculates f(s), F(s) and the stop-loss(d) premium
for Compound-Poisson and Lognormal(mu,sigma)
Fx1 <- plnorm(0:(2ˆ12 - 1) * pas, mu, sigma)
fx1 <- diœ(c(Fx1, 1))
long <- 2ˆ15
fx1l <- c(fx1, rep(0, (long - length(fx1))))
Mx1 <- œt(fx1l)
Ms1 <- exp(lambda * (Mx1 - 1))
fs <- Re(œt(Ms1, inverse = T))[1:long]
fs <- (fs >= 0) * fs
fs <- fs/sum(fs)
Fs <- cumsum(fs)
plot(v1, Fs[v1/pas + 1], col = 1, type = jlj, main = jGraph of
F(s) (Poisson, Lognormal)j, ylab = jF(s)j, xlab = jsj)
```

```
psld <- stoploss(fs, v1, pas)
matplot(v1, psld, col = 1, type = jlj, main = jGraph of stop-loss
premium (Poisson, Lognormal)j, ylab= jpi(d)j, xlab = jdj)
˝
```

```
kendtau <- function(data)–
#This function calculate a non-parametric estimate of Kendall's
tau
#Take a matrix as argument, with a column of X and a column of
Y
#Number of pairs (X,Y)
    len <- nrow(data)
    #Set the ørst vector of row positions (in formula,position=i)
    v1 <- rep(1:(len - 1), (len - 1):1)
    v2 <- c()
    #Set the second vector of row positions (in formula, posi-
tion=j)
    for(i in 2:len) –
        v2 <- c(v2, i:len)
    ˝
#Diœerence between the dataset with row positions v1 and row
positions v2 (for v1<v2)
#Set of possibles diœerences for (i<j)
    diœ <- data[v1, ] - data[v2, ]
    #Use the signs of the diœerences to determine the non-parametric
estimate of the Kendall's tau
    kendalltau <- (2 * sum(sign(diœ[, 1] * diœ[, 2])))/(len * (len -
1))
    return(kendalltau)
˝
```

```
CJ <- function(n = 1000, k = 20, tau = 2/3, F = qgamma, param
= c(2, 2))–
```

```
    #This function simulates n samples of size k of rv's with a speciøed
marginal dist'n (F) and Kendall's tau (tau)
    alpha <- ((1/tau) - 1)/2
    #Create a matrix (n x k) of random exponential(1) all diœerent
    yij <- matrix(rexp(n * k), n, k)
    #Create a matrix (n x k) of random gamma(alpha,1), diœerent for
each row
    zi <- matrix(rep(rgamma(n, alpha), k), n, k)
    #Apply the transformation
    uij <- (1 + yij/zi)^( - alpha)
    #Invert the uij, giving a matrix with diœerent sample on each row
    samp <- matrix(F(uij, param[1], param[2]), n, k)
    return(samp)
    ˝
```

```
    Fn2 <- function(x, sample = x)–
    #This function calculates the empirical distribution of a dataset.
    #It takes as arguments the dataset x, and the sample of points for
which we want to know the empirical cdf (x by default).
    n <- length(x)
    out <- c()
    #Create a matrix with the diœerent values in column 1 and the
empirical cdf of each of these values in column 2.
    Fn <- cbind(rle(sort(x))$v, cumsum(rle(sort(x))$l)/length(x))
    #Calculate the empirical cdf at each point required in sample.
    for(i in 1:length(sample)) –
        out[i] <- max((((sample[i] - Fn[, 1]) >= 0) * 1) * Fn[, 2])
    ˝
    return(out)
    ˝
```

```
    n6 <- function(n = 20, alpha = 2, lambda = 2, s = 0:60)–
    #This function plots the independent cdf of S and the dependent
one on the same graph
    #Number 6(b)
```

```
#Generate 1000 samples of size 20 using the function CJ
emps <- apply(CJ(1000, 20, 2/3, qgamma, param = c(2, 2)), 1,
sum)
#Find the empirical dist'n using the function Fn2
empcdfs <- Fn2(sort(emps), s)
#Set matrices for the matplot
Fs <- cbind(pgamma(s, n * alpha, lambda), empcdfs)
matplot(s, Fs, xlab = JsJ, ylab = JF(s)J, main = JGraph of F(s)
(Cook-Johnson)J, type = JlJ, lty = c(3, 1), col = 1)
legend(35, 0.4, legend = c(JIndependentJ, JCook-JohnsonJ), lty=
c(3, 1))
″
```

Introduction

This project summarizes four articles treating the concept of dependence in risk theory. The ørst article presents principally a set of tools conducting to a better understanding of the other papers. The second and the third articles treat rather theoretical results, the authors being concerned with the bounds of risks in the sense of stop-loss order and with the bounds of the total claim of an insurance portfolio. In risk theory, the bounding of risks is a concept of interest, since it may be useful to classify and compare the risks, and also to determine the possible range for stop-loss premiums. The last article is written in a more practical way, and introduces the concept of ruin theory. A summary of the results of each paper is presented, along with some numerical examples and also a simulation study that aims to verify some results obtained by the authors of the last paper. It is important to note that although these articles present a lot of theorems and corollaries, the proofs are most of the time omitted, as the results stated often come from other articles. This is the reason why the proofs are rarely presented.

Comonotonicity, Correlation Order and Premium Principles

We ørst present a summary of the notions presented in Wang and Dhaene (1998). This paper contains a lot of theoretical concepts that will be useful in the following sections.

7. Stop-loss order and correlation order

A stop-loss premium is paid by an insurer to a reinsurer in order to protect himself against catastrophic claims (or catastrophic years). This means that for a retention level $d \geq 0$, the insurer will pay a maximum amount of d on the total amount of claims during a period, the excess being under the responsibility of the reinsurer. Thus, for a sum of risks S we can deøne the function $(S-d)_+ = \max(0, S-d)$ taking only positive values, leading to the stop-loss premium subject to a retention level d, $E(S-d)_+$.

We now introduce some concepts, beginning with that of stop-loss order.

Definition 7.1. A risk X is said to precede a risk Y in stop-loss order, written $X \leq_{sl} Y$, if for all retentions $d \geq 0$, the net stop-loss premium for risk X is smaller than that for risk Y:

$$E(X-d)_+ \leq E(Y-d)_+ .$$

We illustrate this with a very simple example.

Example 7.1. Suppose that two risks X_1 and X_2 are distributed as exponential distributions with parameter λ_1 and λ_2, respectively. For an exponential(λ), we ønd

$$\begin{aligned} E(X-d)_+ &= \int_d^{\infty} (x-d) f_X(x)\, dx \\ &= \int_d^{\infty} x\lambda e^{-\lambda x} dx - d\int_d^{\infty} \lambda e^{-\lambda x} dx \\ &= \frac{1}{\lambda e^{\lambda d}}. \end{aligned}$$

Thus, if $\lambda_1 \leq \lambda_2$, it implies that $1/\lambda_1 \geq 1/\lambda_2$ and also that $e^{-\lambda_1 d} \geq e^{-\lambda_2 d}$. Hence, $\lambda_1^{-1} e^{-\lambda_1 d} \geq \lambda_2^{-1} e^{-\lambda_2 d}$ and this is true for all $d \geq 0$. We can then say that X_2 precedes X_1 in stop-loss order, written $X_2 \leq_{sl} X_1$. Of course, if $\lambda_1 \geq \lambda_2$, then $\lambda_1^{-1} e^{-\lambda_1 d} \leq \lambda_2^{-1} e^{-\lambda_2 d}$ and $X_1 \leq_{sl} X_2$.

We also formalize in a deønition the notion of correlation for a pair of random variables with given marginals. $\mathcal{R}_2(F_X, F_Y)$ is considered as a class of elements, where the elements F_X and F_Y are the cumulative distribution functions (cdf 's) of the random variables X and Y, respectively.

Definition 7.2. Let (X_1, Y_1) and (X_2, Y_2) be two elements of $\mathcal{R}_2(F_X, F_Y)$. We say that (X_1, Y_1) is less correlated than (X_2, Y_2), written $(X_1, Y_1) \leq_{corr} (X_2, Y_2)$, if either of the following equivalent conditions holds:

1. For all non-decreasing functions f and g for which the covariances exist,
$$Cov\left(f\left(X_1\right), g\left(Y_1\right)\right) \leq Cov\left(f\left(X_2\right), g\left(Y_2\right)\right).$$
2. For all $x, y \geq 0$, the following inequality holds:
$$F_{X_1, Y_1}(x, y) \leq F_{X_2, Y_2}(x, y).$$

In other words, this deønition says that the more correlated of the two pairs is more likely to have closer amounts of claims than the less correlated pair. On the other hand, very diœerent claims amounts is an event that will more probably occurs to the less correlated pair of risks. We now present another concept of dependency.

Definition 7.3. The risks X and Y are said to be positively quadrant dependent, written PQD(X, Y), if either of the following equivalent conditions holds:

1. For all non-decreasing functions for which the covariances exist, we have that
$$Cov\left(f\left(X\right), g\left(Y\right)\right) \geq 0.$$
2. For all $x, y \geq 0$, the following inequality holds:
$$F_{X,Y}(x, y) \geq F_X(x) F_Y(y).$$

Deønition 1.2 and Deønition 1.3 are related in the sense that if a pair of risks (X, Y) is PQD, then this pair has more probability to exceed a value (x, y) than if the risks were independent. Hence, according to Deønition 1.2, (X, Y) are more correlated than independent risks. Note that a similar deønition exists for the opposite concept of negative quadrant dependency (NQD), resulting by changing the sign $\geq$ for $\leq$ in Deønition 1.3.

Example 7.2. Let X_1 and Y_1 be continuous random variables with joint probability density function (pdf)

$$f(x,y)=\begin{cases} x+y & \text{if } 0\le x\le 1, \quad 0\le y\le 1 \\ 0 & \text{elsewhere.}\end{cases}$$

We can calculate the bivariate cdf:

$$F_{X_1,Y_1}(x,y)=\int_0^y\int_0^x (t+s)\,dtds=\int_0^y\left(\frac{x^2}{2}+xs\right)ds$$
$$=\left.\frac{x^2s}{2}+\frac{xs^2}{2}\right|_0^y=\frac{x^2y}{2}+\frac{xy^2}{2}.$$

We can also ønd the marginal pdf's of X_1 and Y_1 by integration:

$$f(x)=\int_0^1 f(x,y)\,dy=\int_0^1 (x+y)\,dy$$
$$=\left.xy+\frac{y^2}{2}\right|_0^1=x+\frac{1}{2},$$

and

$$f(y)=\int_0^1 f(x,y)\,dx=\int_0^1 (x+y)\,dx$$
$$=\left.\frac{x^2}{2}+xy\right|_0^1=y+\frac{1}{2}.$$

If X_2 and Y_2 are considered as independent, their joint pdf is the product of the marginal pdf's, and is given by

$$f_{X_2,Y_2}(x,y)=\left(x+\frac{1}{2}\right)\left(y+\frac{1}{2}\right)$$
$$=xy+\frac{x}{2}+\frac{y}{2}+\frac{1}{4},$$

for $x,y\in[0,1]$.

In the case of independence, the cdf is then:

$$F_{X_2,Y_2}(x,y)=\int_0^y\int_0^x\left(ts+\frac{t}{2}+\frac{s}{2}+\frac{1}{4}\right)dtds$$
$$=\int_0^y\left(\frac{x^2s}{2}+\frac{x^2}{4}+\frac{xs}{2}+\frac{x}{4}\right)ds$$
$$=\left.\frac{x^2s^2}{4}+\frac{x^2s}{4}+\frac{xs^2}{4}+\frac{xs}{4}\right|_0^y$$
$$=\frac{x^2y^2}{4}+\frac{x^2y}{4}+\frac{xy^2}{4}+\frac{xy}{4}.$$

If we compare the two cdf's obtained, we find after some simplifications that

$$\begin{aligned} \frac{x^2y^2}{4}+\frac{x^2y}{4}+\frac{xy^2}{4}+\frac{xy}{4} &\geq \frac{x^2y}{2}+\frac{xy^2}{2} \\ (xy+1) &\geq (x+y), \end{aligned}$$

for $x, y \in [0,1]$. Note that we could also have found the joint cdf of X_2 and Y_2 with the relation $F_{X_2,Y_2}(x,y) = F_{X_2}(x) F_{Y_2}(y)$.

From the previous definitions, we proved that X_1 and Y_1 are NQD, and also that (X_1, Y_1) are less correlated than independent risks, (X_2, Y_2). We know that an equivalent condition to $F_{X_1,Y_1}(x,y) \leq F_{X_1}(x) F_{Y_1}(y)$ for NQD risks is that $Cov(f(X), g(Y)) \leq 0$, for all non-decreasing functions f and g for which the covariances exist. Hence, assuming $f(x) = x$ and $g(y) = y$ yields to $Cov(X_1, Y_1) \leq 0$. We can verify this by calculating $Cov(X_1, Y_1) = E(X_1Y_1) - E(X_1) E(Y_1)$:

$$\begin{aligned} E(X_1Y_1) = \int_0^1\int_0^1 ts(t+s)\,dtds &= \int_0^1\int_0^1 \left(t^2s+s^2t\right) dtds \\ &= \int_0^1 \left(\frac{s}{3}+\frac{s^2}{2}\right) ds = \left.\frac{s^2}{6}+\frac{s^3}{6}\right|_0^1 \\ &= \frac{1}{3}. \end{aligned}$$

Also,

$$\begin{aligned} E(X_1) = \int_0^1 tf(t)\,dt &= \int_0^1 \left(t^2+\frac{t}{2}\right) dt \\ &= \left.\frac{t^3}{3}+\frac{t^2}{4}\right|_0^1 = \frac{7}{12}, \end{aligned}$$

and similarly

$$E(Y_1) = \int_0^1 \left(s^2+\frac{s}{2}\right) ds = \frac{7}{12}.$$

We then obtain for the covariance

$$Cov(X_1, Y_1) = \frac{1}{3} - \left(\frac{7}{12}\right)^2 = -\frac{1}{144},$$

which is negative as expected.

With all these new concepts, it is now possible to introduce a relation between stop-loss order and correlation order.

Theorem 7.1. Let (X_1, Y_1) and (X_2, Y_2) be elements of $\mathcal{R}_2(F_X, F_Y)$. If

$$(X_1, Y_1) \leq_{corr} (X_2, Y_2),$$

then

$$X_1 + Y_1 \leq_{sl} X_2 + Y_2.$$

In words, this means that for two pairs of random variables with some given marginal distributions, the more correlated pair (in the sense of Deønition 1.3) follows the less correlated one in stop-loss order. That is, the more correlated pair has a bigger net stop-loss premium for all retention levels. This makes sense, if we think that the probability of facing catastrophic events in terms of insurance is higher for correlated risks. For instance, an insurer covering two residences near to the same river in a given region will more likely have to indemnify both risks than an insurer covering a residence near to a river and another one close to a ravine. The river represents a common risk of Æood for the houses of the ørst insurer, while two diœerent risks, a Æood and a landslide threaten the houses of the second insurer.

If we consider Example 1.2 then by Theorem 1.1, the independent risk (X_2, Y_2) precedes the risk (X_1, Y_1) in stop-loss order. This can be concluded from the result we found stating that the risk (X_2, Y_2) is less correlated than (X_1, Y_1).

We have seen in the last report that Frchet bounds are used to determine the limits of a bivariate cdf (we will see later that they can also be generalized to a multivariate cdf). Theorem 1.1 along with the concept of Frchet bounds bring another result on the concept of orders. We should precise that the inverse F^{-1} is deøned as $F^{-1}(q) = \inf\{x \in \mathbb{R} : F(x) \geq q\}$, where $0 < q < 1$.

Theorem 7.2. Let U be uniformly distributed on $[0,1]$. Then for any pair of risks (X, Y) the following ordering relations hold:

1. $F_X^{-1}(U), F_Y^{-1}(1-U) \leq_{corr} (X, Y) \leq_{corr} F_X^{-1}(U), F_Y^{-1}(U)$,
2. $F_X^{-1}(U) + F_Y^{-1}(1-U) \leq_{sl} X + Y \leq_{sl} F_X^{-1}(U) + F_Y^{-1}(U)$.

We should notice that Frchet bounds are now expressed as a function of U. With this representation, it is easier to see some concepts closely related to each bound. For the upper bound, X and Y both depend on U, and this underlines the strong dependence structure (comonotonicity). For the lower bound, since X is function of U and Y is function of $1-U$, this introduces a kind of negative association that we will deøne later as mutually exclusive risks.

This theorem states that Frchet bounds also constitute bounds for stop-loss premiums, and this is valid for any pair of risks with given marginals. We can øgure it with the following relation, where Frchet bounds are expressed in terms of the inverse cumulative distribution functions of X and Y:

$$\begin{aligned} & \int_0^1 \left[F_X^{-1}(q) + F_Y^{-1}(1-q) - d\right]_+ dq \\ \leq \; & E(X+Y-d)_+ \\ \leq \; & \int_0^1 \left[F_X^{-1}(q) + F_Y^{-1}(q) - d\right]_+ dq. \end{aligned}$$

We have seen in this section some new tools that will be useful for the next topic. We now move on premium principles and their relations according to the degree of dependence of underlying risks.

8. Premium principles

There exist diœerent methods to øx the premiums that will be charged to the policyholders. The net premium of a risk X (or of a sum of risks) is deøned as the expectation of this risk, $E(X)$. This means that some years, the insurer may get proøt or loss but on average, the insurer will have just enough money to respect his obligations. Since this kind of industry is not really advantageous for the insurer, he rarely charges only the net premium and habitually adds some risk load, which gives the risk-adjusted premium. The procedure allowing to obtain the risk-adjusted premium is called a premium principle, written π. A premium principle π is a mapping that assigns to any risk X a positive value $\pi(X)$, which is called the risk-adjusted premium. It is assumed that risks with the same cumulative distribution functions lead to the same risk-adjusted premiums.

Among the desirable properties for a premium principle, there is one stating that it should preserve stop-loss order, i.e. $X \leq_{sl} Y$ implies that $\pi(X) \leq \pi(Y)$. The following result then follows from Theorem 1.1:

Theorem 8.1. Let π be a premium principle which preserves stop-loss order, and (X_1, Y_1) and (X_2, Y_2) be elements of $\mathcal{R}_2(F_X, F_Y)$. If

$$(X_1, Y_1) \leq_{corr} (X_2, Y_2),$$

then

$$\pi(X_1 + Y_1) \leq \pi(X_2 + Y_2).$$

This results agrees with our previous reasoning on the stop-loss premiums and their dependence structure, i.e. an insurer should increase his premiums as the probabilities of occurrence of the risks subscribed increase. From Theorem 1.2 and Theorem 2.1, we introduce a corollary:

Corollary 8.2. Let π be a premium principle which preserves stop-loss order. Then, we have

$$\pi\left(F_X^{-1}(U)+F_Y^{-1}(1-U)\right)\leq\pi(X+Y)\leq\pi\left(F_X^{-1}(U)+F_Y^{-1}(U)\right).$$

Based on Frchet bounds expressed in terms of the inverse cumulative distribution functions, this corollary says that we can ønd bounds for the premiums obtained by a premium principle preserving stop-loss order. These bounds are found by applying the premium principle to each of the Frchet bounds. Hence, we can see that the upper bound is attained for comonotonic risks. This is not surprising, since such risks are an extension of the concept of perfect correlation, as each one is a bet on the same event and they do not hedge against each other. Hence, it is reasonable for an insurer to charge a bigger premium in such cases since he is more likely to have higher claims than for any other pair of risks. On the other hand, the lower bound is the opposite case, as the second element of such a pair of risks will more probably get a big claim for the ørst element getting a small claim, and vice versa. Hence, this pair of risks consists in an optimal hedge, and this is normal to ønd the lowest premium in this case.

If we consider Example 1.2, a premium principle preserving stop-loss order would imply that the premium for the risk X_1+Y_1 is smaller than that for X_2+Y_2, since X_2 and Y_2 are independent. If this is not the case, the premium principle does not have the property of preserving stop-loss order.

We now introduce the concept of additive premium principle. A premium principle is called additive when the single premium for a pair of risks is the same as the sum of the premiums for each of the risks taken individually. That is,

$$\pi(X+Y)=\pi(X)+\pi(Y).$$

A premium principle is said to be sub-additive if the sum of the premiums for the individual risks is greater than or equal to the single premium for the pair of risks, i.e.

$$\pi(X+Y)\leq\pi(X)+\pi(Y).$$

Conversely, the super-additive principle is the opposite of the sub-additive principle, and is obtained by replacing the sign $\leq$ by $\geq$:

$$\pi(X+Y)\geq\pi(X)+\pi(Y).$$

From Theorem 2.1, another corollary has been found:

Corollary 8.3. If a premium principle preserves stop-loss order and is additive for independent risks, then it is sub-additive for negative quadrant dependent risks, and super-additive for positive quadrant dependent risks:

$$\pi(X+Y) \leq \pi(X) + \pi(Y) \quad \text{if NQD}(X, Y),$$
$$\pi(X+Y) \geq \pi(X) + \pi(Y) \quad \text{if PQD}(X, Y).$$

Since the comonotonic risks are PQD, then a special case of this corollary is that a premium principle preserving stop-loss order that is additive for independent risks is super-additive for comonotonic risks. From this reasoning, it is sensible to believe that a premium principle preserving stop-loss order that is additive for comonotonic risks should be sub-additive for other risks, since we cannot ønd more correlated risks than comonotonic risks. This is formalized in the next corollary.

Corollary 8.4. If a premium principle preserves stop-loss order and is additive for comonotonic risks, then it is sub-additive:

$$\pi(X+Y) \leq \pi(X) + \pi(Y) \quad \text{for all risks } X \text{ and } Y.$$

A consequence of this corollary is that for such a premium principle, it is always advantageous for a policyholder to subscribe a single contract than to be protected by individual policies. We can call this phenomenon a volume discount.

The premium principle to adopt depends on each situation. For instance, an insurer may be tempted to use a super-additive principle for comonotonic risks, since he prefers two independent risks, which are safer in the sense of claims amounts, and may want to reÆect his preference in his prices. A coverage for comonotonic risks would then be more expansive than for individual risks:

$$\pi(X+Y) \geq \pi(X) + \pi(Y).$$

On the other hand, if the coverage for each risk can be split into individual risks, it may be better for the insurer to use a sub-additive principle for comonotonic risks,

$$\pi(X+Y) \leq \pi(X) + \pi(Y),$$

in order to avoid the splitting of risks by the policyholder. A good compromise for the insurer is to use of an additive premium principle for comonotonic risks, as this method avoids the splitting of risks, and does not give any volume discount. However, if we are in the case where the splitting of risks is not allowed, the insurer will obviously make more proøts by using a super-additive premium principle. Note that in this discussion, we do not consider the possibility of concurrence, which may change the rules of the game!

Example 8.1. Consider a premium principle adding the standard error to the net premium, i.e. for a risk X we have $\pi(X) = E(X) + \sigma(X)$. Consider also the random variables presented in Example 1.2. If we calculate the premium for (X_1, Y_1), we get

$$\begin{aligned}\pi(X_1 + Y_1) = E(X_1 + Y_1) + \sigma(X_1 + Y_1) \\ = E(X_1) + E(Y_1) \\ + (Var(X_1) + Var(Y_1) + 2Cov(X_1, Y_1))^{1/2}.\end{aligned}$$

As explained in the previous section, X_1 and Y_1 are NQD, and then $Cov(X_1, Y_1) \leq 0$. Now, we get for (X_2, Y_2)

$$\begin{aligned}\pi(X_2 + Y_2) = E(X_2 + Y_2) + \sigma(X_2 + Y_2) \\ = E(X_2) + E(Y_2) + (Var(X_2) + Var(Y_2))^{1/2},\end{aligned}$$

since X_2 and Y_2 are independent. Hence, it is easy to verify that

$$\pi(X_1 + Y_1) \leq \pi(X_2 + Y_2),$$

and this result agrees with Theorem 2.1.

This result is not additive for independent risks, since $\sigma\left(\sum_{i=1}^{n} X_i\right) \neq \sum_{i=1}^{n} \sigma(X_i)$. However, by Jensen's inequality, we ønd that $\sigma\left(\sum_{i=1}^{n} X_i\right) \leq \sum_{i=1}^{n} \sigma(X_i)$ since this function is convex. This implies that $\pi(X_2 + Y_2) \leq \pi(X_2) + \pi(Y_2)$ and then this premium principle is sub-additive for independent risks, which means it is also sub-additive for NQD risks as (X_1, Y_1).

However, if we consider a premium principle such that $\pi(X + Y) = E(X + Y) + Var(X + Y)$, then it is additive for independent risks since the covariance between X_2 and Y_2 is 0:

$$\begin{aligned}\pi(X_2 + Y_2) = E(X_2 + Y_2) + Var(X_2 + Y_2) \\ = E(X_2) + E(Y_2) + Var(X_2) + Var(Y_2) \\ = \pi(X_2) + \pi(Y_2).\end{aligned}$$

We should recall from Corollary 2.3 that in this case, this premium principle would be sub-additive for NQD risks (X_1, Y_1).

We ønally present a premium principle introduced by Wang, called Wang's premium principle:

$$H_g(X) = \int_0^{\infty} g(1 - F_X(x))\, dx = \int_0^1 F_X^{-1}(1 - q)\, dg(q),$$

where g is a non-decreasing concave function with $g(0) = 0$ and $g(1) = 1$. This principle allows a pretty simple interpretation: the original tail function of the risk, $1 - F_X(x)$, is replaced by a new tail function $g(1 - F_X(x))$, which gives more weight to the right-tail. In other words, this premium principle gives more probability to bigger claims, which is conservative for the insurer. Then, the risk-adjusted premium is computed by ønding the expectation of X under this new tail function. The following theorem states some properties of Wang's premium principle:

Theorem 8.5. Wang's premium principle preserves stop-loss order, i.e.

$$X \leq_{sl} Y \Rightarrow H_g(X) \leq H_g(Y).$$

Moreover, it is additive in the class of comonotonic risks,

$$H_g(X+Y) = H_g(X) + H_g(Y) \quad \text{for comonotonic risks } X \text{ and } Y.$$

It is important to precise that Wang's premium principle is the only way to get an additive premium principle preserving stop-loss order for comonotonic risks. It is not possible to get these two properties simultaneously from a premium principle outside from the class built by Wang.

The Safest Dependence Structure Among Risks

This paper investigates the dependence in Frchet spaces containing mutually exclusive risks. Some new concepts are ørst presented, allowing understanding and deepening the work that has been done on the bounds. The goal of this paper is to bound the aggregate claims of a portfolio, and then deduct from this work some results for the stop-loss premiums. Since the upper bound is a subject already studied, the authors focus on the lower bound. They found that for general risks (under certain conditions ensuring a proper cdf), the lower bound of the portfolio is given by mutually exclusive risks, which are associated with the lower Frchet bound, in the same way comonotonicity is related to the upper Frchet bound.

9. Introduction

The total amount of claims of a portfolio during a given period is the sum S of the risk amounts $X_1, ..., X_n$

$$S = \sum_{i=1}^{n} X_i,$$

for $i = 1, ..., n$. A random variable modelling the total claim of a policy during a period is called a risk and is non-negative with a ønite expectation.

In actuarial literature, the stop-loss premium is a concept of interest as it allows to quantify the risk related to a portfolio. In order to determine it, two elements must be known: the marginal distributions of each risk and the structure of dependence between them. Based on a concern of simpliøcation, a hypothesis of independence between the risks is generally used, which allows modelling the stop-loss premiums with the information on the marginals only. However, it is evident that in the case where the risks are correlated, the stop-loss premium is underestimated and this situation is rather dangerous for the insurer.

The paper of Dhaene and Denuit (1999) is concerned with the safest dependence structure among general risks, that is the structure of dependence giving rise to the smallest stop-loss premium. It has been already shown that the riskiest dependence structure, i.e. the one resulting in the biggest stop-loss premium, is given by the upper Frchet bound, or equivalently by comonotonic risks. However, this paper shows that the safest dependence structure is given, under certain conditions, by the lower Frchet bound. The goal of these conditions is to ensure that the lower Frchet bound is really a proper cumulative distribution function, because contrary to the upper bound, it is not always the case. In order to compare the riskiness of insurance portfolios, the concept used is that of stop-loss order, deøned previously. We now deøne new concepts.

10. Frchet spaces and Frchet bounds

The Frchet space $\mathcal{R}_n(F_{X_1},...,F_{X_n})$ consists of all the n-dimensional random vectors $\mathbf{X}=(X_1,...,X_n)$ having marginal distributions F_{X_1}, ..., F_{X_n}. Note that $F_{X_1},...,F_{X_n}$ are univariate cumulative distribution functions. For our purpose, we obviously work with risks, and then we consider only non-negative random variables with ønite expectation.

We now extend the deønition of the bivariate Frchet bounds presented in the ørst report to the multivariate Frchet bounds. For all $\mathbf{X}$ in $\mathcal{R}_n(F_{X_1},...,F_{X_n})$, the following inequality holds:

$$\max\left\{\sum_{i=1}^{n}F_{X_i}(x_i)-n+1,0\right\}\leq F_{\mathbf{X}}(\mathbf{x})\leq\min\left\{F_{X_1}(x_1),...,F_{X_n}(x_n)\right\},$$

for all $\mathbf{x}=(x_1,...,x_n)\in\mathbb{R}^n$.

It is interesting to notice that the upper bound can be reached. Such random variables in the Frchet space having the upper bound as a multivariate cumulative distribution function are said to be comonotonic. Comonotonic random variables may be expressed in function of inverse marginal distributions, i.e. given a random variable U uniformly distributed on $[0,1]$, the upper bound is the cdf of

$$\left(F_{X_1}^{-1}(x_1),...,F_{X_n}^{-1}(x_n)\right)\in\mathcal{R}_n(F_{X_1},...,F_{X_n}),$$

where the inverses of the F_{X_i}'s are deøned previously.

Contrary to the upper bound, when $n\geq 3$, the lower Frchet bound is not always a proper cdf and then the necessary and suŒcient condition for the lower bound to be a cdf is given by the following theorem.

Theorem 10.1. A necessary and suŒcient condition for the lower Frchet bound to be a cdf in $\mathcal{R}_n(F_{X_1}, ..., F_{X_n})$ is that either

1. $\sum_{j=1}^{n} F_{X_j}(x_j) \leq 1$ for all $\mathbf{x} \in \mathbb{R}^n$ with $0 < F_{X_j}(x_j) < 1$, $j = 1, ..., n$; or
2. $\sum_{j=1}^{n} F_{X_j}(x_j) \geq n - 1$ for all $\mathbf{x} \in \mathbb{R}^n$ with $0 < F_{X_j}(x_j) < 1$, $j = 1, ..., n$.

11. Stochastic Bounds on the Smallest and Largest Claims

It is possible to derive bounds for the distributions of smallest and largest claims of a portfolio. In order to do that, we present a result stating that when all the x_i are equal, the lower Frchet bound is attained.

Theorem 11.1. There exist $\mathbf{X} \in \mathcal{R}_n(F_{X_1}, ..., F_{X_n})$ such that

$$\Pr[\max\{X_1, ..., X_n\} \leq x] = \max\left\{\sum_{i=1}^{n} F_{X_i}(x) - n + 1, 0\right\},$$

for any $n \in \mathbb{N}_0$ and $x \in \mathbb{R}$.

The results for the bounds can now be expressed in a corollary.

Corollary 11.2. For any $\mathbf{X} \in \mathcal{R}_n(F_{X_1}, ..., F_{X_n})$,

$$\begin{aligned} 1 - \min\{F_{X_1}(x), ..., F_{X_n}(x)\} &\leq \Pr[\max\{X_1, ..., X_n\} > x] \\ &\leq \min\left\{1, \sum_{i=1}^{n} (1 - F_{X_i}(x))\right\}, \end{aligned}$$

for all $x \in \mathbb{R}$, and

$$\begin{aligned} \max\{F_{X_1}(x), ..., F_{X_n}(x)\} &\leq \Pr[\min\{X_1, ..., X_n\} \leq x] \\ &\leq \min\left\{1, \sum_{i=1}^{n} F_{X_i}(x)\right\}, \end{aligned}$$

for all $x \in \mathbb{R}$.

These bounds ønd an utility in term of premiums. Since they limit the range of possible values for the distributions of the smallest and largest claims, they may have an impact on the determination of individual premiums of some insurance contract and may be used in the calculation of stop-loss premiums as well.

12. Extremal Dependence Structures

Some studies have been carried out to determine the riskiest and the safest dependence structure among risks. Some results have been presented for the Frchet spaces $\mathcal{R}_n(F_{X_1},...,F_{X_n})$ of all n-dimensional multivariate risks $(X_1,...,X_n)$ with each X_i having a two-point distribution. We consider the case where this distribution has a probability mass at 0, and another one at $\alpha_i > 0$ only, and thus the probability at any other value is 0. It has been investigated that the most dangerous structure of dependence is given by the distributions achieving the upper Frchet bound, and then gives rise to the highest stop-loss premium. This result has also been extended to general risks and is formalized in the following theorem.

Theorem 12.1. Let U be a random variable uniformly distributed on $[0,1]$. Then,

$$\sum_{i=1}^{n} X_i \leq_{sl} \sum_{i=1}^{n} F_{X_i}^{-1}(U),$$

for any multivariate risk $\mathbf{X}$ in $\mathcal{R}_n(F_{X_1},...,F_{X_n})$.

This means that any sum of the X_i's is always smaller in stop-loss order than in the case of comonotonic risks. We are also interested in the safest dependence structure in the Frchet spaces $\mathcal{R}_n(F_{X_1},...,F_{X_n})$. It is normal to believe, by symmetry, that this result should be given by the lower Frchet bound. Since this bound is not always a proper cumulative distribution function, it is not possible to obtain a general result. It is the reason why the study is restricted to a Frchet class $\mathcal{R}_n(F_{X_1},...,F_{X_n})$ for which the condition

$$\sum_{i=1}^{n} q_i \leq 1, \tag{22}$$

where $q_i = 1 - F_{X_i}(0)$, $i = 1,...,n$, is satisøed. This condition ensures that, except for the probability mass at 0, the probability mass of the marginal distributions at the other values is at most 1. Then, from Theorem 4.1(2), (22) is a suŒcient condition for the lower Frchet bound to be a proper cdf.

The lower bound for two-point distributions for the marginals $F_{X_1},...,F_{X_n}$ has already been studied, and the result is stated in the following theorem.

Theorem 12.2. Consider a Frchet space $\mathcal{R}_n(F_{X_1},...,F_{X_n})$ satisfying (22), such that for $i = 1,...,n$, the F_{X_i} are two-point distributions with probability masses in 0 and $\alpha_i > 0$. Consider the risk

$\mathbf{X} \in \mathcal{R}_n\left(F_{X_1}, ..., F_{X_n}\right)$ with dependence structure given by

$$\Pr\left[X_i = \alpha_i, X_j = \alpha_j\right] = 0,$$

for all $i \neq j$. Then,

$$\sum_{i=1}^{n} X_i \leq_{sl} \sum_{i=1}^{n} Y_i,$$

holds for any $\mathbf{Y} \in \mathcal{R}_n\left(F_{Y_1}, ..., F_{Y_n}\right)$.

This means that for two-point distributions with one of the masses at 0, if there is at most one of the risks taking a value bigger than 0, then this distribution is smaller in stop-loss order than any other distribution of the Frchet space.

In order to generalize this result to the case of general risks, the notion of mutually exclusive risks is required.

Definition 12.1. The risks $X_1, ..., X_n$ are said to be mutually exclusive (or, equivalently, the multivariate risk $\mathbf{X}$ is said to possess this property) when

$$\Pr\left[X_i > 0, X_j > 0\right] = 0,$$

for all $i \neq j$.

Dhaene and Denuit (1999) presents numerous examples of such risks, from actuarial to ønancial applications. We can state, for instance, the case of an n-year endowment insurance split as an n-year pure endowment and an n-year insurance issued on the same individual. We can also think of a travel insurance providing a sum in case or disablement and a sum in case of death. Since it is impossible to be both disabled and dead at the same time, these risks are mutually exclusive.

We can now clarify the principal role of condition (22) in the theory of mutually exclusive risks.

Theorem 12.3. A Frchet space $\mathcal{R}_n\left(F_{X_1}, ..., F_{X_n}\right)$ contains mutually exclusive risks if, and only if, it satisøes (22).

Proof. See Dhaene and Denuit (1999). ∎

We know that the concept of comonotonicity corresponds to the upper Frchet bound. Similarly, it can be shown that the concept of mutually exclusive risks corresponds to the lower Frchet bound, and this is formalized in a theorem.

Theorem 12.4. Consider a Frchet space $\mathcal{R}_n(F_{X_1},...,F_{X_n})$ satisfying (22). The risk $\mathbf{X} \in \mathcal{R}_n(F_{X_1},...,F_{X_n})$ is said to be mutually exclusive if, and only if,

$$F_{\mathbf{X}}(\mathbf{x}) = \max\left\{\sum_{i=1}^{n} F_{X_i}(x) - n + 1, 0\right\},$$

for $\mathbf{x} \in \mathbb{R}^n$.

Proof. See Dhaene and Denuit (1999). ∎

Combining the two previous theorems yields that the condition (22) is satisøed if, and only if, the only possible cdf is given by the lower Frchet bound of $\mathcal{R}_n(F_{X_1},...,F_{X_n})$.

To summarize, we began by stating the condition to have a proper cdf as a Frchet lower bound. We then aŒrmed that this condition is a double implication of mutual exclusivity, which is itself a double implication of the Frchet lower bound. If we can verify one of these three cases, then the other two cases are also true, by Theorem 6.3 and Theorem 6.4. We verify these relations with a very simple example.

Example 12.1. Three mutually exclusive risks have the following discrete distribution:

	0	1	2	3
X_1	0.8	0.2	0	0
X_2	0.7	0	0	0.3
X_3	0.6	0	0.4	0

From (22), we ønd that $q_1 = 1 - 0.8 = 0.2$, $q_2 = 0.3$, $q_3 = 0.4$. The condition is then satisøed, since $\sum_{i=1}^{3} q_i = 0.9 < 1$. Since the risks are mutually exclusive, we have

$$\begin{aligned} F_{\mathbf{X}}(1) &= \Pr(X_1 = 1) + \Pr(X_1 = 0, X_2 = 0, X_3 = 0) \\ &= 0.2 + 1 - \sum_{i=1}^{3} q_i \\ &= 0.3. \end{aligned}$$

If we compute this with the Frchet lower bound, we ønd

$$\begin{aligned} F_{\mathbf{X}}(1) &= \max\{\Pr(X_1 \leq 1) + \Pr(X_2 \leq 1) + \Pr(X_3 \leq 1) - 3 + 1, 0\} \\ &= \max\{1 + 0.7 + 0.6 - 3 + 1, 0\} \\ &= 0.3. \end{aligned}$$

We can verify another value, say 2

$$\begin{aligned} F_{\mathbf{X}}(2) = \Pr(X_1 = 1) + \Pr(X_3 = 2) + \Pr(X_1 = 0, X_2 = 0, X_3 = 0) \\ = 0.2 + 0.4 + 1 - 0.9 \\ = 0.7. \end{aligned}$$

With the Frchet lower bound, we get

$$\begin{aligned} F_{\mathbf{X}}(2) = \max\{\Pr(X_1 \leq 2) + \Pr(X_2 \leq 2) + \Pr(X_3 \leq 2) - 3 + 1, 0\} \\ = \max\{1 + 0.7 + 1 - 3 + 1, 0\} \\ = 0.7. \end{aligned}$$

If we repeat this procedure for the values 0 and 3, we ønd that $F_{\mathbf{X}}(0) = 0.1$ and $F_{\mathbf{X}}(3) = 1$.

We should notice that when (22) is not satisøed, this procedure makes no sense, since $\Pr(X_1 = 0, ..., X_n = 0) = 1 - \sum_{i=1}^{n} q_i$ is a negative value.

We now bring our attention on the stop-loss premium. As a special case of a theorem stating that the expected utility is additive for a sum of mutually exclusive risks, it follows that

$$E\left[\left(\sum_{i=1}^{n} X_i - d\right)_+\right] = \sum_{i=1}^{n} E(X_i - d)_+$$

holds when X is mutually exclusive for any deductible $d \geq 0$.

Example 12.2. We consider the three risks of the previous example and we assume a deductible $d = 1$. We then have

$$\begin{aligned} E\left[\left(\sum_{i=1}^{3} X_i - 1\right)_+\right] &= (1-1)\Pr\left(\sum_{i=1}^{3} X_i = 1\right) \\ &+ (2-1)\Pr\left(\sum_{i=1}^{3} X_i = 2\right) \\ &+ (3-1)\Pr\left(\sum_{i=1}^{3} X_i = 3\right) \\ &= 0(0.2) + 1(0.4) + 2(0.3) \\ &= 1, \end{aligned}$$

and also

$$\sum_{i=1}^{3} E\left(X_i - 1\right)_+ = (1-1)\Pr\left(X_1 = 1\right) + (2-1)\Pr\left(X_1 = 2\right) + (3-1)\Pr\left(X_1 = 3\right) + ... + (3-1)\Pr\left(X_3 = 3\right)$$
$$= 0\,(0.2) + 1\,(0.4) + 2\,(0.3)$$
$$= 1.$$

This can be veriøed for any deductible $d \geq 0$. Some values are presented in the following table.

d	$E\left[\left(\sum_{i=1}^{3} X_i - d\right)_+\right]$	$\sum_{i=1}^{3} E\left(X_i - d\right)_+$
0	1.9	1.9
1	1	1
2	0.3	0.3
3	0	0

Theorem 6.2 can now be generalized.

Theorem 12.5. Consider a Frchet space $\mathcal{R}_n\left(F_{X_1}, ..., F_{X_n}\right)$ satisfying (22). Let $\mathbf{X}$ be a mutually exclusive risk in $\mathcal{R}_n\left(F_{X_1}, ..., F_{X_n}\right)$. Then,

$$\sum_{i=1}^{n} X_i \leq_{sl} \sum_{i=1}^{n} Y_i$$

holds for any $\mathbf{Y} \in \mathcal{R}_n\left(F_{Y_1}, ..., F_{Y_n}\right)$.

Proof. See Dhaene and Denuit (1999). ∎

Hence, mutually exclusive risks correspond to the lower Frchet bound, and lead to the safest dependence structure of a portfolio. This means that this kind of dependency gives rise to the smallest stop-loss premium of an insurance portfolio.

Does Positive Dependence Between Individual Risks Increase Stop-Loss Premiums?

This paper is, in a sense, very close of the previous article presented in this report. The authors are still concerned with the lower bound of an aggregation of risks, but instead of considering general risks, they now focus on positive cumulative dependent risks, a concept that will be presented in a short time. They found that in this case, such a portfolio is bounded below by the concept of independence between the risks. We ørst specify some necessary notation, and then pursue with the main result.

13. Introduction

The risk $\mathbf{X}^{\perp} = \left(X_1^{\perp}, ..., X_n^{\perp}\right)$ represents the independent version of the risk $\mathbf{X} = (X_1, ..., X_n)$. This means that for $i = 1, ..., n$, the random variables X_i and $X_i^{\perp}$ have the same marginal distribution, but their joint distribution is not the same. Since the random variables $X_1^{\perp}, ..., X_n^{\perp}$ are mutually independent, their joint distribution is given by the product of the marginals. Furthermore, the risk $\mathbf{X}^U = \left(X_1^U, ..., X_n^U\right)$ represents the comonotonic version of $\mathbf{X}$, which means that $X_i^U = F_{X_i}^{-1}(U)$, for $i = 1, ..., n$. Note that U denotes a random variable uniformly distributed on $[0, 1]$ and $F_{X_i}^{-1}$ is the quantile function associated to the distribution function F_{X_i} of X_i.

14. Positive Cumulative Dependence

The concept of positive quadrant dependence (PQD) has already been presented. The notion of positive cumulative dependence is now considered. For $\mathcal{I} \subset \{1, ..., n\}$, deøne $S_{\mathcal{I}}$ as the sum of the X_i's whose index is in $\mathcal{I}$, i.e. $S_{\mathcal{I}} = \sum_{i \in \mathcal{I}} X_i$. Risks are said to be positive cumulative dependent (PCD) if for any $\mathcal{I}$ and $j \notin \mathcal{I}$, $S_{\mathcal{I}}$ and X_j are PQD. This concept allows extending the concept of positive quadrant dependence to arbitrary dimensions, while keeping the intuitive meaning of this deønition. This means that the knowledge that one of the positive quadrant dependent random variables is large (X_j) increases the probability of the others $(S_{\mathcal{I}})$ to be large too. Thus, the probability of the aggregate claim (excluding the known risk) is more likely to be large, inÆuencing the stop-loss premium in the same way.

15. Main Result

The main result of this paper is stated in the following theorem.

Theorem 15.1. Let us consider PCD risks $X_1, ..., X_n$ with marginal distribution functions $F_{X_1}, ..., F_{X_n}$. Then, we have

$$X_1^{\perp} + ... + X_n^{\perp} \leq_{sl} X_1 + ... + X_n \leq_{sl} X_1^U + ... + X_n^U.$$

Proof. This theorem is proved by induction, see Denuit et al. (2001). ∎

This theorem provides bounds for positive cumulative dependent risks. Also, for PCD øxed marginals, the riskiest dependence structure is given by comonotonic risks, as mentioned previously. However, for the same marginals, the safest dependence structure is provided by mutual independence. It follows that making the assumption of mutual independence for PCD risks leads to an underestimation of the stop-loss premiums. As we have seen in Dhaene and Denuit (1999), the dependence does not always exist when the risks are not known to be PCD.

An obvious application of this result follows for a given class of premium principles. Let $H(\cdot)$ be a premium calculation principle assigning a premium amount $H(X)$ to any risk X. We assume that the distribution function of X completely determines the premium for X, and also that the premium principle $H(\cdot)$ preserves stop-loss order

$$X \leq_{sl} Y \Rightarrow H(X) \leq_{sl} H(Y).$$

Hence, for PCD risks $X_1, ..., X_n$, the stop-loss calculation premium principle and the previous theorem (9.1) yield the relation

$$H\left(\sum_{i=1}^{n} X_i^{\perp}\right) \leq H\left(\sum_{i=1}^{n} X_i\right) \leq H\left(\sum_{i=1}^{n} X_i^U\right).$$

This inequality means that the stop-loss premium for PCD risks is bounded by the comonotonic and the independent versions of the risks for øxed marginal distributions.

Example 15.1. Consider these two independent random variables:

$$\Pr(X_i = 0) = \Pr(X_i = 1) = \frac{1}{2}, \quad i = 1, 2.$$

Then, the distribution of $S_1 = X_1 + X_2$ is

$$\Pr(S_1 = 0) = \Pr(S_1 = 2) = \frac{1}{4} \quad \text{and} \quad \Pr(S_1 = 1) = \frac{1}{2}.$$

Now, consider two dependent random variables such that $Y_1 = X_1$ and

$$\Pr(Y_2 = 0) = \Pr(Y_2 = 1) = \frac{1}{4} \quad \text{and} \quad \Pr(Y_2 = Y_1) = \frac{1}{2}.$$

The distribution of $S_2 = Y_1 + Y_2$ is given by

$$\Pr(S_2 = 0) = \Pr(S_2 = 2) = \frac{3}{8} \quad \text{and} \quad \Pr(S_2 = 1) = \frac{1}{4}.$$

Finally, consider two random variables such that $Z_1 = X_1$ and

$$\Pr(Z_2 = Z_1) = 1.$$

Since the distribution is reaching the upper Frchet bound for bivariate risks, i.e. $F_{Z_1,Z_2}(z_1, z_2) = \min(F_{Z_1}(z_1), F_{Z_2}(z_2))$, the random variables Z_1 and Z_2 are comonotonic according to the deønition in Wang and Dhaene (1998). The distribution of $S_3 = Z_1 + Z_2$ is given by

$$\Pr(S_3 = 0) = \Pr(S_3 = 2) = \frac{1}{2} \quad \text{and} \quad \Pr(S_3 = 1) = 0.$$

Moreover, it is important to precise that X_1, Y_1, Z_1 have the same marginal distribution, while X_2, Y_2, Z_2 also have an identical marginal. In fact, they all have the same marginal distribution, but this is not required.

We can now calculate the stop-loss premiums for S_i, $i = 1, 2, 3$. We have to split the deductibles in two groups, that is $0 \leq d \leq 1$ and $1 < d \leq 2$. For $0 \leq d \leq 1$, we ønd

$$\begin{aligned} E(S_1 - d)_+ &= (1-d)\frac{1}{2} + (2-d)\frac{1}{4} = 1 - \frac{3}{4}d, \\ E(S_2 - d)_+ &= (1-d)\frac{1}{4} + (2-d)\frac{3}{8} = 1 - \frac{5}{8}d, \\ E(S_3 - d)_+ &= (1-d)\,0 + (2-d)\frac{1}{2} = 1 - \frac{1}{2}d. \end{aligned}$$

For $1 < d \leq 2$, we ønd

$$\begin{aligned} E(S_1 - d)_+ &= (2-d)\frac{1}{4} = \frac{1}{2} - \frac{1}{4}d, \\ E(S_2 - d)_+ &= (2-d)\frac{3}{8} = \frac{3}{4} - \frac{3}{8}d, \\ E(S_3 - d)_+ &= (2-d)\frac{1}{2} = 1 - \frac{1}{2}d. \end{aligned}$$

Figure 4. Stop-loss premiums for independent, dependent and comonotonic versions of a pair of risks.

By looking at these linear equations, this is evident that the result of Theorem 9.1 is verified. This is illustrated in Figure 1.

Consider now a premium principle π such that $\pi(S) = E(S) + Var(S)$. We need to calculate the variances for S_i, $i = 1, 2, 3$:

$$\begin{aligned} Var(S_1) &= E\left(S_1^2\right) - E(S_1)^2 \\ &= 1.5 - 1^2 = 0.5. \end{aligned}$$

For S_2, we find

$$\begin{aligned} Var(S_2) &= E\left(S_2^2\right) - E(S_2)^2 \\ &= 1.75 - 1^2 = 0.75, \end{aligned}$$

and finally, we have for S_3

$$\begin{aligned} Var(S_3) &= E\left(S_3^2\right) - E(S_3)^2 \\ &= 2 - 1^2 = 1. \end{aligned}$$

Since $E(S_i) = 1$ for $i = 1, 2, 3$, it is evident that $\pi(S_1) \leq \pi(S_2) \leq \pi(S_3)$ and the last relation mentioned is verified. As stated previously, this principle preserves stop-loss order and is additive for independent risks. It follows that it is super-additive for the other two cases presented.

The Discrete-Time Risk Model with Correlated Classes of Business

Cossette and Marceau (2000) examines the discrete-time risk model with correlated classes of business. The authors treat two diœerent ways to introduce dependence between the diœerent classes, and study the impact of these relations on the ønite-time ruin probabilities.

We consider throughout the concept of book of business, which is deøned as the union of disjoint classes of business, each having an aggregate distribution.

For a matter of simpliøcation, classes of business in an insurance book of business are traditionally assumed independent in risk theory. This assumption, however, is not always realistic as in practice there exist a lot of situations in which it is not veriøed. For instance, in the case of a natural disaster as a hurricane, the damages covered by homeowner and private passenger automobile insurance cannot be considered independent.

The probability of ruin in the discrete-time risk model studied in this paper is presented in Bowers et al. (1997). A brief description of the discrete-time model is introduced and the probability of ruin over ønite and inønite-time is deøned. A Poisson common shock model and a Negative Binomial (NB) component model are used to introduce a relation of dependence between the diœerent classes of business, and we aim to verify the impact discussed in the paper through simulation methods.

16. Discrete-time model

Assume the discrete-time process $\{U_n,\, n = 0, 1, 2, ...\}$ where U_n is the surplus for a book of business of an insurer at time n $(n = 0, 1, 2, ...)$, which is deøned as

$$U_n = u + cn - S_n, \tag{23}$$

where u is the initial surplus, c the premium income received during each period and S_n the total claim amount over the ørst n periods. It is also assumed that

$$S_n = W_1 + W_2 + ... + W_n, \tag{24}$$

where W_i represents the total claim amounts for the book of business in the period i and $\{W_i,\, i = 1, 2, ...\}$ is a sequence of independent and identically distributed (iid) random variables with $E(W_i) = \mu_W < c$. The probability distribution and density function of W_i $(i = 1, 2, ...)$ are denoted by F_W and f_W, respectively.

Given (24), (23) can be written as

$$U_n = u + (c - W_1) + (c - W_2) + ... + (c - W_n), \tag{25}$$

that is in grouping inÆows and outÆows for each period, where u is the initial surplus. The premium c is received at the beginning of the period, while the claims are paid at the end of the period. The insurer begins with an initial surplus u and then receive the premiums c. We assume $\mu_W < c$ and thus we deøne the security loading η_j such that $c = (1 + \eta_j)\mu_W$, where η_j is strictly positive. No interest income is assumed. At the end of the period, he pays the claims incurred and is left with a new surplus. If he does not have enough of the amount $u + c$ to pay the claims, his new surplus is negative and ruin occurs. If he has enough money to pay damages, he is left with a new positive surplus U_1 that will be used, along with c, to pay the claims of the second period, and so on.

When the surplus process goes under 0, that is when the cash inÆows of the insurer do not suŒce to pay the claims, then ruin occurs. Assume T is the time of ruin, deøned as

$$T = \inf(n, U_n < 0),$$

assuming that $T = \infty$ if $U_n \geq 0$ for all $n = 1, 2, ...$

Let $\psi(u, 1, n)$ be the ønite-time ruin probability over the periods 1 to n

$$\psi(u, 1, n) = P(T \leq n),$$

that is the probability that ruin occurs before or at time n (during the n ørst periods). When $n \to \infty$ in $\psi(u, 1, n)$, we have

$$\psi(u) = P(T < \infty),$$

which represents the inønite-time ruin probability. Conversely, the inønite and ønite-time horizon non-ruin probabilities are deøned as

$$\phi(u) = 1 - \psi(u), \quad \phi(u, 1, n) = 1 - \psi(u, 1, n),$$

respectively. Given (25), we can write this as a joint probability that the surplus stays over 0 and we have

$$\begin{aligned}\phi(u,1,n) &= P(U_1 \geq 0, U_2 \geq 0, ..., U_n \geq 0) \\ &= P(W_1 \leq u + c, W_1 + W_2 \leq u + 2c, \\ &\qquad ..., W_1 + W_2 + ... + W_n \leq u + nc).\end{aligned}$$

Let the non-ruin probability over the periods j to n be

$$\begin{aligned}\phi(y,j,n) &= P(W_j \leq y + c, W_j + W_{j+1} \leq y + 2c, \\ &\qquad ..., W_j + W_{j+1} + ... + W_{n-j} \leq y + (n-j)c),\end{aligned}$$

where y is the value of the surplus process at time j. It follows from $W_1, ..., W_n$ being iid that

$$\phi(y,2,n) = \phi(y,1,n-1),$$

that is for the same surplus y at period j, the non-ruin probability over a øxed number of periods is the same for all j. It is then possible to express the non-ruin probabilities using a renewal equation (see Cossette and Marceau (2000)). However, since exact calculations are diŒcult to carry on with this formulation, an algorithm approximating $\phi(u,1,n)$ has been proposed, which requires the discretization of the distribution function F_W. More details can be found in Cossette and Marceau (2000).

It is important to mention that the ruin deøned previously does not correspond to the bankruptcy of the insurance company. It refers to the insurance activities of a speciøc portfolio of risks.

17. Aggregation of dependent classes of business

17.1. Introduction. It is assumed that the book of business of the insurer is constituted of m dependent classes of business and that the total claim amounts for the book of business in period i is given by

$$W_i = W_{i,1} + W_{i,2} + ... + W_{i,m},$$

for $i = 1, 2, ...,$ where $W_{i,j}$ represents the total claim amounts of the period i for the jth class of business. For $i \neq i'$ (i.e. for two diœerent periods), W_i and $W_{i'}$ are supposed independent and identically distributed. The common probability distribution function of the random variables W_i $(i = 1, 2, ...)$ is denoted by F_W and we assume that W is a random variable with this probability distribution function. For a øxed period i $(i = 1, 2, ...)$, the diœerent classes of business are assumed dependent.

For the class of business j $(j = 1, ..., m)$ in the period i $(i = 1, 2, ...)$, $X_{i,j,k}$ represents the kth individual claim, and $N_{i,j}$ the number of claims. Then,

$$W_{i,j} = \sum_{k=1}^{N_{i,j}} X_{i,j,k}.$$

For a given class of business (j øxed), $F_{X^{(i)}}$ (with $F_{X^{(i)}}(0) = 0$), denotes the common distribution function of the iid random variables $X_{i,j,k}$ $(i = 1, 2, ...; k = 1, 2, ..., N_{i,j})$. Let $X^{(j)}$ be a random variable with this distribution function.

For j øxed, $N_{i,j}$ $(i = 1, 2, ...)$ are identically distributed random variables. Let $N^{(j)}$ be a random variable with their common distribution function. Similarly, the random variables $W_{i,j}$ $(i = 1, 2, ...)$, are supposed identically distributed. Let $W^{(j)}$ be a random variable with their common distribution function. The usual assumption that $X^{(j)}$ and $N^{(j)}$ are independent is also made.

For the class of business j and for any period i $(i = 1, 2, ...)$, the premium income is

$$\begin{aligned} c_j &= E\left(W^{(j)}\right)\left(1 + \eta_j\right) \\ &= E\left(X^{(j)}\right) E\left(N^{(j)}\right)\left(1 + \eta_j\right), \end{aligned}$$

for $j = 1, ..., m$, where η_j is the positive risk margin for the jth class of business. The premium income for the book of business in the period i $(i = 1, 2, ...)$ is $c = c_1 + ... + c_m$.

17.2. Poisson model with common shock. The common shock model is presented in Wang (1998). Consider a book of business divided in three $(m = 3)$ dependent classes of business. Note that it is easy to generalize this model to any number m of dependent classes of business. It is assumed that a common shock aœects the three classes of business and that another common shock has an impact on each couple of classes.

Given the previous assumptions of identical distribution of the random variables $N_{i,1}, N_{i,2}$ and $N_{i,3}$ for any øxed period i $(i = 1, 2, ...)$, $N^{(j)}$ $(j = 1, 2, 3)$ is deøned as follows:

$$\begin{aligned} N^{(1)} &= N^{(11)} + N^{(12)} + N^{(13)} + N^{(123)}, \\ N^{(2)} &= N^{(22)} + N^{(12)} + N^{(23)} + N^{(123)}, \\ N^{(3)} &= N^{(33)} + N^{(13)} + N^{(23)} + N^{(123)}, \end{aligned}$$

where $N^{(uv)} \sim \text{Poisson}(\lambda_{uv})$ for $u, v = 1, 2, 3$ and $N^{(123)} \sim \text{Poisson}(\lambda_{123})$.

Since the distribution of the convolution of n independent Poisson random variables X_i with parameter λ_i is Poisson with parameter $\sum_{i=1}^{n} \lambda_i$ (distribution inønitely divisible), then

$$N^{(r)} \sim \text{Poisson}\left(\lambda_r\right),$$

for $r = 1, 2, 3$, with

$$\begin{aligned} \lambda_1 &= \lambda_{11} + \lambda_{12} + \lambda_{13} + \lambda_{123}, \\ \lambda_2 &= \lambda_{22} + \lambda_{12} + \lambda_{23} + \lambda_{123}, \\ \lambda_3 &= \lambda_{33} + \lambda_{13} + \lambda_{23} + \lambda_{123}. \end{aligned}$$

Also,

$$Cov\left(N^{(u)}, N^{(v)}\right) = Var\left(N^{(uv)}\right) + Var\left(N^{(123)}\right), \tag{26}$$

for $u \neq v$, and then

$$\begin{aligned} Cov\left(N^{(1)}, N^{(2)}\right) &= \lambda_{12} + \lambda_{123}, \\ Cov\left(N^{(1)}, N^{(3)}\right) &= \lambda_{13} + \lambda_{123}, \\ Cov\left(N^{(2)}, N^{(3)}\right) &= \lambda_{23} + \lambda_{123}. \end{aligned}$$

17.2.1. Application and Simulation. Cossette and Marceau (2000) studies the impact of the probability of ruin for a Poisson model with common shock between two classes of business of an insurance book of business. We verify the numerical results through a simulation study. Consider the following example:

$$\begin{array}{lrl} \text{Book of business \#1:} & X^{(1)} \sim & \text{Weibull}\,(0.5, 1/0.5625) \\ & N^{(1)} \sim & \text{Poisson}\,(4) \\ \text{Book of business \#2:} & N^{(1)} \sim & \text{Exponential}\,(1.125) \\ & N^{(2)} \sim & \text{Poisson}\,(4) \end{array}.$$

The moments of these random variables, as well as the moments of the random variable W has been obtained through simulation and appear in Table 1 for each book of business.

We present in Table 2 some correlation parameters for the cases where the coeŒcient of correlation between $N^{(1)}$ and $N^{(2)}$, $\rho\left(N^{(1)}, N^{(2)}\right)$, takes the values 0, 0.25, and 0.75. To obtain these values, we ørst determine the value of λ_{12}. From (26), we know that $\lambda_{12} = Var\left(N^{(12)}\right) = Cov\left(N^{(1)}, N^{(2)}\right)$. By the deønition of the covariance, we know that $Cov\left(N^{(1)}, N^{(2)}\right) = \rho\left(N^{(1)}, N^{(2)}\right)\sigma_{N^{(1)}}\sigma_{N^{(2)}}$, and thus the parameter λ_{12} is found to be $\lambda_{12} = \rho\left(N^{(1)}, N^{(2)}\right)\sigma_{N^{(1)}}\sigma_{N^{(2)}}$. Then, with the value of the coeŒcient of correlation along with the moments found previously, we can easily determine the value of λ_{12}. Note that when $\rho\left(N^{(1)}, N^{(2)}\right) = 0$, we are in the case of uncorrelated numbers of claims.

The third table contains the probability of ruin $\psi(u, 1, 20)$, that is the probability that the ruin occurs within the ørst 20 periods, for different degrees of dependence (given by $\rho\left(N^{(1)}, N^{(2)}\right)$ taking the values 0, 0.25, or 0.75). To simulate the probabilities of ruin, we ørst simulate 20 numbers of claims for each of the variables $N_i^{(12)}, N_i^{(11)}, N_i^{(22)}$ $(i = 1, ..., 20)$, in order to determine the number of claims occurring in each of the period considered. We then simulate $N_i^{(1)}$ and $N_i^{(2)}$ claim amounts for each period $(i = 1, ..., 20)$, in order to determine the total amount of claims for each variable $N_i^{(j)}$ $(j = 1, 2; i = 1, ..., 20)$. Since we now have the total cash outÆows $N_i^{(1)} + N_i^{(2)}$ for each period $(i = 1, ..., 20)$, we can compare it with the surplus process, using cash inÆows for diœerent values of the surplus u, and with a relative security margin of 15%. If there is at least one period where the outÆows are bigger than the surplus, then ruin occurs. If ruin does not occur, then the surplus is always at least as big as the amount of cash outÆows. By repeating this procedure numerous times, we can determine the proportion of times ruin occurs.

Note that we have experimented some problems in reproducing the results of Cossette and Marceau (2000) by simulation, due to some errors in the deønition of the distributions and also a lack of setting parameters (as the margin security). In order to øx up these problems, we have set our own parameters, and it is why some values may diœer from the results of the paper. However, having the same expectation as them, we are in general pretty close to the results they obtained. Also, with the results obtained by simulation, we are pretty conødent that our parameters are the same as the ones they used (or pretty close).

In general, we observe that the ruin probability increases with the degree of dependence between the number of claims, and vanishes as the initial surplus gets bigger.

	Class (1)	Class (2)
$E\left(X^{(i)}\right)$	1.12989	1.12655
$E\left(X^{(i)\,2}\right)$	7.69551	2.53913
$E\left(N^{(i)}\right)$	3.99858	3.99462
$Var\left(N^{(i)}\right)$	3.99132	3.99023
$E\left(W^{(i)}\right)$	4.51798	4.50016
$Var\left(W^{(i)}\right)$	30.7711	10.1428

Table 1. Moments of $X^{(i)}$, $N^{(i)}$, *and* $W^{(i)}$

$\rho\left(N^{(1)}, N^{(2)}\right)$	0	0.25	0.75
λ_{12}	0.000000	1.000000	3.000000
$Cov\left(N^{(1)}, N^{(2)}\right)$	−0.000717	1.008920	3.022179
$Cov\left(W^{(1)}, W^{(2)}\right)$	0.083057	1.310217	3.651093
$\rho\left(W^{(1)}, W^{(2)}\right)$	0.004726	0.073979	0.210962

Table 2. Correlation parameters

u	$\psi(u,1,20,0)$	$\psi(u,1,20,0.25)$	$\psi(u,1,20,0.75)$
0	0.626	0.654	0.659
10	0.330	0.351	0.390
20	0.154	0.190	0.220
30	0.078	0.094	0.109
40	0.043	0.047	0.062
50	0.023	0.023	0.030
60	0.011	0.011	0.014
70	0.006	0.003	0.008
80	0.002	0.002	0.004
90	0.000	0.000	0.001
100	0.000	0.000	0.001
110	0.000	0.000	0.000
120	0.000	0.000	0.000
130	0.000	0.000	0.000
140	0.000	0.000	0.000
150	0.000	0.000	0.000

Table 3. Ruin probabilities $\psi(u,1,20)$ *for the Poisson model*

17.3. Negative Binomial model with common component.
It is well known that for a Poisson random variable, the expectation is equal to the variance. When modelling the number of claims N in practice, however, this is not always the case and sometimes, a distribution with variance bigger than the expectation ($\mathrm{Var}(N) > E(N)$) may be needed. In such situations, the Negative Binomial is often used to model claim numbers since it has this property. The probability function of a random variable N having this distribution is

$$P(N=n) = \binom{\alpha+n-1}{\alpha-1}\left(\frac{1}{1+\beta}\right)^{\alpha}\left(\frac{\beta}{1+\beta}\right)^{n},$$

for $\alpha, \beta > 0$, $n = 0, 1, 2, \ldots$. The mean is given by $\mu = \alpha\beta$ and the variance is $\sigma^2 = \alpha\beta(1+\beta) = \mu(1+\beta)$.

It is possible to adapt the construction of the common shock model to a Negative Binomial (see Wang (1998)). We consider the special case of a book of business subdivided in three dependent classes of business. The number of claims in the jth $(j = 1, 2, 3)$ class of business is assumed to be the sum of two random variables. The ørst random variable, $N^{(jj)}$, is speciøc for each class and is independent of the speciøc random variables of the other classes. The second random variable, $N^{(j0)}$, $(j = 1, 2, 3)$, is assumed to be dependent on the second random variable of the other classes. For $(i = 1, 2, \ldots)$, $N^{(j)}$ $(j = 1, 2, 3)$ is deøned as

$$N^{(j)} = N^{(jj)} + N^{(j0)},$$

where

$$\begin{aligned} N^{(jj)} &\sim NB\left(\alpha_{jj}, \beta_j\right), \\ N^{(j0)} &\sim NB\left(\alpha_0, \beta_j\right), \end{aligned}$$

for $j = 1, 2, 3$.

Since the Negative Binomial is a distribution inønitely divisible, the sum of n independent Negative Binomial random variables X_i with parameters (α_i, β) is Negative Binomial with parameters $\left(\sum_{i=1}^{n} \alpha_i, \beta\right)$. Hence, the distribution of $N^{(j)}$ becomes

$$N^{(j)} \sim NB\left(\alpha_j, \beta_j\right),$$

for $j = 1, 2, 3$, where $\alpha_j = \alpha_{jj} + \alpha_0$. For a øxed period i $(i = 1, 2, \ldots)$, it is assumed that the random variables $N^{(jj)}$ $(j = 1, 2, 3)$ are independent. However, a relation of dependence is introduced between the random variables $N^{(j0)}$ $(j = 1, 2, 3)$, as they are modeled by a common Poisson-Gamma mixture:

1. $N^{(j0)} \mid \Theta = \theta \sim \mathrm{Poisson}\left(\theta\beta_j\right) \quad (j = 1, 2, 3)$

2. $\Theta \sim \text{Gamma}(\alpha_0, 1)$
3. $N^{(j0)} \mid \Theta = \theta$ are independent $(j = 1, 2, 3)$

It is possible to develop an expression for the covariance of the random variables $N^{(j)}$ $(j = 1, 2, 3)$:

$$\begin{aligned} Cov\left(N^{(u)}, N^{(v)}\right) &= Cov\left(N^{(uu)} + N^{(u0)}, N^{(vv)} + N^{(v0)}\right) \\ &= Cov\left(N^{(u0)}, N^{(v0)}\right), \end{aligned}$$

since the other random variables of this expression are independent. By conditioning on Θ, we get

$$\begin{aligned} Cov\left(N^{(u0)}, N^{(v0)}\right) &= Cov\left(E\left(N^{(u0)} \mid \Theta\right), E\left(N^{(v0)} \mid \Theta\right)\right) \\ &\quad +E\left(Cov\left(N^{(u0)}, N^{(v0)} \mid \Theta\right)\right) \\ &= Cov\left(\Theta\beta_u, \Theta\beta_v\right), \end{aligned}$$

since $N^{(j0)} \mid \Theta = \theta$ $(j = 1, 2, 3)$ are independent. We then obtain

$$\begin{aligned} Cov\left(N^{(u0)}, N^{(v0)}\right) &= \beta_u\beta_v Var\left(\Theta\right) \\ &= \alpha_0\beta_u\beta_v. \end{aligned} \tag{27}$$

17.3.1. Application and Simulation. Cossette and Marceau (2000) studies the impact of the probability of ruin for a Negative Binomial model with common component between two classes of business of an insurance book of business. We verify the numerical results through a simulation study. Consider the following example:

$$\begin{aligned} \text{Book of business \#1:}\quad & X^{(1)} \sim \text{Weibull}\left(0.5, 1/0.5625\right) \\ & N^{(1)} \sim NB\left(1, 4\right) \\ \text{Book of business \#2:}\quad & N^{(1)} \sim \text{Exponential}\left(1.125\right) \\ & N^{(2)} \sim NB\left(1, 4\right) \end{aligned}$$

The moments of these random variables, as well as the moments of the random variable W has been obtained through simulation and appear in Table 4 for each book of business.

We present in Table 5 some correlation parameters for the cases where the coeŒcient of correlation between $N^{(1)}$ and $N^{(2)}$, $\rho\left(N^{(1)}, N^{(2)}\right)$, takes the values 0, 0.25, and 0.75. To obtain these values, we ørst determine the value of α_0. From (27), we know that $\alpha_0\beta_1\beta_2 = Cov\left(N^{(1)}, N^{(2)}\right)$. By the deønition of the covariance, we know that $Cov\left(N^{(1)}, N^{(2)}\right) = \rho\left(N^{(1)}, N^{(2)}\right)\sigma_{N^{(1)}}\sigma_{N^{(2)}}$, and thus the parameter α_0 is found to be

$$\alpha_0 = \frac{\rho\left(N^{(1)}, N^{(2)}\right)\sigma_{N^{(1)}}\sigma_{N^{(2)}}}{\beta_1\beta_2}.$$

Then, with the value of the coeŒcient of correlation along with the moments found previously, we can determine easily the value of α_0. Note that when $\rho\left(N^{(1)}, N^{(2)}\right) = 0$, we are in the case of uncorrelated numbers of claims.

Due to the impossibility of the package SPLUS to simulate Negative Binomial distributions with a non-integer value for the parameter α, we had to use the fact that a Negative Binomial is a Poisson-Gamma mixture. We simulate a value of a Gamma with the appropriate parameters, which will determine the parameter of the Poisson we have to simulate. However, this double simulation brings a little bit more variance in our correlation parameters. We also used this method to determine the probabilities of ruin, but the impact of the double simulation is practically unobservable for this case.

The sixth table contains the probability of ruin $\psi(u, 1, 20)$, that is the probability that the ruin occurs within the ørst 20 periods, for different degrees of dependence (given by $\rho\left(N^{(1)}, N^{(2)}\right)$ taking the values 0, 0.25, or 0.75). To simulate the probabilities of ruin, we ørst simulate 20 numbers of claims for each of the variables $N_i^{(0)}, N_i^{(11)}, N_i^{(22)}$ $(i = 1, ..., 20)$, through the Poisson-Gamma mixture, to determine the number of claims occurring in each of the period considered. We then simulate a number $N_i^{(j)}$ $(j = 1, 2; i = 1, ..., 20)$ of claims amounts for each variable and each period, in order to determine the total amount of claims for each variable $N_i^{(j)}$ $(j = 1, 2; i = 1, ..., 20)$. Since we now have the total cash outÆows $N_i^{(1)} + N_i^{(2)}$ for each period $(i = 1, ..., 20)$, we can compare it with the surplus process, using cash inÆows for different values of the surplus u, and with a relative security margin of 15%. If there is at least one period where the outÆows are bigger than the surplus, then ruin occurs. If ruin does not occur, the surplus is always at least as big as the amount of cash outÆows. By repeating this procedure numerous times, we can determine the proportion of times ruin occurs.

Note that we have experimented some problems in reproducing the results of Cossette and Marceau (2000) by simulation, due to some errors in the deønition of the distributions and also a lack of setting parameters (as the margin security). In order to øx up these problems, we have set our own parameters, and it is why some values may diœer from the results of the paper. However, having the same expectation as them, we are in general pretty close to what they obtained. Also, with the results obtained by simulation, we are pretty conødent that our parameters are the same as the ones they used (or pretty close).

In general, we observe that the ruin probability increases with the degree of dependence between the number of claims, and vanishes as the initial surplus gets bigger. We should also notice that the ruin probabilities and the correlation parameters are higher for the Negative Binomial case, compared to the Poisson case.

	Class (1)	Class (2)
$E\left(X^{(i)}\right)$	1.12450	1.12410
$E\left(X^{(i)\,2}\right)$	7.69298	2.52608
$E\left(N^{(i)}\right)$	4.03035	4.00423
$Var\left(N^{(i)}\right)$	20.1298	19.9685
$E\left(W^{(i)}\right)$	4.53212	4.50117
$Var\left(W^{(i)}\right)$	51.3633	30.2877

Table 4. Moments of $X^{(i)}$, $N^{(i)}$, *and* $W^{(i)}$

$\rho\left(N^{(1)}, N^{(2)}\right)$	0	0.25	0.75
α_0	0	0.3125	0.9375
$Cov\left(N^{(1)}, N^{(2)}\right)$	−0.096754	5.348388	15.77930
$Cov\left(W^{(1)}, W^{(2)}\right)$	0.037190	5.627755	19.97331
$\rho\left(W^{(1)}, W^{(2)}\right)$	0.000922	0.139748	0.520756

Table 5. Correlation parameters

u	$\psi(u,1,20,0)$	$\psi(u,1,20,0.25)$	$\psi(u,1,20,0.75)$
0	0.690	0.694	0.696
10	0.460	0.492	0.529
20	0.322	0.347	0.402
30	0.213	0.234	0.293
40	0.163	0.169	0.209
50	0.079	0.116	0.139
60	0.044	0.080	0.105
70	0.019	0.055	0.074
80	0.010	0.031	0.057
90	0.006	0.015	0.043
100	0.005	0.008	0.028
110	0.002	0.005	0.020
120	0.001	0.004	0.012
130	0.001	0.002	0.006
140	0.000	0.002	0.002
150	0.000	0.001	0.001

Table 6. Ruin probabilities $\psi(u,1,20)$ *for the NB model*

Conclusion

We have presented a set of tools for studying the properties of correlated risks. Among other things, we introduced the concept of stop-loss order and studied many notions and results related to that. We presented bounds for dependent risks, found from the concept of Frchet bounds. These inequalities ønd useful applications for the concept of stop-loss premiums in risk theory, which is an important concept in insurance. We ønally introduced the ruin theory and presented the model for classes of business. We simulated some results of the discrete-time version of ruin theory based on an example presenting two ways to introduce dependence in a model.

Bibliography

[1] Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and C.J. Nesbitt (1997). Actuarial Mathematics. Schaumburg, Ill.: Society of Actuaries.

[2] Cossette, H., Marceau, E. (2000). The Discrete-time Risk Model with Correlated Classes of Business. Insurance: Mathematics and Economics 26, 133-149.

[3] Denuit, M., Dhaene, J., Ribas, C. (2001). Does Positive Dependence Between Individual Risks Increase Stop-Loss Premiums? Insurance: Mathematics and Economics 28, 305-308.

[4] Dhaene, J., Denuit, M. (1999). The Safest Dependence Structure Among Risks. Insurance: Mathematics and Economics 25, 11-21.

[5] Wang, S. (1998). Aggregation of Correlated Risk Portfolios: Models and Algorithms. CAS Report to the risk theory commitee.

[6] Wang, S., Dhaene, J. (1998). Comonotonicity, Correlation Order and Premium Principles. Insurance: Mathematics and Economics 22, 235-242.

SPLUS Functions

18. Poisson model with common shock

18.1. Moments.

```
function(part = ¡A¡, S = 20000, we1 = 0.5, we2
= 0.5625, po1 = 4, ex = 1/1.125, po2 = 4, rho = 0.25)
–
if(part == ¡A¡) –
X1 <- rweibull(S, we1, we2)
EX1 <- mean(X1)
EX12 <- mean(X1^2)
N1 <- rpois(S, po1)
EN1 <- mean(N1)
VN1 <- var(N1)
EW1 <- EX1 * EN1
VW1 <- EX12 * EN1
X2 <- rexp(S, ex)
EX2 <- mean(X2)
EX22 <- mean(X2^2)
N2 <- rpois(S, po2)
EN2 <- mean(N2)
VN2 <- var(N2)
EW2 <- EX2 * EN2
VW2 <- EX22 * EN2
resul <- matrix(c(EX1, EX12, EN1, VN1, EW1, VW1, EX2, EX22,
EN2, VN2, EW2, VW2), 6, 2)
return(resul)
˝
if(part == ¡B¡) –
N0 <- 0
if(rho != 0) –
N0 <- rpois(S, rho * po1)
˝
N11 <- rpois(S, (1 - rho) * po1)
N22 <- rpois(S, (1 - rho) * po2)
```

```
N1 <- N0 + N11
N2 <- N0 + N22
CN12 <- var(matrix(c(N1, N2), S, 2))
W1 <- c()
W2 <- c()
for(i in 1:S) –
W1 <- c(W1, sum(rweibull(N1[i], we1, we2)))
W2 <- c(W2, sum(rexp(N2[i], ex)))
”
CW12 <- var(matrix(c(W1, W2), S, 2))
COW12 <- cor(matrix(c(W1, W2), S, 2))
return(CN12[1, 2], CW12[1, 2], COW12[1, 2])
”
”
N2 <- N0 + N22
CN12 <- var(matrix(c(N1, N2), S, 2))
W1 <- c()
W2 <- c()
for(i in 1:S) –
W1 <- c(W1, sum(rweibull(N1[i], we1, we2)))
W2 <- c(W2, sum(rexp(N2[i], ex)))
”
CW12 <- var(matrix(c(W1, W2), S, 2))
COW12 <- cor(matrix(c(W1, W2), S, 2))
return(CN12[1, 2], CW12[1, 2], COW12[1, 2])
”
”
CW12 <- var(matrix(c(W1, W2), S, 2))
COW12 <- cor(matrix(c(W1, W2), S, 2))
return(CN12[1, 2], CW12[1, 2], COW12[1, 2])
”
”
```

18.2. Ruin. function(part = ¡A¡, S = 1000, we1 = 0.5, we2 = 0.5625, po1 = 4, ex = 1/1.125, po2 = 4, rho = 0.25, u = c(0, 10, 30, 90), theta = 0.15)

```
–
if(part == ¡A¡) –
N0 <- 0
if(rho != 0) –
N0 <- rpois(20 * S, rho * po1)
”
```

```
N1 <- N0 + rpois(20 * S, (1 - rho) * po1)
N2 <- N0 + rpois(20 * S, (1 - rho) * po2)
W1 <- c()
W2 <- c()
for(i in 1:(20 * S)) –
W1 <- c(W1, sum(rweibull(N1[i], we1, we2)))
W2 <- c(W2, sum(rexp(N2[i], ex)))
″
S1 <- apply(matrix(W1, S, 20, T), 1, cumsum)
S2 <- apply(matrix(W2, S, 20, T), 1, cumsum) #
u <- c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140,
150)
PB <- c()
for(i in 1:length(u)) –
IN <- u[i] + matrix(rep((1:20) * (1/ex) * 2 * po1 * (1 + theta),
each = S), S, 20)
PBS <- apply((IN > t(S1) + t(S2)) * 1, 1, prod)
PB <- c(PB, 1 - mean(PBS))
″
return(cbind(u, PB))
″
″
```

19. Negative Binomial model with common component

19.1. Moments.

```
function(part = ȷAȷ, S = 1000, we1 = 0.5, we2
= 0.5625, nb1 = 1, nb2 = 4, ex = 1/1.125, rho = 0.25)
–
if(part == ȷAȷ) –
X1 <- rweibull(S, we1, we2)
EX1 <- mean(X1)
EX12 <- mean(X1ˆ2)
VX1 <- var(X1)
N1 <- rnbinom(S, nb1, 1/(1 + nb2))
EN1 <- mean(N1)
VN1 <- var(N1)
EW1 <- EX1 * EN1
VW1 <- EX1ˆ2 * VN1 + VX1 * EN1
X2 <- rexp(S, ex)
EX2 <- mean(X2)
EX22 <- mean(X2ˆ2)
VX2 <- var(X2)
N2 <- rnbinom(S, nb1, 1/(1 + nb2))
```

```
EN2 <- mean(N2)
VN2 <- var(N2)
EW2 <- EX2 * EN2
VW2 <- EX2^2 * VN2 + VX2 * EN2
resul <- matrix(c(EX1, EX12, EN1, VN1, EW1, VW1, EX2, EX22,
EN2, VN2, EW2, VW2), 6, 2)
return(resul)
”
if(part == ¡B¡) –
W1 <- c()
W2 <- c()
a0 <- (rho * nb1 * nb2 * (1 + nb2))/nb2^2
G1 <- rgamma(S, nb1 - a0, 1)
G2 <- rgamma(S, nb1 - a0, 1)
N0 <- 0
if(rho != 0) –
G0 <- rgamma(S, a0, 1)
N0 <- rpois(S, G0 * nb2)
”
N1 <- rpois(S, G1 * nb2) + N0
N2 <- rpois(S, G2 * nb2) + N0 # pet <- mean((N1 - mean(N1)) *
(N2 - mean(N2)))
for(i in 1:S) –
W1 <- c(W1, sum(rweibull(N1[i], we1, we2)))
W2 <- c(W2, sum(rexp(N2[i], ex)))
”
CN12 <- var(cbind(N1, N2))
CW12 <- var(cbind(W1, W2))
COW12 <- cor(cbind(W1, W2))
return(c(CN12[1, 2], CW12[1, 2], COW12[1, 2]))
”
”
```

19.2. Ruin.

```
function(S = 1000, we1 = 0.5, we2 = 0.5625, nb1 =
1, ex = 1/1.125, nb2 = 4, rho = 0.25, u = c(0, 30, 90), theta = 0.15)
–
N0 <- 0
a0 <- (rho * nb1 * nb2 * (1 + nb2))/nb2^2
if(rho != 0) –
G0 <- rgamma(20 * S, a0, 1)
N0 <- rpois(20 * S, G0 * nb2)
”
```

```
G1 <- rgamma(20 * S, nb1 - a0, 1)
G2 <- rgamma(20 * S, nb1 - a0, 1)
N1 <- rpois(20 * S, G1 * nb2) + N0
N2 <- rpois(20 * S, G2 * nb2) + N0
W1 <- c()
W2 <- c()
for(i in 1:(20 * S)) –
W1 <- c(W1, sum(rweibull(N1[i], we1, we2)))
W2 <- c(W2, sum(rexp(N2[i], ex)))
″
S1 <- apply(matrix(W1, S, 20, T), 1, cumsum)
S2 <- apply(matrix(W2, S, 20, T), 1, cumsum) #
u <- c(0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140,
150)
PB <- c()
for(i in 1:length(u)) –
IN <- u[i] + matrix(rep((1:20) * (1/ex) * 2 * nb1 * nb2 * (1 +
theta), each = S), S, 20)
PBS <- apply((IN > t(S1) + t(S2)) * 1, 1, prod)
PB <- c(PB, 1 - mean(PBS))
″
return(cbind(u, PB))
″
```

Introduction

The paper of Hu and Wu (1999) treats the notion of multivariate dependence between individuals and its eœect on the related stop-loss premiums. It focuses on the case of a portfolio of m life insurance policies, each having a positive face amount during a certain reference period. The ørst type of dependence considered is the dependence giving rise to the safest aggregate claims, and we saw in the previous report that this type of dependence is given by the lower Frchet bound or, equivalently, by mutually exclusive risks. Then, the notion of superadditive dependence ordering is applied to these results. This paper concludes with a numerical example, from which we reproduce the results.

We ørst present the two-point distributions of mutually exclusive risks presented by the authors. In the second section, we present some results relating the distributions introduced with the concept of stop-loss order. We then introduce the notion of superadditive dependence, and its link with stop-loss order as well as with the distributions presented in the ørst section. We ønally reproduce the results of the numerical example presented in Hu and Wu (1999), illustrating the eœects of dependence on stop-loss premiums.

Particular Types of Dependence

This section aims to present some distributions of mutually exclusive risks that will be analyzed in the next sections. The marginals considered are two-point distributions. We ørst present two distributions of nonexchangeable risks, and then one distribution of exchangeable risks. For the non-exchangeable risks, we present the distributions of the aggregate claims, and we verify that they really are lower Frchet bounds. In the case of the exchangeable risks, we only state the joint distribution.

Let $(X_1, X_2, ..., X_m)$ be a portfolio consisting of m risks $X_1, X_2, ..., X_m$ with X_i having a two-point distribution in 0 and $\alpha_i > 0$, that is

$$\Pr(X_i = 0) = p_i \qquad \text{and} \qquad \Pr(X_i = \alpha_i) = 1 - p_i = q_i, \qquad (28)$$

for $i = 1, 2, ..., m$.

The cumulative distribution function (cdf) of X_i, $i = 1, 2, ..., m$ is thus

$$F_i(x_i) = \Pr(X_i \leq x_i) = \begin{cases} 0 & x_i < 0 \\ p_i & 0 \leq x_i < \alpha_i \\ 1 & x_i \geq \alpha_i \end{cases}.$$

Let $\mathcal{H}(F_1, F_2, ..., F_m)$ denote the set of all m-dimensional random vectors with marginal distributions $F_1, F_2, ..., F_m$. For the present case, let

$$\mathcal{H}(q_1, ..., q_m; \alpha_1, ..., \alpha_m) \equiv \mathcal{H}_m$$

denote the class of random vectors $(X_1, X_2, ..., X_m)$, where each of the X_i's is distributed as (28). For a matter of convenience, we also assume that the risks $X_1, X_2, ..., X_m$ are classiøed so that the face amounts are in a nondecreasing order, that is

$$\alpha_1 < \alpha_2 < ... < \alpha_m.$$

The authors also precise that when $\alpha_1, \alpha_2, ..., \alpha_m$ have ties, all the coming results are valid under minor modiøcations.

20. Nonexchangeable Risks

We assume that random variables are exchangeable if any permutation of those has the same distribution. Nonexchangeable risks are then risks that are not exchangeable. Since we are working with distributions varying in the face amount (α_i) and also in the probability of this face amount (q_i), it is easy to see that risks having distribution (28) are nonexchangeable.

We now consider the safest dependence structure in the case where the marginal distributions of the risks X_i's are given by (28), given as known by the lower Frchet bound. We mentioned in the last report that the lower Frchet bound is equivalent to mutually exclusive risks in the same way that the upper Frchet bound is related to the concept of comonotonicity. We also saw that for the lower Frchet bound to be a proper cdf, some conditions have to be satisøed and these are stated in Theorem 4.1 of the second report. We thus consider two cases, one for each of the equivalent conditions required to have a proper cdf, that is $\sum_{i=1}^{m} p_i \geq m-1$ and $\sum_{i=1}^{m} p_i \leq 1$.

20.1. First Case.

We ørst consider the case where the condition $\sum_{i=1}^{m} p_i \geq m-1$ is satisøed. Let the distribution of the individual risks be given by

$$\begin{aligned}
\Pr\left(X_i=\alpha_i, X_j=\alpha_j\right) &= 0, \quad \forall i \neq j, & (29)\\
\Pr\left(\begin{array}{c} X_j=\alpha_j, X_i=0, \\ i=1,...,j-1,j+1,...,m \end{array}\right) &= 1-p_j, \quad \forall j=1,...,m, & (30)\\
\Pr\left(X_1=0, X_2=0,...,X_m=0\right) &= \sum_{i=1}^{m} p_i-(m-1). & (31)
\end{aligned}$$

We ørst derive the distribution of the aggregate claims, and then we verify that it is lower Frchet bound. We denote the distribution of the aggregate claims under this smallest dependence structure by S_* :

$$S_*=\sum_{i=1}^{m} X_i.$$

Since each of the risks is a two-point distribution and since the risks are mutually exclusive, then the range of the possible outcomes for S_* is limited and is given by $\{0, \alpha_1, \alpha_2, ..., \alpha_m\}$. We then have for the probability function of S_*:

$$\begin{aligned}
\Pr\left(S_*=0\right) &= \Pr\left(X_1=0, X_2=0,...,X_m=0\right)\\
&= \sum_{i=1}^{m} p_i-(m-1),
\end{aligned}$$

and

$$\Pr(S_* = \alpha_j) = \Pr\begin{pmatrix} X_1 = 0, ..., X_{j-1} = 0, X_j = \alpha_j, \\ X_{j+1} = 0, ..., X_m = 0 \end{pmatrix} = q_j,$$

for $j = 1, ..., m$.

It is now possible to ønd the cdf of S_*. We denote the distribution of S_* by H_* and we consider three cases. For $0 \leq s < \alpha_1$, $H_*(s)$ is just the probability of getting 0 for each risk:

$$\begin{aligned} H_*(s) &= \Pr(S_* = 0) \\ &= \Pr(X_1 = 0, X_2 = 0, ..., X_m = 0) \\ &= \sum_{i=1}^{m} p_i - (m-1). \end{aligned}$$

For $\alpha_{j-1} \leq s < \alpha_j$, $j = 2, ..., m$, $H_*(s)$ is the probability of getting an aggregate claim amount in $\{0, \alpha_1, ..., \alpha_{j-1}\}$:

$$\begin{aligned} H_*(s) &= \Pr(S_* \leq s) \\ &= \Pr(S_* \leq \alpha_{j-1}) \\ &= \Pr(S_* = 0) + \Pr(S_* = \alpha_1) + ... + \Pr(S_* = \alpha_{j-1}), \end{aligned}$$

which gives

$$\begin{aligned} H_*(s) &= \sum_{i=1}^{m} p_i - (m-1) + \sum_{i=1}^{j-1} q_i \\ &= \sum_{i=1}^{m} p_i - (m-1) + \sum_{i=1}^{j-1} (1 - p_i) \\ &= \sum_{i=1}^{m} p_i - (m-1) + (j-1) - \sum_{i=1}^{j-1} p_i \\ &= \sum_{i=j}^{m} p_i - (m-j). \end{aligned}$$

Finally, when $s \geq \alpha_m$, $H_*(s)$ is obviously a certain event and it then follows that

$$\begin{aligned} H_*(s) &= \Pr(S_* \leq s) \\ &= \Pr(S_* \leq \alpha_m) \\ &= 1. \end{aligned}$$

From the previous results, the distribution of S_* is then given by

$$H_*(s) = \begin{cases} \sum_{i=1}^{m} p_i - (m-1), & \text{for } 0 \leq s < \alpha_1, \\ \sum_{i=j}^{m} p_i - (m-j), & \text{for } \alpha_{j-1} \leq s < \alpha_j,\ j = 2, ..., m, \\ 1, & \text{for } s \geq \alpha_m. \end{cases} \tag{32}$$

As stated previously, H_* given by (32) is the lower Frchet bound of $\mathcal{H}_m$ when $\sum_{i=1}^{m} p_i \geq m-1$. As proved in the previous report, the lower Frchet bound of multivariate distributions of random vectors in $\mathcal{H}(F_1, F_2, ..., F_m)$ is

$$F_L(x_1, ..., x_m) = \max\left\{0, \sum_{i=1}^{m} F_i(x_i) - (m-1)\right\}.$$

We verify that we really have the same distribution using the lower Frchet bound. For the ørst case, $0 \leq s < \alpha_1$, $H_*(s)$ is equivalent to $F_L(0, ..., 0)$:

$$\begin{aligned} H_*(s) &= H_*(0) \\ &= F_L(0, ..., 0) \\ &= \max\left\{0, \sum_{i=1}^{m} p_i - (m-1)\right\}, \end{aligned}$$

and since $\sum_{i=1}^{m} p_i \geq m-1$, we can get rid of the maximum function and write

$$H_*(s) = \sum_{i=1}^{m} p_i - (m-1).$$

For the second case, $\alpha_{j-1} \leq s < \alpha_j$, $j = 2, ..., m$, $H_*(s)$ is equivalent to $F_L(\alpha_1, ..., \alpha_{j-1}, 0, ..., 0)$, which can be decomposed in numerous sums. The risks being mutually exclusive, the probability of getting the majority of these individual sums is equal to 0. In fact, since we cannot have more than one claim in the same period, we are left only with the terms having only one claim in the reference period, that is

$$\begin{aligned} F_L(\alpha_1, ..., \alpha_{j-1}, 0, ..., 0) &= \Pr(X_1 = 0, ..., X_m = 0) \\ &\quad + \Pr(X_1 = \alpha_1, X_2 = 0, ..., X_m = 0) \\ &\quad + \Pr(X_1 = 0, X_2 = \alpha_2, X_3 = 0, ..., X_m = 0) + ... \\ &\quad + \Pr\left(\begin{array}{c} X_1 = 0, ..., X_{j-2} = 0, X_{j-1} = \alpha_{j-1}, \\ X_j = 0, ..., X_m = 0 \end{array}\right) \\ &= H_*(s). \end{aligned}$$

Since $F_i(\alpha_i) = \Pr(X_i \leq \alpha_i) = 1$, $i = 1, ..., m$, we have

$$\begin{aligned} H_*(s) &= F_L(\alpha_1, ..., \alpha_{j-1}, 0, ..., 0) \\ &= \max\left\{0, \sum_{i=1}^{j-1} 1 + \sum_{i=j}^{m} p_i - (m-1)\right\}. \end{aligned}$$

Again, we can take oœ the maximum function because of the condition $\sum_{i=j}^{m} p_i \geq m-1$, and get

$$\begin{aligned} H_*(s) &= (j-1) + \sum_{i=j}^{m} p_i - (m-1) \\ &= \sum_{i=j}^{m} p_i - (m-j). \end{aligned}$$

For the last case when $s \geq \alpha_m$, $H_*(s)$ is equivalent to $F_L(\alpha_1, ..., \alpha_m)$, by applying the same reasoning as for the second case. Similarly, we then have

$$\begin{aligned} H_*(s) &= F_L(\alpha_1, ..., \alpha_m) \\ &= \max\left\{0, \sum_{i=1}^{m} 1 - (m-1)\right\} \\ &= \max\{0, m-(m-1)\} \\ &= 1. \end{aligned}$$

We then veriøed that (32) is a lower Frchet bound.

20.2. Second Case. We now consider the case where the condition $\sum_{i=1}^{m} p_i \leq 1$ is satisøed. Let the distribution of the individual risks be given by

$$\Pr(X_i = 0, X_j = 0) = 0, \qquad \forall i \neq j, \tag{33}$$

$$\Pr\begin{pmatrix} X_j = 0, X_i = \alpha_i, \\ i = 1, ..., j-1, j+1, ..., m \end{pmatrix} = p_j, \qquad \forall j = 1, ..., m, \tag{34}$$

$$\Pr(X_1 = \alpha_1, X_2 = \alpha_2, ..., X_m = \alpha_m) = 1 - \sum_{i=1}^{m} p_i. \tag{35}$$

Again, we derive the distribution of the aggregate claims and verify that it is lower Frchet bound. We still denote the distribution of the aggregate claims under this smallest dependence structure by S_* :

$$S_* = \sum_{i=1}^{m} X_i.$$

Since each of the risks is a two-point distribution and since the risks are mutually exclusive (now in terms of 0), then the range of the possible outcomes for S_* is limited and is given by $\{\alpha - \alpha_m, ..., \alpha - \alpha_1, \alpha\}$, where $\alpha = \sum_{i=1}^{m} \alpha_i$. Note that the risks are now mutually exclusive in the sense that we cannot have more than one risk that has no claim during a given reference period. We then have, for the probability function of S_*:

$$\begin{aligned}\Pr(S_* = \alpha) &= \Pr(X_1 = \alpha_1, X_2 = \alpha_2, ..., X_m = \alpha_m) \\ &= 1 - \sum_{i=1}^{m} p_i,\end{aligned}$$

and

$$\begin{aligned}\Pr(S_* = \alpha - \alpha_j) &= \Pr\begin{pmatrix} X_1 = \alpha_1, ..., X_{j-1} = \alpha_{j-1}, \\ X_j = 0, X_{j+1} = \alpha_{j+1}, ..., X_m = \alpha_m \end{pmatrix} \\ &= p_j,\end{aligned}$$

for $j = 1, ..., m$.

We denote the cdf of S_* by H_* and we consider three cases. For $0 \leq s < \alpha - \alpha_m$, $H_*(s)$ can be interpreted as the probability that at least two risks do not get a claim in a given reference period, since the smallest amount of aggregate claims when there is only one risk that do not get a claim is $\alpha - \alpha_m$ (i.e. when $X_m = 0$). The probability that at least two risks do not get a claim being 0, we then have

$$H_*(s) = 0.$$

For $\alpha - \alpha_j \leq s < \alpha - \alpha_{j-1}$, $j = 1, ..., m$, (with $\alpha_0 = 0$) $H_*(s)$ is the probability of getting an aggregate claim amount in $\{\alpha - \alpha_m, ..., \alpha - \alpha_j\}$:

$$\begin{aligned}H_*(s) &= \Pr(S_* \leq s) \\ &= \Pr(S_* \leq \alpha - \alpha_j) \\ &= \Pr(S_* = \alpha - \alpha_m) + ... + \Pr(S_* = \alpha - \alpha_j),\end{aligned}$$

which gives

$$\begin{aligned}H_*(s) &= p_m + p_{m-1} + ... + p_j \\ &= \sum_{i=j}^{m} p_i.\end{aligned}$$

Finally, when $s \geq \alpha$, $H_*(s)$ is obviously a certain event, and it then follows that

$$\begin{aligned} H_*(s) &= \Pr(S_* \leq s) \\ &= \Pr(S_* \leq \alpha) \\ &= 1. \end{aligned}$$

From the preceding results, the distribution of S_* is then given by

$$H_*(s) = \begin{cases} 0, & \text{for } 0 \leq s < \alpha - \alpha_m, \\ \sum_{i=j}^{m} p_i, & \text{for } \alpha - \alpha_j \leq s < \alpha - \alpha_{j-1},\ j = 1, ..., m, \\ 1, & \text{for } s \geq \alpha, \end{cases} \quad (36)$$

with $\alpha_0 = 0$.

As stated previously, H_* given by (36) is the lower Frchet bound of $\mathcal{H}_m$ when $\sum_{i=1}^{m} p_i \leq 1$. The lower Frchet bound of multivariate distributions of random vectors in $\mathcal{H}(F_1, F_2, ..., F_m)$ is

$$F_L(x_1, ..., x_m) = \max\left\{0, \sum_{i=1}^{m} F_i(x_i) - (m-1)\right\}.$$

Again, we verify that the distribution of S_* really is a lower Frchet bound. For the ørst case, $0 \leq s < \alpha - \alpha_m$, $H_*(s)$ is equivalent to $F_L(0, ..., 0)$ since

$$\begin{aligned} H_*(s) &= H_*(0) \\ &= F_L(0, ..., 0) \\ &= \max\left\{0, \sum_{i=1}^{m} p_i - (m-1)\right\}, \end{aligned}$$

and since $\sum_{i=1}^{m} p_i \leq 1$, the expression $\sum_{i=1}^{m} p_i - (m-1)$ is negative, so we can get rid of the maximum function and write

$$H_*(s) = 0.$$

For the second case, $\alpha - \alpha_j \leq s < \alpha - \alpha_{j-1}$, $j = 1, ..., m$, $H_*(s)$ is equivalent to a sum of F_L's, which in turn can be decomposed in numerous sums:

$$\begin{aligned} H_*(s) &= F_L(\alpha_1, ..., \alpha_{j-1}, 0, \alpha_{j+1}, ..., \alpha_m) \\ &\quad + F_L(\alpha_1, ..., \alpha_j, 0, \alpha_{j+2}, ..., \alpha_m) \\ &\quad + ... + F_L(\alpha_1, ..., \alpha_{m-1}, 0) \\ &= \Sigma_j. \end{aligned}$$

The risks being mutually exclusive, the majority of the individual sums of each F_L is equal to 0. In fact, since we cannot have more than two risks that do not have a claim in the same period, we are left only with one term for each F_L, the term where only one risk do not have any claim in the reference period, that is

$$\begin{aligned}\Sigma_j &= \Pr\left(X_1 = \alpha_1, ..., X_{j-1} = \alpha_{j-1}, X_j = 0, X_{j+1} = \alpha_{j+1}, ..., X_m = \alpha_m\right) \\ &\quad + \Pr\left(X_1 = \alpha_1, ..., X_j = \alpha_j, X_{j+1} = 0, X_{j+2} = \alpha_{j+2}, ..., X_m = \alpha_m\right) \\ &\quad + ... + \Pr\left(X_1 = \alpha_1, ..., X_{m-1} = \alpha_{m-1}, X_m = 0\right) \\ &= H_*\left(s\right).\end{aligned}$$

Since $F_i\left(\alpha_i\right) = \Pr\left(X_i \leq \alpha_i\right) = 1$, $i = 1, ..., m$, we then have

$$\begin{aligned}H_*\left(s\right) &= \Sigma_j \\ &= \sum_{k=j}^{m} \max\left\{0, \sum_{i=1, i\neq k}^{m} 1 + p_k - (m-1)\right\} \\ &= \sum_{k=j}^{m} \max\left\{0, (m-1) + p_k - (m-1)\right\} \\ &= \sum_{k=j}^{m} \max\left\{0, p_k\right\},\end{aligned}$$

and since $p_k \geq 0$, we can get rid of the maximum function and write

$$H_*\left(s\right) = \sum_{i=j}^{m} p_i.$$

For the last case when $s \geq \alpha$, $H_*\left(s\right)$ is equivalent to $F_L\left(\alpha_1, ..., \alpha_m\right)$ and we then have

$$\begin{aligned}H_*\left(s\right) &= F_L\left(\alpha_1, ..., \alpha_m\right) \\ &= \max\left\{0, \sum_{i=1}^{m} 1 - (m-1)\right\} \\ &= \max\left\{0, m - (m-1)\right\} \\ &= 1.\end{aligned}$$

We then veriøed that (36) is a lower Frchet bound.

21. Exchangeable Risks

We now present a distribution in the family of multivariate exchangeable Bernoulli distributions with marginal probability π. Some results in the next sections are related with this distribution. We now assume that $\alpha_1 = ... = \alpha_m = 1$ and $q_1 = ... = q_m = \pi$. Since the α_i's and the π_i's take the same values for $i = 1, ..., m$, then it is easy to see that they are exchangeable risks. Let the m-variate exchangeable Bernoulli risks $(X_1, X_2, ..., X_m)$ have the following distribution:

$$\Pr(X_1 = \delta_1, ..., X_m = \delta_m) = \begin{cases} (r+1-m\pi)\binom{m}{r}, & \text{if } \sum_{i=1}^{m}\delta_i = r, \\ (m\pi - r)\binom{m}{r+1}, & \text{if } \sum_{i=1}^{m}\delta_i = r+1, \\ 0, & \text{otherwise,} \end{cases} \tag{37}$$

where r is an integer such that $r \leq m\pi < r+1$, and $\delta_1, ..., \delta_m$ take values 0 and 1.

Since this function does not respect the condition of discrete pdf 's stating that the sum of the probabilities should be 1, then there is obviously a mistake in the formula. However, since we did not point out exactly what it is, then we do not develop much about that.

The Safest Aggregate Claims

In the second report, we treated the riskiest and the safest dependence structures among risks. We concluded that the riskiest dependence structure in the sense that it leads to the largest stop-loss premiums is given by the upper Frchet bound or, equivalently, by comonotonic risks. Similarly, the safest dependence structure, that is the dependence structure leading to the smallest stop-loss premiums is given by the lower Frchet bound under certain conditions to ensure that it is really a proper cdf. As we have seen, this is equivalent to the concept mutually exclusive risks. The results presented in this section rely the concept of stop-loss order with the distributions presented before.

The following theorem states that the aggregate claims S_* for the two cases of nonexchangeable risks presented in the previous section give rise to the minimal stop-loss premiums.

Theorem 21.1. Let $\mathbf{X} \in \mathcal{H}_m$ with the distribution given by (29)-(31) or (33)-(35). Then, for any $\mathbf{Y} \in \mathcal{H}_m$, we have

$$S \geq_{sl} S_*,$$

where $S_* = \sum_{i=1}^{m} X_i$ and $S = \sum_{i=1}^{m} Y_i$.

Proof. See Hu and Wu (1999). ∎

It is straight-forward to verify that this is true for $m = 1$. We then have $S_* = X_1$ and $S = Y_1$. For this case, note that both conditions for the lower Frchet bound to be a proper cdf are satisøed

$$m - 1 \leq \sum_{i=1}^{m} p_i \leq 1,$$

which implies that

$$0 \leq p_1 \leq 1.$$

Since the aggregate claims consist in fact in just one risk, we cannot consider any dependence relation between risks, and then the stop-loss premiums have to be the same for any retention level d, since X_1 and Y_1 have the same marginal distribution. We then have

$$\begin{aligned} E\left(S_* - d\right)_+ &= E\left(X_1 - d\right)_+ \\ &= q_1\left(\alpha_1 - d\right)_+ + p_1\left(0 - d\right)_+ \\ &= q_1\left(\alpha_1 - d\right)_+, \end{aligned}$$

and

$$\begin{aligned} E\left(S - d\right)_+ &= E\left(Y_1 - d\right)_+ \\ &= q_1\left(\alpha_1 - d\right)_+ + p_1\left(0 - d\right)_+ \\ &= q_1\left(\alpha_1 - d\right)_+. \end{aligned}$$

We can then say that $S \geq_{sl} S_*$.

We now present a result similar to that of Theorem 2.1, but now for the multivariate Bernoulli distribution (37) of exchangeable risks.

Theorem 21.2. Let $\mathbf{X}$ and $\mathbf{Y}$ be two elements of $\mathcal{H}\left(\pi, ..., \pi; 1, ..., 1\right)$, and let the distribution of $\mathbf{X}$ be given by (37). Then $S \geq_{sl} S_*$, where $S_* = \sum_{i=1}^{m} X_i$ and $S = \sum_{i=1}^{m} Y_i$.

Proof. See Hu and Wu (1999). ∎

This means that the distribution for exchangeable risks presented in Section 1.2 is the safest distribution among those having the same marginal distributions. In other words, any other ȷcombinationȷ in terms of joint distribution is riskier than that we presented.

Superadditive Dependence Ordering and Stop-loss Order

In the second report, we presented the notion of correlation ordering between bivariate random vectors and we related this concept to that of stop-loss order. We now present the relations between the superadditive dependence ordering and the stop-loss ordering for multivariate risks, and then we specify these general results for the distributions of the ørst section.

We ørst deøne the concordance ordering. We should notice that the correlation ordering is in fact the concordance ordering for the bivariate case. However, although the correlation order between some risks implies the stop-loss order between their sums, the same implication does not hold for multivariate risks.

Definition 21.1. Let $\mathbf{X}$ and $\mathbf{Y}$ belong to $\mathcal{H}(F_1, F_2, ..., F_m)$. $\mathbf{Y}$ is said to be more concordant than $\mathbf{X}$ (denoted by $\mathbf{Y} \geq_c \mathbf{X}$) if, for any $\mathbf{x} = (x_1, ..., x_m)$,

$$\Pr(Y_i \leq x_i, i = 1, ..., m) \geq \Pr(X_i \leq x_i, i = 1, ..., m),$$

and

$$\Pr(Y_i > x_i, i = 1, ..., m) \geq \Pr(X_i > x_i, i = 1, ..., m).$$

The following example (developed from Hu and Wu (1999)) illustrates that concordance ordering does not necessarily implies stop-loss ordering.

Example 21.1. Let $\mathbf{X}$ and $\mathbf{Y}$ be four-dimensional risk vectors with support on $\{0,1\}^4$. Let f_1, f_2 be their probability mass functions and let $d = f_2 - f_1$. Let ε be a small positive constant and let d be deøned by $d(i_1, i_2, i_3, i_4) = \varepsilon$ if there are an even number of zeros among i_1, i_2, i_3, i_4, and $d(i_1, i_2, i_3, i_4) = -\varepsilon$ if there are an odd number of zeros among i_1, i_2, i_3, i_4. Assume that f_1, f_2 are nonzero where necessary in order for d to be well-deøned. Denote $p_{ijkl} = f_1(i, j, k, l)$ for all i, j, k, l.

Then,

$$E\left(\sum_{u=1}^{4} X_u - 2\right)_+ = 2p_{1111} + \sum_{i+j+k+l=3} p_{ijkl},$$

and

$$\begin{aligned}
E\left(\sum_{u=1}^{4} Y_u - 2\right)_+ &= 2\left(p_{1111}+\varepsilon\right) + \sum_{i+j+k+l=3}\left(p_{ijkl}-\varepsilon\right) \\
&= 2p_{1111} + 2\varepsilon + \sum_{i+j+k+l=3} p_{ijkl} - \binom{4}{3}\varepsilon \\
&= 2p_{1111} + 2\varepsilon + \sum_{i+j+k+l=3} p_{ijkl} - 4\varepsilon \\
&= 2\left(p_{1111}-\varepsilon\right) + \sum_{i+j+k+l=3} p_{ijkl},
\end{aligned}$$

where i, j, k, l take the values 0 or 1. We then necessarily know that X is not smaller in stop-loss order than $\mathbf{Y}$, since the stop-loss premium with a retention level $d = 2$ is bigger for $\mathbf{X}$. On the other hand, we can see with the following relations that $\mathbf{X}$ is smaller in concordance order than $\mathbf{Y}$, that is $\mathbf{X} \leq_c \mathbf{Y}$. To see it is really true, we verify the two conditions stated in the deønition of the concordance order for each possible case $\mathbf{x} = (x_1, x_2, x_3, x_4)$.

For $\mathbf{x} = (0,0,0,0)$, we have

$$\Pr\left(\mathbf{X} \leq \mathbf{x}\right) = p_{0000} \leq p_{0000} + \varepsilon = \Pr\left(\mathbf{X} \leq \mathbf{y}\right),$$

for the ørst condition, and

$$\Pr\left(\mathbf{X} > \mathbf{x}\right) = p_{1111} \leq p_{1111} + \varepsilon = \Pr\left(\mathbf{X} > \mathbf{y}\right),$$

for the second one. For $\mathbf{x} = (1,0,0,0)$, we get

$$\begin{aligned}
\Pr\left(\mathbf{X} \leq \mathbf{x}\right) &= p_{1000} + p_{0000} \\
&\leq \left(p_{1000} - \varepsilon\right) + \left(p_{0000} + \varepsilon\right) \\
&= p_{1000} + p_{0000} \\
&= \Pr\left(\mathbf{X} \leq \mathbf{y}\right),
\end{aligned}$$

and also

$$\Pr\left(\mathbf{X} > \mathbf{x}\right) = 0 \leq 0 = \Pr\left(\mathbf{X} > \mathbf{y}\right),$$

since the probability of taking a value strictly bigger than 1 is 0 for any variable. We should note that every case symmetric to this one, that is

$\mathbf{x}=(0,1,0,0)$, $\mathbf{x}=(0,0,1,0)$, and $\mathbf{x}=(0,0,0,1)$, leads to the same result. For the third case, i.e. $\mathbf{x}=(1,1,0,0)$, we have

$$\begin{aligned}
\Pr(\mathbf{X}\leq\mathbf{x}) &= p_{1100}+p_{1000}+p_{0100}+p_{0000}\\
&\leq (p_{1100}+\varepsilon)+(p_{1000}-\varepsilon)+(p_{0100}-\varepsilon)+(p_{0000}+\varepsilon)\\
&= p_{1100}+p_{1000}+p_{0100}+p_{0000}\\
&= \Pr(\mathbf{Y}\leq\mathbf{x}),
\end{aligned}$$

and for the same reasons as before, we get for the second condition

$$\Pr(\mathbf{X}>\mathbf{x})=0\leq 0=\Pr(\mathbf{X}>\mathbf{y}).$$

Again, the other 5 symmetric cases lead to the same result. The fourth case, $\mathbf{x}=(1,1,1,0)$, implies that

$$\begin{aligned}
\Pr(\mathbf{X}\leq\mathbf{x}) &= p_{1110}+p_{1100}+p_{1010}+p_{0110}\\
&\quad +p_{1000}+p_{0100}+p_{0010}+p_{0000}\\
&\leq (p_{1110}-\varepsilon)+(p_{1100}+\varepsilon)+(p_{1010}+\varepsilon)\\
&\quad +(p_{0110}+\varepsilon)+(p_{1000}-\varepsilon)+(p_{0100}-\varepsilon)\\
&\quad +(p_{0010}-\varepsilon)+(p_{0000}+\varepsilon)\\
&= p_{1110}+p_{1100}+p_{1010}+p_{0110}\\
&\quad +p_{1000}+p_{0100}+p_{0010}+p_{0000}\\
&= \Pr(\mathbf{Y}\leq\mathbf{x}),
\end{aligned}$$

and again the other condition brings that

$$\Pr(\mathbf{X}>\mathbf{x})=0\leq 0=\Pr(\mathbf{X}>\mathbf{y}).$$

There are three other cases that are symmetric to this one, with the same results. Finally, the last case $\mathbf{x}=(1,1,1,0)$ gives

$$\begin{aligned}
\Pr(\mathbf{X}\leq\mathbf{x}) &= p_{1111}+p_{1110}+p_{1101}+p_{1011}+p_{0111}\\
&\quad +p_{1100}+p_{1010}+p_{1001}+p_{0110}+p_{0101}+p_{0011}\\
&\quad +p_{1000}+p_{0100}+p_{0010}+p_{0001}+p_{0000}\\
&\leq (p_{1111}+\varepsilon)+(p_{1110}-\varepsilon)+(p_{1101}-\varepsilon)+(p_{1011}-\varepsilon)\\
&\quad +(p_{0111}-\varepsilon)+(p_{1100}+\varepsilon)+(p_{1010}+\varepsilon)+(p_{1001}+\varepsilon)\\
&\quad +(p_{0110}+\varepsilon)+(p_{0101}+\varepsilon)+(p_{0011}+\varepsilon)+(p_{1000}-\varepsilon)\\
&\quad +(p_{0100}-\varepsilon)+(p_{0010}-\varepsilon)+(p_{0001}-\varepsilon)+(p_{0000}+\varepsilon)\\
&= p_{1111}+p_{1110}+p_{1101}+p_{1011}+p_{0111}\\
&\quad +p_{1100}+p_{1010}+p_{1001}+p_{0110}+p_{0101}+p_{0011}\\
&\quad +p_{1000}+p_{0100}+p_{0010}+p_{0001}+p_{0000}\\
&= \Pr(\mathbf{Y}\leq\mathbf{x}).
\end{aligned}$$

Obviously, the second condition for this last case is as the previous ones:

$$\Pr(\mathbf{X} > \mathbf{x}) = 0 \leq 0 = \Pr(\mathbf{X} > \mathbf{y}).$$

Since all the possible combinations for $\mathbf{x}$ always respect both conditions of the concordance order, then we can aŒrm that $\mathbf{X}$ is smaller in concordance order than $\mathbf{Y}$. It then follows that concordance order does not imply stop-loss order.

Before introducing the main result, we should introduce the deøn-ition of superadditive functions and also the concept of superadditive dependence ordering.

Definition 21.2. A real-valued function ϕ deøned on $\Re^m$ is said to be superadditive if

$$\phi(\mathbf{x} \vee \mathbf{y}) + \phi(\mathbf{x} \wedge \mathbf{y}) \geq \phi(\mathbf{x}) + \phi(\mathbf{y}), \qquad \forall \mathbf{x}, \mathbf{y} \in \Re^m.$$

Here, $\vee$ and $\wedge$ denote, respectively, the componentwise maximum and the componentwise minimum. Also, a function $\phi(x_1, ..., x_m)$ is superadditive if, and only if, $\phi(.., x_i, ..., x_j, ...)$ is superadditive in (x_i, x_j) for any $i \neq j$ with the other variables øxed. If ϕ has continuous second partial derivatives, then the notion of superadditivity is equivalent to

$$\frac{\partial^2 \phi}{\partial x_i \partial x_j} \geq 0,$$

for all $i \neq j$. We can now introduce the superadditive dependence ordering.

Definition 21.3. Let $\mathbf{X}$ and $\mathbf{Y}$ belong to $\mathcal{H}(F_1, F_2, ..., F_m)$. $\mathbf{Y}$ is said to be more superadditively dependent than $\mathbf{X}$ (denoted by $\mathbf{Y} \geq_{sa} \mathbf{X}$) if $E(\phi(\mathbf{Y})) \geq E(\phi(\mathbf{X}))$ for all superadditive functions ϕ for which the expectations exist.

Now that we know these deønitions, we can present the following result.

Theorem 21.3. Let m-dimensional random vectors $\mathbf{X}$ and $\mathbf{Y}$ be deøned as

$$\mathbf{X} = (g_1(U_1, V_1, W), g_2(U_2, V_2, W), ..., g_m(U_m, V_m, W)),$$

and

$$\mathbf{Y} = (g_1(U_1, V_1, W), g_2(U_2, V_1, W), ..., g_m(U_m, V_1, W)),$$

where $\{U_i, i = 1, ..., m\}$ and $\{V_i, i = 1, ..., m\}$ are, respectively, identically distributed, and $\{W, U_i, V_i, i = 1, ..., m\}$ are independent. If $g_i(u, v, w)$,

$i = 1, ..., m$, are all increasing or all decreasing in v for every (u, w), then $\mathbf{Y} \geq_{sa} \mathbf{X}$.

Proof. See Hu and Wu (1999). ∎

With this last result, we can now present a special case making a link between the superadditive dependence ordering and the stop-loss order. The following theorem follows from the fact that $U \geq_{sl} V$ if, and only if, $E(h(U)) \geq E(h(V))$ for all increasing convex functions $h(x)$. Since the function $\phi(\mathbf{x}) = (x_1 + x_2 + ... + x_m)_+$ is superadditive (it is easily seen from the second partial derivatives condition), the next result directly follows.

Theorem 21.4. Let $\mathbf{X}, \mathbf{Y} \in \mathcal{H}(F_1, F_2, ..., F_m)$. If $\mathbf{Y} \geq_{sa} \mathbf{X}$, then $\sum_{i=1}^{m} Y_i \geq_{sl} \sum_{i=1}^{m} X_i$.

Proof. See Hu and Wu (1999). ∎

From Theorem 2.3, we can strengthen Theorem 2.1, which provided a lower bound in stop-loss order for the marginal distributions given by (28). If we consider a special case of Theorem 2.3, that is the case where $g_i(u, v, w)$ depends only on v, then we have

$$\begin{aligned} \mathbf{X} &= (h_1(V_1), h_2(V_2), ..., h_m(V_m)), \\ \mathbf{X}^* &= (h_1(V_1), h_2(V_1), ..., h_m(V_1)). \end{aligned}$$

We know from Theorem 2.3 that if $h_i(v), i = 1, ..., m$, are all increasing or all decreasing in v, then $\mathbf{X}^* \geq_{sa} \mathbf{X}$. If $\mathbf{X} \in \mathcal{H}(F_1, F_2, ..., F_m)$, then we know that every function is non-decreasing, and then it is easy to see that $\mathbf{X}^*$ represents the comonotonic version of the risks (upper Frchet bound distribution). Since the stop-loss premium is a superadditive function, then $\sum_{i=1}^{m} X_i^* \geq_{sl} \sum_{i=1}^{m} X_i$. This then introduces an upper bound for stop-loss ordering (and thus also for our previous distributions). Stop-loss orders are thus bounded both over and below by comonotonic risks and mutually exclusive risks, respectively. This is what Theorem 2.4 tells us.

We now present a last result, relating the superadditive ordering and stop-loss ordering for mutually exclusive risks.

Theorem 21.5. Let $\mathbf{X}$ and $\mathbf{Y}$ be two elements of $\mathcal{H}_m$ with $\sum_{i=1}^{m} p_i \geq m-1$ or $\sum_{i=1}^{m} p_i \leq 1$, and let the distribution of $\mathbf{X}$ be given by (29)-(31) or (33)-(35). Then, $\mathbf{Y} \geq_{sa} \mathbf{X}$.

Proof. See Hu and Wu (1999). ∎

This theorem is saying that the distributions of mutually exclusive risks presented previously are smaller in superadditive dependence than any other distribution with the same marginals, but with a diœerent joint distribution. The least superadditive dependence is thus given by the lower Frchet bound, as for the stop-loss order. For the case where the marginal distributions of the risks are given by (28), the least superadditively dependent distributions (and also the smaller in stop-loss order) are the distributions (29)-(31) or (33)-(35), depending on which condition is satisøed.

A Numerical Example

We now illustrate the effect of introducing negative and positive dependence between risks in an insurance portfolio by a numerical example. In fact, we reproduce the results obtained in Hu and Wu (1999). We use the life insurance portfolio consisting of 31 risks presented in Table 1. We should notice that $\sum_{i=1}^{31} q_i \leq 1$. Each risk has a two-point distribution with a mass at 0, that is each risk either produces no claim or a fixed positive claim amount (amount at risk) during a reference period. The claim probability is the probability that the risk produces a claim during the reference period. We should notice that the risks are assigned by column, i.e. the first column (amount at risk 1) has only two risks, called X_1 and X_2, the second column (amount at risk 2) has 8 risks, called $X_3, ..., X_{10}$, and so forth.

Claim probability	Amount at risk				
	1	2	3	4	5
0.01	2	3	1	2	-
0.02	-	1	2	2	1
0.03	-	2	4	2	2
0.04	-	2	2	2	1

Table 1. Number of policies with given amount at risk and claim probability

The joint distribution of most negatively and positively dependent risks are the lower and upper Frchet bounds, respectively. Let S_* and S^* denote the aggregate claims of the portfolio consisting of 31 mutually exclusive risks and comonotonic risks respectively. We develop the distribution for both cases.

We first consider the distribution for mutually exclusive risks. Since $\sum_{i=1}^{31} q_i \leq 1$, then we are in case 2, i.e. in the same case as Section 2.2. Extending the theory from that section to this particular example, we get

$$\begin{aligned}
\Pr(S_* = 0) &= \Pr(X_1 = 0, ..., X_{31} = 0) \\
&= 1 - \sum_{i=1}^{31} q_i \\
&= 1 - (8(0.01) + 6(0.02) + 10(0.03) + 7(0.04)) \\
&= 0.22.
\end{aligned}$$

We also have

$$\begin{aligned}
\Pr(S_* = 1) &= \Pr(X_1 = 1, X_2 = 0, ..., X_{31} = 0) \\
&\quad + \Pr(X_1 = 0, X_2 = 1, X_3 = 0, ..., X_{31} = 0) \\
&= q_1 + q_2 \\
&= 0.01 + 0.01 \\
&= 0.02,
\end{aligned}$$

$$\begin{aligned}
\Pr(S_* = 2) &= \sum_{i=3}^{10} q_i \\
&= 3(0.01) + 1(0.02) + 2(0.03) + 2(0.04) \\
&= 0.19,
\end{aligned}$$

$$\begin{aligned}
\Pr(S_* = 3) &= \sum_{i=11}^{19} q_i \\
&= 1(0.01) + 2(0.02) + 4(0.03) + 2(0.04) \\
&= 0.25,
\end{aligned}$$

$$\begin{aligned}
\Pr(S_* = 4) &= \sum_{i=20}^{27} q_i \\
&= 2(0.01) + 2(0.02) + 2(0.03) + 2(0.04) \\
&= 0.20,
\end{aligned}$$

$$\begin{aligned}
\Pr(S_* = 5) &= \sum_{i=28}^{31} q_i \\
&= 0(0.01) + 1(0.02) + 2(0.03) + 1(0.04) \\
&= 0.12.
\end{aligned}$$

It then follows that the distribution of S_* is given by

$$\Pr(S_* = k) = \begin{cases} 0.22, & \text{for } k = 0, \\ 0.02, & \text{for } k = 1, \\ 0.19, & \text{for } k = 2, \\ 0.25, & \text{for } k = 3, \\ 0.20, & \text{for } k = 4, \\ 0.12, & \text{for } k = 5. \end{cases}$$

We now consider the distribution for comonotonic risks, i.e. risks given by the upper Frchet bound:

$$F_U(x_1, ..., x_m) = \min(F_1(x_1), ..., F_m(x_m)).$$

The easiest way to get what we want is ørst, to ønd the cdf of these comonotonic risks, and then deduce the probability function. The way it works is that when a claim having probability 0.04 occurs, then all claims with probability 0.04 occur, since they are comonotonic and we suppose they all depend on the same variable. Also, if a claim having probability 0.03 occurs, then all other claims with probability 0.03 occur, but also all claims with probability 0.04, since these events are more likely to happen, and so on. There are then only a few total amount of claims possible, that is $\{0, 23, 57, 78, 97\}$. We ørst consider the probability that the aggregate claims are $0 \leq s < 23$, in other words the probability that none of the claims occurs:

$$\begin{aligned} \Pr(S^* \leq 0) &= F_U(0, ..., 0) \\ &= \min(0.96, 0.97, 0.98, 0.99) \\ &= 0.96, \end{aligned}$$

since among the whole bunch of variables, there are only four diœerent probabilities. For $23 \leq s < 57$, the event $S^* \leq s$ is realized if some or all of the claims with probability 0.04 arise. Since the risks are comonotonic, either the event with probability 0.04 occurs, and every risk with this probability has a claim, or none of them arises. It is thus impossible, in fact, to obtain a realization of S^* between 23 and 57. The amount will be either 23 or 57. This is traduced by an event with probability

$$\begin{aligned} \Pr(S^* \leq s) &= \min(1, 0.97, 0.98, 0.99) \\ &= 0.97, \end{aligned}$$

since the marginals are 1 for the risks occurring, and are $1 - q_i = p_i$ for the risks that do not arise. For $57 \leq s < 78$, the event $S^* \leq s$ is realized if all of the claims with probability 0.03 arise, and this implies of course that all the claims with probability 0.04 also occur. The possible realization for S^* is then 57, and this is traduced by an event with probability

$$\begin{aligned}\Pr(S^* \leq s) &= \min(1, 1, 0.98, 0.99) \\ &= 0.98.\end{aligned}$$

For $78 \leq s < 97$, the event $S^* \leq s$ is realized if all of the claims with probability 0.04, 0.03 and 0.02 arise. The amount of claim 78 is again the only possible value for S^* in this case. This is traduced by an event with probability

$$\begin{aligned}\Pr(S^* \leq s) &= \min(1, 1, 1, 0.99) \\ &= 0.99.\end{aligned}$$

Finally, for $s \geq 97$, the event $S^* \leq s$ is realized if every single claim occurs. This is traduced by an event with probability

$$\begin{aligned}\Pr(S^* \leq 97) &= \min(1, 1, 1, 1) \\ &= 1.\end{aligned}$$

It then follows that the cdf of S^* is given by

$$\Pr(S^* \leq s) = \begin{cases} 0.96, & \text{for } 0 \leq s < 23, \\ 0.97, & \text{for } 23 \leq s < 57, \\ 0.98, & \text{for } 57 \leq s < 78, \\ 0.99, & \text{for } 78 \leq s < 97, \\ 1, & \text{for } s \geq 97, \end{cases}$$

and thus, the probability distribution of S^* is

$$\Pr(S^* = k) = \begin{cases} 0.96, & \text{for } k = 0, \\ 0.01, & \text{for } k = 23, 57, 78, 97. \end{cases}$$

From the distributions we found, it is now easy to calculate the stop-loss premiums for many retention levels d. For instance, we have for a retention level of 2 for mutually exclusive risks

$$\begin{aligned}E(S_* - 2)_+ &= (0-2)_+ 0.22 + (1-2)_+ 0.02 + (2-2)_+ 0.19 \\ &\quad + (3-2)_+ 0.25 + (4-2)_+ 0.20 + (5-2)_+ 0.12 \\ &= 0.25 + 2(0.20) + 3(0.12) \\ &= 1.01,\end{aligned}$$

and for a retention level of 10, we get for comonotonic risks

$$\begin{aligned}E(S^* - 10)_+ &= (0-10)_+ 0.96 + (23-10)_+ 0.01 + (57-10)_+ 0.01 \\ &\quad + (78-10)_+ 0.01 + (97-10)_+ 0.01 \\ &= 13(0.01) + 47(0.01) + 68(0.01) + 87(0.01) \\ &= 2.15.\end{aligned}$$

The results for retention levels $d = 0, ..., 11$ for both extremal dependence structures are presented in Table 2. Results for the independence case are also presented. Since those are rather tedious to do by hand, we implemented a function in SPLUS using the Fast Fourier Transform (FFT) to convolve the risks and then get the stop-loss premiums for diœerent retention levels. These results are also included in Table 2 and the SPLUS function is presented in the appendix. It is easily noticed that the stop-loss premiums increase in going from the most negative dependence structure to the most positive dependence structure. Also, this diœerence is more important as the retention level d gets farther from 0.

Retention d	Mutually Exclusive	Independent	Comonotonic
0	2.55	2.55	2.55
1	1.77	2.00	2.51
2	1.01	1.47	2.47
3	0.44	1.02	2.43
4	0.12	0.69	2.39
5	0	0.46	2.35
6	0	0.31	2.31
7	0	0.20	2.27
8	0	0.12	2.23
9	0	0.08	2.19
10	0	0.05	2.15
11	0	0.03	2.11

Table 2. Stop-loss premiums for the portfolio in Table1.

Appendix

22. Stop-loss premium function (independence)

```
convol <- function(d = 0, data = tab1, repartition = repar)–
# Mylne Bdard
# 20/02/2002
# This function returns the stop-loss premium for a retention level
d and
# discrete distributions given by the matrix data (one distribution
by row),
# and the number of individuals for each distribution given by
repartition.
    long <- 2^12
    zeros <- matrix(0, nrow(data), long - ncol(data))
    datal <- cbind(data, zeros)
    Mx <- apply(datal, 1, œt)
    Mx2 <- Mx^matrix(repartition, long, nrow(data), byrow = T)
    Ms <- apply(Mx2, 1, prod)
    fs <- Re(œt(Ms, inverse = T))[1:long]
    fs <- (fs >= 0) * fs
    fs <- fs/sum(fs)
    amount <- (0:(long - 1)) - d
    ES <- sum((amount >= 0) * amount * fs)
    return(ES)
”
```

Introduction

The paper of Cossette et al. (2000) treats of the impact of dependence among multiple claims in a single event (loss). In casualty insurance, policies often involve correlated random variables. For instance, an insurance company issuing a ȷstandardȷ travel insurance contract has to consider a certain amount of correlation, either positively or negatively, between claims under the diœerent coverages oœered. One can observe that medical costs and disablement payments are positively associated, while some claims in a loss can even be mutually exclusive, as disablement and death payments. The authors aim to derive bounds on the cumulative distribution function (cdf) of the aggregate claim S in order to quantify the impact of correlation among the multiple claims related to a single event. The proposed methods allow developing bounds when the marginal distributions of the claim amounts are speciøed or when only partial information is available (e.g. ørst moments of the distributions).

In the ørst section, we shall set some notation, to make the further developments clearer. In the second and third sections, we present the results for the bounds of a single loss when the marginal are known, and when only the ørst moments are known, respectively. With these results, we can then develop the bounds for the aggregate claims, that is the total amount of loss, depending on the amount of information we have about the distributions. We ønally present a numerical illustration (reproduced from Cossette et al. (2000)), which leads to the development of the bounds for given distributions when full or partial information is available.

Notation

For a given insurance portfolio, deøne S to be the aggregate claim amount during a øxed period of time

$$S = \sum_{i=1}^{N} X_i, \tag{38}$$

where N represents the number of accidents and X_i, $i = 1, 2, ...$, the ith loss amount. We suppose that the random variable N is independent of the X_i's. Each $X_i, i = 1, 2, ...$, is the aggregation of the claims under a øxed number of diœerent coverages. Since each loss consists in a øxed number m of claims (one from each of the individual coverages), then X_i can be decomposed as

$$X_i = X_i^{(1)} + X_i^{(2)} + ... + X_i^{(m)}, \tag{39}$$

for $i = 1, 2, ...$, where $X_i^{(j)}$ is the jth claim on the ith loss, and $X_i^{(j)} = 0$ means that there is no claim. Although the random variables $X_i^{(j)}$, $j = 1, 2, ..., m$, are clearly dependent for a øxed i since they result from a same event, their correlation structure is generally not entirely known.

We denote the cdf of $X_i^{(j)}$ as $F_{X^{(j)}}$, $j = 1, 2, ..., m$, $i = 1, 2, ...$ The random vectors $\left(X_i^{(1)}, X_i^{(2)}, ..., X_i^{(m)}\right)$, $i = 1, 2, ...$, are formed of the claim amounts arising from each coverage in a single loss, and are assumed independent and identically distributed (iid) with unknown common joint cdf $F_{X^{(1)},X^{(2)},...,X^{(m)}}$. This assumption of homogeneity is realist since insurers generally tend to group similar risks in a same portfolio. Finally, the X_i's deøned in (39) have the common cdf F_X.

A simple way to view this model is to notice that it consists in the classical risk model, but where each one of the N independent random losses is decomposed as a øxed sum of m dependent components.

The following sections present methods from Cossette et al. (2000) to derive bounds of the cdf of S, F_S. These methods allow to handle two kinds of situations: the case where the marginal distributions $F_{X^{(j)}}$, $j = 1, 2, ..., m$, are speciøed, and the case where the marginal distributions

$F_{X(j)}$, $j = 1, 2, ..., m$, are unknown but their ørst few moments are given. Note that some variations of the method with the moments exist, depending on the number of moments available and on the presence of an upper bound for this distribution. The authors mention that the bounds on S derived in their paper are the best-possible bounds in the classical sense of stochastic dominance.

Stochastic Bounds on a Single Loss: Known Marginals

We now present a method to obtain bounds on F_X when the marginal distributions are speciøed, in order to be able to further derive bounds on F_S. We ørst consider the case where the type of dependence existing among the random variables is unknown, and we then present the case where we have some information about this correlation.

23. Unknown Dependence Structure

When no assumption is made in regard to the correlation structure between $X_i^{(1)}, X_i^{(2)}, ..., X_i^{(m)}$, $i = 1, 2, ...$, then for all $s \geq 0$, there exist $F_{\min}$ and $F_{\max}$ such that

$$F_{\min}(s) \leq F_X(s) \leq F_{\max}(s), \tag{40}$$

where

$$F_{\min}(s) = \sup_{(x_1, x_2, ..., x_m) \in \Sigma(s)} \max\left\{ \sum_{j=1}^{m} \Pr\left(X_i^{(j)} < x_j \right) - (m-1), 0 \right\},$$

$$F_{\max}(s) = \inf_{(x_1, x_2, ..., x_m) \in \Sigma(s)} \min\left\{ \sum_{j=1}^{m} F_{X^{(j)}}(x_j), 1 \right\},$$

and

$$\Sigma(s) = \left\{ (x_1, x_2, ..., x_m) \in \mathbb{R}^m \left| x_1 + x_2 + ... + x_m = s \right. \right\}.$$

As stated previously, these bounds are the best-possible bounds on F_X in the full information case. From their general form, we should notice the similarity between these bounds and the Frchet bounds. The provenance of this result can be found in Denuit et al. (1999).

For some distributions as the uniform, the normal, the Cauchy and the exponential family, it is possible to ønd explicit expressions for $F_{\min}$ and $F_{\max}$. Otherwise, they have to be approximated numerically. When we have to ønd these bounds, the case where $m = 2$ is the most simple. After having expressed x_2 as $s - x_1$, we can ønd the supremum (inømum) for each s by maximizing (minimizing) the given functions. For $F_{X^{(j)}}$ being the cdf of $X_i^{(j)}$ (as speciøed in the ørst section), this is written as

$$\begin{aligned} F_{\min}(s) &= \sup_{(x_1,x_2)\in\Sigma(s)} \max\left\{F^-_{X^{(1)}}(x_1) + F^-_{X^{(2)}}(x_2) - 1, 0\right\} \\ &= \sup_{x_1\in\mathbb{R}} \max\left\{F^-_{X^{(1)}}(x_1) + F^-_{X^{(2)}}(s - x_1) - 1, 0\right\}, \end{aligned}$$

and

$$\begin{aligned} F_{\max}(s) &= \inf_{(x_1,x_2)\in\Sigma(s)} \min\left\{F_{X^{(1)}}(x_1) + F_{X^{(2)}}(x_2), 1\right\} \\ &= \inf_{x_1\in\mathbb{R}} \min\left\{F_{X^{(1)}}(x_1) + F_{X^{(2)}}(s - x_1), 1\right\}. \end{aligned}$$

Note that $F^-_{X^{(j)}}(s)$ is the left limit of the distribution, and is equivalent to $\Pr\left(X_i^{(j)} < s\right)$. When $m \geq 3$, we have to use a recursive method in order to ønd these bounds. For instance, when $m = 3$, we ørst ønd

$$F_{\max(2)}(s - x_1) = \inf_{x_2+x_3=s-x_1} \min\left\{F_{X^{(2)}}(x_2) + F_{X^{(3)}}(x_3), 1\right\},$$

and then we can use this to compute

$$\begin{aligned} F_{\max}(s) &= \inf_{(x_1,x_2,x_3)\in\Sigma(s)} \min\left\{F_{X^{(1)}}(x_1) + F_{X^{(2)}}(x_2) + F_{X^{(3)}}(x_3), 1\right\} \\ &= \inf_{x_1\in\mathbb{R}} \min\left\{F_{X^{(1)}}(x_1) + F_{\max(2)}(s - x_1), 1\right\}. \end{aligned}$$

We proceed in a similar way for the lower bound, as well as for an other value of m.

In the case where the $X_i^{(j)}$'s have an identical distribution (Uniform, Exponential or Pareto), we get fairly simple expressions for the bounds $F_{\min}$ and $F_{\max}$. We now develop the bounds for each of these three cases. In order to simplify the expressions, we let $s_i = x_1 + ... + x_i$, for $i = 1, ..., m$.

23.1. Uniform Distribution on (a, b). Suppose $X_i^{(j)}$, $j = 1, ..., m$, has a Uniform(a, b) distribution

$$F_{X^{(j)}}(x_j) = \frac{x_j - a}{b - a}, \qquad a \leq x_j \leq b.$$

23.1.1. Upper Bound. First, we develop the upper bound $F_{\max}$:

$$F_{\max}(s) = \inf_{x_1+...+x_m=s} \min\left\{F_{X^{(1)}}(x_1) + ... + F_{X^{(m)}}(x_m), 1\right\}.$$

We use the recursive method presented previously. We ørst ønd

$$\begin{aligned}
& F_{\max(2)}(s-s_{m-2}) \\
= & \inf_{x_{m-1}+x_m=s-s_{m-2}} \min\left\{F_{X^{(m-1)}}(x_{m-1})+F_{X^{(m)}}(x_m),1\right\} \\
= & \inf_{x_{m-1}\in\mathbb{R}} \min\left\{F_{X^{(m-1)}}(x_{m-1})+F_{X^{(m)}}(s-s_{m-1}),1\right\} \\
= & \inf_{x_{m-1}\in\mathbb{R}} \min\left\{\frac{x_{m-1}-a}{b-a}+\frac{s-s_{m-1}-a}{b-a},1\right\} \\
= & \inf_{x_{m-1}\in\mathbb{R}} \min\left\{\frac{s-s_{m-2}-2a}{b-a},1\right\}. \qquad (41)
\end{aligned}$$

Since (41) does not depend on x_{m-1} anymore, then we do not have to diœerentiate the function to ønd the inømum. We then have

$$\begin{aligned}
F_{\max(2)}(s-s_{m-2}) &= \min\left\{\frac{s-s_{m-2}-2a}{b-a},1\right\} \\
&= \min\left\{\frac{s-x_1-\ldots-x_{m-2}-2a}{b-a},1\right\}.
\end{aligned}$$

Note that we cannot get rid of the minimum, since for $a\leq x_{m-1},x_m\leq b$,

$$\frac{s-x_1-\ldots-x_{m-2}-2a}{b-a}=\frac{x_{m-1}+x_m-2a}{b-a},$$

is not necessarily smaller than 1. The next step is then to ønd

$$\begin{aligned}
& F_{\max(3)}(s-s_{m-3}) \\
= & \inf_{x_{m-2}+x_{m-1}+x_m=s-s_{m-3}} \min\left\{\begin{array}{c} F_{X^{(m-2)}}(x_{m-2})+F_{X^{(m-1)}}(x_{m-1}) \\ +F_{X^{(m)}}(x_m),1 \end{array}\right\} \\
= & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{F_{X^{(m-2)}}(x_{m-2})+F_{\max(2)}(s-s_{m-2}),1\right\} \\
= & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{\frac{x_{m-2}-a}{b-a}+\min\left\{\frac{s-s_{m-2}-2a}{b-a},1\right\},1\right\}.
\end{aligned}$$

Since we have in general

$$\min\left\{F_X+\min\left\{F_Y,1\right\},1\right\}=\min\left\{F_X+F_Y,1\right\},$$

we can write

$$\begin{aligned}
F_{\max(3)}(s-s_{m-3}) &= \inf_{x_{m-2}\in\mathbb{R}} \min\left\{\frac{x_{m-2}-a}{b-a}+\frac{s-s_{m-2}-2a}{b-a},1\right\} \\
&= \inf_{x_{m-2}\in\mathbb{R}} \min\left\{\frac{s-s_{m-3}-3a}{b-a},1\right\}.
\end{aligned}$$

Again, we can get rid of the inømum, but not of the minimum

$$\begin{aligned} F_{\max(3)}\left(s-s_{m-3}\right) &= \min\left\{\frac{s-s_{m-3}-3a}{b-a},1\right\} \\ &= \min\left\{\frac{s-x_1-...-x_{m-3}-3a}{b-a},1\right\}. \end{aligned}$$

By applying this method recursively, we ønd

$$\begin{aligned} F_{\max(m-1)}\left(s-s_{1}\right) &= \min\left\{\frac{s-s_{1}-(m-1)\,a}{b-a},1\right\} \\ &= \min\left\{\frac{s-x_1-(m-1)\,a}{b-a},1\right\}, \end{aligned}$$

and ønally get

$$\begin{aligned} F_{\max(m)}(s) &= \inf_{x_1+...+x_m=s}\min\left\{F_{X^{(1)}}(x_1)+...+F_{X^{(m)}}(x_m),1\right\} \\ &= \inf_{x_1\in\mathbb{R}}\min\left\{F_{X^{(1)}}(x_1)+F_{\max(m-1)}(s-x_1),1\right\} \\ &= \inf_{x_1\in\mathbb{R}}\min\left\{\frac{x_1-a}{b-a}+\frac{s-x_1-(m-1)\,a}{b-a},1\right\} \\ &= \inf_{x_1\in\mathbb{R}}\min\left\{\frac{s-ma}{b-a},1\right\}. \end{aligned}$$

Again, we drop the inømum function, and get as ønal expression

$$F_{\max}(s) = F_{\max(m)}(s) = \min\left\{\frac{s-ma}{b-a},1\right\}.$$

23.1.2. Lower Bound. We now want to ønd the lower bound, $F_{\min}$. Since the uniform function is continuous, then $F_{X^{(j)}} = F^{-}_{X^{(j)}}$, and

$$F_{\min}(s) = \sup_{x_1+...+x_m=s}\max\left\{F_{X^{(1)}}(x_1)+...+F_{X^{(m)}}(x_m)-(m-1),0\right\}.$$

We use the same recursive method as that for the upper bound. We first find

$$
\begin{aligned}
& F_{\min(2)}\left(s-s_{m-2}\right) \\
= & \sup_{x_{m-1}+x_m=s-s_{m-2}} \max\left\{F_{X^{(m-1)}}\left(x_{m-1}\right)+F_{X^{(m)}}\left(x_m\right)-(2-1),0\right\} \\
= & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{F_{X^{(m-1)}}\left(x_{m-1}\right)+F_{X^{(m)}}\left(s-s_{m-1}\right)-1,0\right\} \\
= & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{\frac{x_{m-1}-a}{b-a}+\frac{s-s_{m-1}-a}{b-a}-1,0\right\} \\
= & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{\frac{s-s_{m-2}-2a}{b-a}-1,0\right\}. \qquad (42)
\end{aligned}
$$

Since (42) does not depend on x_{m-1} anymore, then we do not have to differentiate the function to find the supremum:

$$
\begin{aligned}
F_{\min(2)}\left(s-s_{m-2}\right) & = \max\left\{\frac{s-s_{m-2}-2a}{b-a}-1,0\right\} \\
& = \max\left\{\frac{s-x_1-\ldots-x_{m-2}-2a}{b-a}-1,0\right\}.
\end{aligned}
$$

However, since $a \leq x_{m-1}, x_m \leq b$, then

$$
\frac{s-x_1-\ldots-x_{m-2}-2a}{b-a}-1=\frac{x_{m-1}+x_m-(a+b)}{b-a},
$$

may be negative, and we have to keep the maximum function. The next step is then to find

$$
\begin{aligned}
& F_{\min(3)}\left(s-s_{m-3}\right) \\
= & \sup_{x_{m-2}+x_{m-1}+x_m=s-s_{m-3}} \max\left\{\begin{array}{c} F_{X^{(m-2)}}\left(x_{m-2}\right)+F_{X^{(m-1)}}\left(x_{m-1}\right) \\ +F_{X^{(m)}}\left(x_m\right)-(3-1),0 \end{array}\right\} \\
= & \sup_{x_{m-2}\in\mathbb{R}} \max\left\{F_{X^{(m-2)}}\left(x_{m-2}\right)+F_{\min(2)}\left(s-s_{m-2}\right)-1,0\right\} \\
= & \sup_{x_{m-2}\in\mathbb{R}} \max\left\{\frac{x_{m-2}-a}{b-a}+\max\left\{\frac{s-s_{m-2}-2a}{b-a}-1,0\right\}-1,0\right\}.
\end{aligned}
$$

We know that in general

$$
\max\left\{F_X+\max\left\{F_Y-1,0\right\}-1,0\right\}=\max\left\{F_X+F_Y-2,0\right\},
$$

because if F_Y-1 is positive, then there is no problem, and if F_Y-1 is negative, the expression is not disturbed since F_X-1 will be negative anyway and the whole expression ends up to be 0. We then write

$$\begin{aligned} F_{\min(3)}\left(s-s_{m-3}\right) &= \sup_{x_{m-2}\in\mathbb{R}} \max\left\{\frac{x_{m-2}-a}{b-a}+\frac{s-s_{m-2}-2a}{b-a}-2,0\right\} \\ &= \sup_{x_{m-2}\in\mathbb{R}} \max\left\{\frac{s-s_{m-3}-3a}{b-a}-2,0\right\}. \end{aligned}$$

Again, we can get rid of the supremum, but we have to keep the maximum

$$\begin{aligned} F_{\min(3)}\left(s-s_{m-3}\right) &= \max\left\{\frac{s-s_{m-3}-3a}{b-a}-2,0\right\} \\ &= \max\left\{\frac{s-x_1-\ldots-x_{m-3}-3a}{b-a}-2,0\right\}. \end{aligned}$$

By applying this method recursively, we ﬁnd

$$\begin{aligned} F_{\min(m-1)}\left(s-s_1\right) &= \max\left\{\frac{s-s_1-(m-1)\,a}{b-a}-(m-2)\,,0\right\} \\ &= \max\left\{\frac{s-x_1-(m-1)\,a}{b-a}-(m-2)\,,0\right\}, \end{aligned}$$

and ﬁnally get

$$\begin{aligned} & F_{\min(m)}\left(s\right) \\ = & \sup_{x_1+\ldots+x_m=s} \max\left\{F_{X^{(1)}}\left(x_1\right)+\ldots+F_{X^{(m)}}\left(x_m\right)-(m-1)\,,0\right\} \\ = & \sup_{x_1\in\mathbb{R}} \max\left\{F_{X^{(1)}}\left(x_1\right)+F_{\max(m-1)}\left(s-x_1\right)-1,0\right\} \\ = & \sup_{x_1\in\mathbb{R}} \max\left\{\frac{x_1-a}{b-a}+\frac{s-x_1-(m-1)\,a}{b-a}-(m-1)\,,0\right\} \\ = & \sup_{x_1\in\mathbb{R}} \max\left\{\frac{s-ma}{b-a}-(m-1)\,,0\right\}. \end{aligned}$$

Again, we drop the supremum and we get as ﬁnal expression

$$F_{\min}\left(s\right)=F_{\min(m)}\left(s\right)=\max\left\{\frac{s-ma}{b-a}-(m-1)\,,0\right\}.$$

23.2. Exponential Distribution. Suppose that the random variables $X_i^{(j)}$'s have an Exponential(λ) distribution. For $j=1,\ldots,m$, we have

$$F_{X^{(j)}}\left(x_j\right)=1-e^{-\lambda x_j}.$$

23.2.1. Upper Bound. First, we develop the upper bound, $F_{\max}$. We have

$$F_{\max}(s) = \inf_{x_1+...+x_m=s} \min\left\{F_{X^{(1)}}(x_1) + ... + F_{X^{(m)}}(x_m), 1\right\}.$$

We use the recursive method presented previously. We ørst ønd

$$\begin{aligned}
&F_{\max(2)}(s - s_{m-2}) \\
=\;& \inf_{x_{m-1}+x_m=s-s_{m-2}} \min\left\{F_{X^{(m-1)}}(x_{m-1}) + F_{X^{(m)}}(x_m), 1\right\} \\
=\;& \inf_{x_{m-1}\in\mathbb{R}} \min\left\{F_{X^{(m-1)}}(x_{m-1}) + F_{X^{(m)}}(s - s_{m-1}), 1\right\} \\
=\;& \inf_{x_{m-1}\in\mathbb{R}} \min\left\{1 - e^{-\lambda x_{m-1}} + 1 - e^{-\lambda(s-s_{m-1})}, 1\right\} \\
=\;& \inf_{x_{m-1}\in\mathbb{R}} \min\left\{2 - e^{-\lambda x_{m-1}} - e^{-\lambda(s-s_{m-1})}, 1\right\}. \qquad (43)
\end{aligned}$$

Since (43) is still function of x_{m-1}, we have to diœerentiate to ønd the inømum. We set the derivative with respect to x_{m-1} equal to 0, and we ønd $\lambda e^{-\lambda x_{m-1}} - \lambda e^{-\lambda(s-x_1-...-x_{m-1})} = 0$. Then, we isolate x_{m-1} and get

$$\begin{aligned}
x_{m-1} &= \frac{1}{2}(s - x_1 - ... - x_{m-2}) \\
&= \frac{1}{2}(s - s_{m-2}). \qquad (44)
\end{aligned}$$

Replacing (44) in $F_{\max(2)}$ yields

$$\begin{aligned}
&F_{\max(2)}(s - s_{m-2}) \\
=\;& \min\left\{2 - e^{-\frac{\lambda}{2}(s-s_{m-2})} - e^{-\lambda\left(s-s_{m-2}-\frac{1}{2}(s-s_{m-2})\right)}, 1\right\} \\
=\;& \min\left\{2 - e^{-\frac{\lambda}{2}(s-s_{m-2})} - e^{-\frac{\lambda}{2}(s-s_{m-2})}, 1\right\} \\
=\;& \min\left\{2 - 2e^{-\frac{\lambda}{2}(s-s_{m-2})}, 1\right\}.
\end{aligned}$$

The next step is then to ønd

$$\begin{aligned}
&F_{\max(3)}(s - s_{m-3}) \\
=\;& \inf_{x_{m-2}+x_{m-1}+x_m=s-s_{m-3}} \min\left\{\begin{array}{c} F_{X^{(m-2)}}(x_{m-2}) + F_{X^{(m-1)}}(x_{m-1}) \\ +F_{X^{(m)}}(x_m), 1 \end{array}\right\} \\
=\;& \inf_{x_{m-2}\in\mathbb{R}} \min\left\{F_{X^{(m-2)}}(x_{m-2}) + F_{\max(2)}(s - s_{m-2}), 1\right\} \\
=\;& \inf_{x_{m-2}\in\mathbb{R}} \min\left\{1 - e^{-\lambda x_{m-2}} + \min\left\{2 - 2e^{-\frac{\lambda}{2}(s-s_{m-2})}, 1\right\}, 1\right\}.
\end{aligned}$$

We can simplify the minimum functions as before, and write

$$\begin{aligned} & F_{\max(3)}\left(s-s_{m-3}\right) \\ = & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{1-e^{-\lambda x_{m-2}}+2-2e^{-\frac{\lambda}{2}\left(s-s_{m-2}\right)},1\right\} \\ = & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{3-e^{-\lambda x_{m-2}}-2e^{-\frac{\lambda}{2}\left(s-s_{m-2}\right)},1\right\}. \end{aligned}$$

Again, we have to diœerentiate to get the inømum. We set the derivative with respect to x_{m-2} equal to 0 and ønd $\lambda e^{-\lambda x_{m-2}}-\lambda e^{-\frac{\lambda}{2}(s-x_1-\ldots-x_{m-2})}=0$. Then, we isolate x_{m-2} to get the minimum:

$$\begin{aligned} x_{m-2} &= \frac{1}{3}\left(s-x_1-\ldots-x_{m-3}\right) \\ &= \frac{1}{3}\left(s-s_{m-3}\right). \end{aligned} \tag{45}$$

Replacing (45) in $F_{\max(3)}$ yields

$$\begin{aligned} & F_{\max(3)}\left(s-s_{m-3}\right) \\ = & \min\left\{3-e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)}-2e^{-\frac{\lambda}{2}\left(s-s_{m-3}-\frac{1}{3}\left(s-s_{m-3}\right)\right)},1\right\} \\ = & \min\left\{3-e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)}-2e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)},1\right\} \\ = & \min\left\{3-3e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)},1\right\}. \end{aligned}$$

By applying this method recursively, we ønd

$$\begin{aligned} F_{\max(m-1)}\left(s-s_1\right) &= \min\left\{(m-1)-(m-1)\,e^{-\frac{\lambda}{(m-1)}\left(s-s_1\right)},1\right\} \\ &= \min\left\{(m-1)-(m-1)\,e^{-\frac{\lambda}{(m-1)}\left(s-x_1\right)},1\right\}, \end{aligned}$$

and ønally get

$$\begin{aligned} & F_{\max(m)}(s) \\ = & \inf_{x_1+\ldots+x_m=s} \min\left\{F_{X^{(1)}}\left(x_1\right)+\ldots+F_{X^{(m)}}\left(x_m\right),1\right\} \\ = & \inf_{x_1\in\mathbb{R}} \min\left\{F_{X^{(1)}}\left(x_1\right)+F_{\max(m-1)}\left(s-x_1\right),1\right\} \\ = & \inf_{x_1\in\mathbb{R}} \min\left\{1-e^{-\lambda x_1}+(m-1)-(m-1)\,e^{-\frac{\lambda}{(m-1)}\left(s-x_1\right)},1\right\} \\ = & \inf_{x_1\in\mathbb{R}} \min\left\{m-e^{-\lambda x_1}-(m-1)\,e^{-\frac{\lambda}{(m-1)}\left(s-x_1\right)},1\right\}. \end{aligned}$$

We diœerentiate with respect to x_1, set the derivative equal to 0 and get $\lambda e^{-\lambda x_1}-\lambda e^{-\frac{\lambda}{(m-1)}(s-x_1)}=0$. We then isolate x_1 to get the inømum, and ønd $x_1 = \frac{s}{m}$. The upper bound is thus

$$\begin{aligned} F_{\max}(s) &= F_{\max(m)}(s) \\ &= \min\left\{m - me^{-\frac{\lambda s}{m}}, 1\right\} \\ &= \min\left\{m\left(1 - e^{-\frac{\lambda s}{m}}\right), 1\right\}. \end{aligned}$$

23.2.2. Lower Bound. We now want to ønd the lower bound, $F_{\min}$. Again, the Exponential distribution is continuous, so there is no need to specify the left limit of the distribution :

$$F_{\min}(s) = \sup_{x_1+\ldots+x_m=s} \max\left\{F_{X^{(1)}}(x_1) + \ldots + F_{X^{(m)}}(x_m) - (m-1), 0\right\}.$$

We use the same recursive method as that for the upper bound. We ørst ønd

$$\begin{aligned} & F_{\min(2)}(s - s_{m-2}) \\ = & \sup_{x_{m-1}+x_m=s-s_{m-2}} \max\left\{F_{X^{(m-1)}}(x_{m-1}) + F_{X^{(m)}}(x_m) - (2-1), 0\right\} \\ = & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{F_{X^{(m-1)}}(x_{m-1}) + F_{X^{(m)}}(s - s_{m-1}) - 1, 0\right\} \\ = & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{1 - e^{-\lambda x_{m-1}} + 1 - e^{-\lambda(s-s_{m-1})} - 1, 0\right\} \\ = & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{2 - e^{-\lambda x_{m-1}} - e^{-\lambda(s-s_{m-1})} - 1, 0\right\}. \end{aligned} \tag{46}$$

Since (46) is the same expression as for the upper bound (except for an additional constant), the derivative with respect to x_{m-1} is the same and leads to the same result:

$$x_{m-1} = \frac{1}{2}(s - x_1 - \ldots - x_{m-2}).$$

If we replace this expression in $F_{\min(2)}$, we get

$$\begin{aligned} & F_{\min(2)}(s - s_{m-2}) \\ = & \max\left\{2 - e^{-\frac{\lambda}{2}(s-s_{m-2})} - e^{-\lambda\left(s-s_{m-2}-\frac{1}{2}(s-s_{m-2})\right)} - 1, 0\right\} \\ = & \max\left\{2 - e^{-\frac{\lambda}{2}(s-s_{m-2})} - e^{-\frac{\lambda}{2}(s-s_{m-2})} - 1, 0\right\} \\ = & \max\left\{2 - 2e^{-\frac{\lambda}{2}(s-s_{m-2})} - 1, 0\right\}. \end{aligned}$$

The next step is then to find

$$\begin{aligned}
&F_{\min(3)}\left(s-s_{m-3}\right)\\
=&\sup_{x_{m-2}+x_{m-1}+x_m=s-s_{m-3}}\max\left\{\begin{array}{c}F_{X^{(m-2)}}\left(x_{m-2}\right)+F_{X^{(m-1)}}\left(x_{m-1}\right)\\+F_{X^{(m)}}\left(x_m\right)-(3-1),0\end{array}\right\}\\
=&\sup_{x_{m-2}\in\mathbb{R}}\max\left\{F_{X^{(m-2)}}\left(x_{m-2}\right)+F_{\min(2)}\left(s-s_{m-2}\right)-1,0\right\}\\
=&\sup_{x_{m-2}\in\mathbb{R}}\max\left\{1-e^{-\lambda x_{m-2}}+\max\left\{2-2e^{-\frac{\lambda}{2}\left(s-s_{m-2}\right)}-1,0\right\}-1,0\right\}.
\end{aligned}$$

By simplifying the maximum functions, we have

$$\begin{aligned}
&F_{\min(3)}\left(s-s_{m-3}\right)\\
=&\sup_{x_{m-2}\in\mathbb{R}}\max\left\{1-e^{-\lambda x_{m-2}}+2-2e^{-\frac{\lambda}{2}\left(s-s_{m-2}\right)}-1-1,0\right\}\\
=&\sup_{x_{m-2}\in\mathbb{R}}\max\left\{3-e^{-\lambda x_{m-2}}-2e^{-\frac{\lambda}{2}\left(s-s_{m-2}\right)}-2,0\right\}.
\end{aligned}$$

Again we differentiate with respect to x_{m-2} and set the derivative equal to 0. This leads to the same result as for the upper bound:

$$x_{m-2}=\frac{1}{3}\left(s-x_1-\ldots-x_{m-3}\right). \tag{47}$$

Replacing (47) in $F_{\min(3)}$ yields

$$\begin{aligned}
&F_{\min(3)}\left(s-s_{m-3}\right)\\
=&\max\left\{3-e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)}-2e^{-\frac{\lambda}{2}\left(s-s_{m-3}-\frac{1}{3}\left(s-s_{m-3}\right)\right)}-2,0\right\}\\
=&\max\left\{3-e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)}-2e^{-\frac{\lambda}{3}\left(s-s_{m-3}\right)}-2,0\right\}\\
=&\max\left\{3-3e^{-\frac{\lambda}{3}\left(s-x_1-\ldots-x_{m-3}\right)}-2,0\right\}.
\end{aligned}$$

By applying this method recursively, we find

$$F_{\min(m-1)}\left(s-s_1\right)=\max\left\{(m-1)-(m-1)\,e^{-\frac{\lambda}{(m-1)}\left(s-s_1\right)}-(m-2),0\right\},$$

and finally get

$$\begin{aligned}
& F_{\min(m)}(s) \\
= & \sup_{x_1+\ldots+x_m=s} \max\left\{F_{X^{(1)}}(x_1)+\ldots+F_{X^{(m)}}(x_m)-(m-1),0\right\} \\
= & \sup_{x_1\in\mathbb{R}} \max\left\{F_{X^{(1)}}(x_1)+F_{\max(m-1)}(s-x_1)-1,0\right\} \\
= & \sup_{x_1\in\mathbb{R}} \max\left\{\begin{array}{c} 1-e^{-\lambda x_1}+(m-1)-(m-1)e^{-\frac{\lambda}{(m-1)}(s-x_1)} \\ -(m-2)-1,0 \end{array}\right\} \\
= & \sup_{x_1\in\mathbb{R}} \max\left\{m-e^{-\lambda x_1}-(m-1)e^{-\frac{\lambda}{(m-1)}(s-x_1)}-(m-1),0\right\}.
\end{aligned}$$

We differentiate with respect to x_1, and set the derivative equal to 0 to find $\lambda e^{-\lambda x_1}-\lambda e^{-\frac{\lambda}{(m-1)}(s-x_1)}=0$. We then isolate x_1 to get the supremum, located at $x_1=\frac{s}{m}$. The lower bound is thus

$$\begin{aligned}
F_{\min}(s) &= F_{\min(m)}(s) \\
&= \max\left\{m-me^{-\frac{\lambda s}{m}}-(m-1),0\right\} \\
&= \max\left\{1-me^{-\frac{\lambda s}{m}},0\right\}.
\end{aligned}$$

23.3. Pareto Distribution. Suppose $X_i^{(j)}$, $j=1,\ldots,m$, have a Pareto(α,λ) distribution

$$F_{X^{(j)}}(x_j)=1-\left(\frac{\lambda}{\lambda+x_j}\right)^{\alpha}.$$

23.3.1. Upper Bound. We first develop the upper bound, $F_{\max}$. We have

$$F_{\max}(s)=\inf_{x_1+\ldots+x_m=s}\min\left\{F_{X^{(1)}}(x_1)+\ldots+F_{X^{(m)}}(x_m),1\right\}.$$

We use the recursive method presented previously. We first find

$$\begin{aligned}
& F_{\max(2)}(s-s_{m-2}) \\
= & \inf_{x_{m-1}+x_m=s-s_{m-2}} \min\left\{F_{X^{(m-1)}}(x_{m-1})+F_{X^{(m)}}(x_m),1\right\} \\
= & \inf_{x_{m-1}\in\mathbb{R}} \min\left\{F_{X^{(m-1)}}(x_{m-1})+F_{X^{(m)}}(s-s_{m-1}),1\right\} \\
= & \inf_{x_{m-1}\in\mathbb{R}} \min\left\{1-\left(\frac{\lambda}{\lambda+x_{m-1}}\right)^{\alpha}+1-\left(\frac{\lambda}{\lambda+(s-s_{m-1})}\right)^{\alpha},1\right\} \\
= & \inf_{x_{m-1}\in\mathbb{R}} \min\left\{2-\left(\frac{\lambda}{\lambda+x_{m-1}}\right)^{\alpha}-\left(\frac{\lambda}{\lambda+(s-s_{m-1})}\right)^{\alpha},1\right\}. \quad (48)
\end{aligned}$$

Since (48) is still function of x_{m-1}, we have to differentiate to find the infimum. We set the derivative with respect to x_{m-1} equal to 0, and we ønd $\frac{\alpha\lambda^{\alpha}}{(x_{m-1}+\lambda)^{\alpha+1}} - \frac{\alpha\lambda^{\alpha}}{((s-s_{m-1})+\lambda)^{\alpha+1}} = 0$. Then, we isolate x_{m-1} and get, as for the Exponential case

$$\begin{aligned} x_{m-1} &= \frac{1}{2}(s - x_1 - ... - x_{m-2}) \\ &= \frac{1}{2}(s - s_{m-2}). \end{aligned} \tag{49}$$

Replacing (49) in $F_{\max(2)}$ yields

$$\begin{aligned} & F_{\max(2)}(s - s_{m-2}) \\ = & \min\left\{ \begin{array}{c} 2 - \left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha} \\ -\left(\frac{\lambda}{\lambda+(s-s_{m-2})-\frac{1}{2}(s-s_{m-2})}\right)^{\alpha}, 1 \end{array} \right\} \\ = & \min\left\{2 - \left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha} - \left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha}, 1\right\} \\ = & \min\left\{2 - 2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha}, 1\right\}. \end{aligned}$$

The next step is then to ønd

$$\begin{aligned} & F_{\max(3)}(s - s_{m-3}) \\ = & \inf_{x_{m-2}+x_{m-1}+x_m = s-s_{m-3}} \min\left\{ \begin{array}{c} F_{X^{(m-2)}}(x_{m-2}) + F_{X^{(m-1)}}(x_{m-1}) \\ +F_{X^{(m)}}(x_m), 1 \end{array} \right\} \\ = & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{F_{X^{(m-2)}}(x_{m-2}) + F_{\max(2)}(s - s_{m-2}), 1\right\} \\ = & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{ \begin{array}{c} 1 - \left(\frac{\lambda}{\lambda+x_{m-2}}\right)^{\alpha} \\ +\min\left\{2 - 2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha}, 1\right\}, 1 \end{array} \right\}. \end{aligned}$$

We simplify the minimum functions to get

$$\begin{aligned} & F_{\max(3)}(s - s_{m-3}) \\ = & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{1 - \left(\frac{\lambda}{\lambda+x_{m-2}}\right)^{\alpha} + 2 - 2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha}, 1\right\} \\ = & \inf_{x_{m-2}\in\mathbb{R}} \min\left\{3 - \left(\frac{\lambda}{\lambda+x_{m-2}}\right)^{\alpha} - 2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha}, 1\right\}. \end{aligned}$$

Again, we have to diœerentiate to get the inømum. We set the derivative with respect to x_{m-2} equal to 0, and ønd

$$\frac{\alpha\lambda^{\alpha}}{(x_{m-2}+\lambda)^{\alpha+1}} - \frac{2\alpha(2\lambda)^{\alpha}}{((s - x_1 - ... - x_{m-2}) + 2\lambda)^{\alpha+1}} = 0.$$

Then, we isolate x_{m-2} to get the minimum:

$$\begin{aligned} x_{m-2} &= \frac{1}{3}\left(s - x_1 - \ldots - x_{m-3}\right) \\ &= \frac{1}{3}\left(s - s_{m-3}\right). \end{aligned} \tag{50}$$

Replacing (50) in $F_{\max(3)}$ leads to

$$\begin{aligned} & F_{\max(3)}\left(s - s_{m-3}\right) \\ = \; & \min\left\{3 - \left(\frac{3\lambda}{3\lambda + \left(s - s_{m-3}\right)}\right)^{\alpha} - 2\left(\frac{2\lambda}{2\lambda + \frac{2}{3}\left(s - s_{m-3}\right)}\right)^{\alpha}, 1\right\} \\ = \; & \min\left\{3 - \left(\frac{3\lambda}{3\lambda + \left(s - s_{m-3}\right)}\right)^{\alpha} - 2\left(\frac{3\lambda}{3\lambda + \left(s - s_{m-3}\right)}\right)^{\alpha}, 1\right\} \\ = \; & \min\left\{3 - 3\left(\frac{3\lambda}{3\lambda + \left(s - s_{m-3}\right)}\right)^{\alpha}, 1\right\}. \end{aligned}$$

By applying this method recursively, we find

$$\begin{aligned} & F_{\max(m-1)}\left(s - s_1\right) \\ = \; & \min\left\{(m-1) - (m-1)\left(\frac{(m-1)\lambda}{(m-1)\lambda + \left(s - s_1\right)}\right)^{\alpha}, 1\right\}, \end{aligned}$$

and we finally get

$$\begin{aligned} & F_{\max(m)}(s) \\ = \; & \inf_{x_1 + \ldots + x_m = s} \min\left\{F_{X^{(1)}}\left(x_1\right) + \ldots + F_{X^{(m)}}\left(x_m\right), 1\right\} \\ = \; & \inf_{x_1 \in \mathbb{R}} \min\left\{F_{X^{(1)}}\left(x_1\right) + F_{\max(m-1)}\left(s - x_1\right), 1\right\} \\ = \; & \inf_{x_1 \in \mathbb{R}} \min\left\{\begin{array}{c} 1 - \left(\frac{\lambda}{\lambda + x_1}\right)^{\alpha} + (m-1) \\ -(m-1)\left(\frac{(m-1)\lambda}{(m-1)\lambda + (s - x_1)}\right)^{\alpha}, 1 \end{array}\right\} \\ = \; & \inf_{x_1 \in \mathbb{R}} \min\left\{m - \left(\frac{\lambda}{\lambda + x_1}\right)^{\alpha} - (m-1)\left(\frac{(m-1)\lambda}{(m-1)\lambda + \left(s - x_1\right)}\right)^{\alpha}, 1\right\}. \end{aligned}$$

We differentiate with respect to x_1, set the derivative equal to 0 and find

$$\frac{\alpha\lambda^{\alpha}}{\left(x_1 + \lambda\right)^{\alpha+1}} - \frac{(m-1)\,\alpha\left((m-1)\,\lambda\right)^{\alpha}}{\left(\left(s - x_1\right) + (m-1)\,\lambda\right)^{\alpha+1}} = 0.$$

Isolating x_1 yields $x_1 = \frac{s}{m}$, and we then have for the upper bound

$$\begin{aligned} F_{\max}(s) &= F_{\max(m)}(s) \\ &= \min\left\{m - m\left(\frac{m\lambda}{m\lambda+s}\right)^{\alpha}, 1\right\} \\ &= \min\left\{m\left(1-\left(\frac{m\lambda}{m\lambda+s}\right)^{\alpha}\right), 1\right\}. \end{aligned}$$

23.3.2. Lower Bound. We now find the lower bound

$$F_{\min}(s) = \sup_{x_1+\ldots+x_m=s} \max\left\{F_{X^{(1)}}(x_1) + \ldots + F_{X^{(m)}}(x_m) - (m-1), 0\right\}.$$

We use the same recursive method as for the upper bound. We first find

$$\begin{aligned} & F_{\min(2)}(s - s_{m-2}) \\ = & \sup_{x_{m-1}+x_m=s-s_{m-2}} \max\left\{F_{X^{(m-1)}}(x_{m-1}) + F_{X^{(m)}}(x_m) - (2-1), 0\right\} \\ = & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{F_{X^{(m-1)}}(x_{m-1}) + F_{X^{(m)}}(s - s_{m-1}) - 1, 0\right\} \\ = & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{1-\left(\frac{\lambda}{\lambda+x_{m-1}}\right)^{\alpha} + 1 - \left(\frac{\lambda}{\lambda+(s-s_{m-1})}\right)^{\alpha} - 1, 0\right\} \\ = & \sup_{x_{m-1}\in\mathbb{R}} \max\left\{2-\left(\frac{\lambda}{\lambda+x_{m-1}}\right)^{\alpha} - \left(\frac{\lambda}{\lambda+(s-s_{m-1})}\right)^{\alpha} - 1, 0\right\} \quad (51) \end{aligned}$$

Since (51) is the same expression as for the upper bound (except for an additional constant), the derivative with respect to x_{m-1} is the same and leads to the same result:

$$\begin{aligned} x_{m-1} &= \frac{1}{2}(s - x_1 - \ldots - x_{m-2}) \\ &= \frac{1}{2}(s - s_{m-2}). \qquad (52) \end{aligned}$$

Replacing (52) in $F_{\min(2)}$ leads to

$$\begin{aligned} & F_{\min(2)}(s - s_{m-2}) \\ = & \max\left\{\begin{array}{c} 2-\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha} \\ -\left(\frac{\lambda}{\lambda+(s-s_{m-2})-\frac{1}{2}(s-s_{m-2})}\right)^{\alpha} - 1, 0 \end{array}\right\} \\ = & \max\left\{2-\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha} - \left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha} - 1, 0\right\} \\ = & \max\left\{2-2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^{\alpha} - 1, 0\right\}. \end{aligned}$$

The next step is then to ønd

$$\begin{aligned}
& F_{\min(3)}\left(s-s_{m-3}\right) \\
= & \sup_{x_{m-2}+x_{m-1}+x_m=s-s_{m-3}} \max\left\{\begin{array}{c} F_{X^{(m-2)}}\left(x_{m-2}\right)+F_{X^{(m-1)}}\left(x_{m-1}\right) \\ +F_{X^{(m)}}\left(x_m\right)-(3-1), 0 \end{array}\right\} \\
= & \sup_{x_{m-2}\in\mathbb{R}} \max\left\{F_{X^{(m-2)}}\left(x_{m-2}\right)+F_{\min(2)}\left(s-s_{m-2}\right)-1, 0\right\} \\
= & \sup_{x_{m-2}\in\mathbb{R}} \max\left\{\begin{array}{c} 1-\left(\frac{\lambda}{\lambda+x_{m-2}}\right)^\alpha \\ +\max\left\{2-2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^\alpha-1, 0\right\}-1, 0 \end{array}\right\}.
\end{aligned}$$

By simplifying the maximum functions, we get

$$\begin{aligned}
& F_{\min(3)}\left(s-s_{m-3}\right) \\
= & \sup_{x_{m-2}\in\mathbb{R}} \max\left\{\begin{array}{c} 1-\left(\frac{\lambda}{\lambda+x_{m-2}}\right)^\alpha+2 \\ -2\left(\frac{2\lambda}{2\lambda+(s-s_{m-2})}\right)^\alpha-1-1, 0 \end{array}\right\} \\
= & \sup_{x_{m-2}\in\mathbb{R}} \max\left\{3-\left(\frac{\lambda}{\lambda+x_{m-2}}\right)^\alpha-2\left(\frac{2\lambda}{2\lambda+\left(s-s_{m-2}\right)}\right)^\alpha-2, 0\right\}.
\end{aligned}$$

Again we diœerentiate with respect to x_{m-2} and set the derivative equal to 0. This leads to the same result as for the upper bound:

$$\begin{aligned}
x_{m-2} &= \frac{1}{3}\left(s-x_1-\ldots-x_{m-3}\right) \\
&= \frac{1}{3}\left(s-s_{m-3}\right). \qquad (53)
\end{aligned}$$

Replacing (53) in $F_{\min(3)}$ yields

$$\begin{aligned}
& F_{\min(3)}\left(s-s_{m-3}\right) \\
= & \max\left\{3-\left(\frac{3\lambda}{3\lambda+\left(s-s_{m-3}\right)}\right)^\alpha-2\left(\frac{2\lambda}{2\lambda+\frac{2}{3}\left(s-s_{m-3}\right)}\right)^\alpha-2, 0\right\} \\
= & \max\left\{3-\left(\frac{3\lambda}{3\lambda+\left(s-s_{m-3}\right)}\right)^\alpha-2\left(\frac{3\lambda}{3\lambda+\left(s-s_{m-3}\right)}\right)^\alpha-2, 0\right\} \\
= & \max\left\{3-3\left(\frac{3\lambda}{3\lambda+\left(s-s_{m-3}\right)}\right)^\alpha-2, 0\right\}.
\end{aligned}$$

By applying this method recursively, we ønd

$$\begin{aligned}
& F_{\min(m-1)}\left(s-s_1\right) \\
= & \max\left\{(m-1)-(m-1)\left(\frac{(m-1)\lambda}{(m-1)\lambda+\left(s-s_1\right)}\right)^\alpha-(m-2), 0\right\},
\end{aligned}$$

and finally get

$$\begin{aligned}
&F_{\min(m)}(s)\\
=\;&\sup_{x_1+\ldots+x_m=s}\max\left\{F_{X^{(1)}}(x_1)+\ldots+F_{X^{(m)}}(x_m)-(m-1),0\right\}\\
=\;&\sup_{x_1\in\mathbb{R}}\max\left\{F_{X^{(1)}}(x_1)+F_{\max(m-1)}(s-x_1)-1,0\right\}\\
=\;&\sup_{x_1\in\mathbb{R}}\max\left\{\begin{array}{c}1-\left(\frac{\lambda}{\lambda+x_1}\right)^{\alpha}+(m-1)\\ -(m-1)\left(\frac{(m-1)\lambda}{(m-1)\lambda+(s-x_1)}\right)^{\alpha}-(m-2)-1,0\end{array}\right\}\\
=\;&\sup_{x_1\in\mathbb{R}}\max\left\{\begin{array}{c}m-\left(\frac{\lambda}{\lambda+x_1}\right)^{\alpha}\\ -(m-1)\left(\frac{(m-1)\lambda}{(m-1)\lambda+(s-x_1)}\right)^{\alpha}-(m-1),0\end{array}\right\}.
\end{aligned}$$

We differentiate with respect to x_1, set the derivative equal to 0 and find

$$\frac{\alpha\lambda^{\alpha}}{(x_1+\lambda)^{\alpha+1}}-\frac{(m-1)\,\alpha\,((m-1)\,\lambda)^{\alpha}}{((s-x_1)+(m-1)\,\lambda)^{\alpha+1}}=0.$$

Isolating x_1 leads to $x_1=\frac{s}{m}$, and we then have for the lower bound

$$\begin{aligned}
F_{\max}(s) &= F_{\max(m)}(s)\\
&= \min\left\{m-m\left(\frac{m\lambda}{m\lambda+s}\right)^{\alpha}-(m-1),1\right\}\\
&= \min\left\{1-m\left(\frac{m\lambda}{m\lambda+s}\right)^{\alpha},1\right\}.
\end{aligned}$$

24. Partial Knowledge of Dependence

We now consider the case where we have some knowledge about the correlation structure between the $X_i^{(j)}$'s for a fixed i. Suppose that we are aware of the existence of a multivariate cdf H satisfying

$$H(x_1,x_2,\ldots,x_m)\le F_{\left(X^{(1)},X^{(2)},\ldots,X^{(m)}\right)}(x_1,x_2,\ldots,x_m),\qquad(54)$$

for all $x_1,x_2,\ldots,x_m\in\mathbb{R}$, and a joint decumulative distribution function $\overline{G}$ such that

$$\Pr\left(X_1^{(1)}>x_1,X_1^{(2)}>x_2,\ldots,X_1^{(m)}>x_m\right)\ge\overline{G}(x_1,x_2,\ldots,x_m),\qquad(55)$$

for all $x_1, x_2, ..., x_m \in \mathbb{R}$. When (54) and (55) hold, then Denuit et al. (1999) showed that

$$\begin{aligned}\sup_{(x_1,x_2,...,x_m)\in\Sigma(s)} H(x_1, x_2, ..., x_m) &\leq F_X(s) \\ &\leq 1 - \sup_{(x_1,x_2,...,x_m)\in\Sigma(s)} \overline{G}(x_1, x_2, ..., x_m),\end{aligned}$$

for all $s \in \mathbb{R}$.

In the special case where

$$H(x_1, x_2, ..., x_m) = \prod_{j=1}^{m} F_{X^{(j)}}(x_j), \tag{56}$$

for all $x_1, x_2, ..., x_m \in \mathbb{R}$, then the vectors $\left(X_i^{(1)}, X_i^{(2)}, ..., X_i^{(m)}\right)$ are said to be positively lower orthant dependent. If

$$\overline{G}(x_1, x_2, ..., x_m) = \prod_{j=1}^{m} \left(1 - F_{X^{(j)}}(x_j)\right), \tag{57}$$

for all $x_1, x_2, ..., x_m \in \mathbb{R}$, then the vectors $\left(X_i^{(1)}, X_i^{(2)}, ..., X_i^{(m)}\right)$ are said to be positively upper orthant dependent. When both (54) and (55) are fulølled with (56) and (57) respectively, the $\left(X_i^{(1)}, X_i^{(2)}, ..., X_i^{(m)}\right)$ are said to be positively orthant dependent (POD).

In the second report, we quoted the deønition of positive quadrant dependence, who is valid in the bivariate case. We can see that the concept of positively lower and upper orthant dependence is just the multivariate extension of the positive quadrant dependence. However, in this multivariate extension, (56) and (57) are not equivalent, by opposition to the bivariate case. Intuitively, (56) and (57) mean that the risks are more likely simultaneously to have small values and large ones, respectively, compared with a vector of independent risks with the same marginal distributions.

Stochastic Bounds on a Single Loss: Unknown Marginals

We now present bounds on F_X when only the support and the ørst few moments of the marginals $F_{X(j)}$, $j = 1, 2, ..., m$, are known. Again, we present the bounds for the $X_i^{(j)}$'s and from them it is possible to determine the bounds of F_X.

25. General Case: Risk Y

We ørst consider the general case with a non-negative random variable Y for which we know only the mean μ and the standard deviation σ. Then for all $s \geq 0$, there exist two cdf 's, $M^{(\mu,\sigma)}$ and $W^{(\mu,\sigma)}$, such that

$$M^{(\mu,\sigma)}(s) \leq F_Y(s) \leq W^{(\mu,\sigma)}(s). \tag{58}$$

We can ønd explicit expressions for these extremal distributions in Table 1, where δ_2 stands for $E(Y^2)$, the second moment of Y.

Value of s	$M^{(\mu,\sigma)}(s)$	$W^{(\mu,\sigma)}(s) - M^{(\mu,\sigma)}(s)$	$W^{(\mu,\sigma)}(s)$
$0 < s < \mu$	0	$\frac{\sigma^2}{(s-\mu)^2+\sigma^2}$	$\frac{\sigma^2}{(s-\mu)^2+\sigma^2}$
$\mu < s < \frac{\delta_2}{\mu}$	$\frac{s-\mu}{s}$	$\frac{\mu}{s}$	1
$s > \frac{\delta_2}{\mu}$	$\frac{(s-\mu)^2}{(s-\mu)^2+\sigma^2}$	$\frac{\sigma^2}{(s-\mu)^2+\sigma^2}$	1

Table 1. Extremal distributions in (58), two moments known, inønite spectrum

Obviously, the lower bound 0, and the two upper bounds 1 are just the natural bounds for a cdf. Also, since the random variable considered is non-negative, we can use the generalization of Markov's inequality:

$$\Pr(g(X) \geq a) \leq \frac{E(g(X))}{a},$$

for all $a > 0$. We then have

$$\Pr(Y \geq s) \leq \frac{E(Y)}{s} = \frac{\mu}{s},$$

implying that

$$\begin{aligned}
\Pr(Y < s) &= 1 - \Pr(Y \geq s) \\
&\geq 1 - \frac{\mu}{s} \\
&= \frac{s-\mu}{s}.
\end{aligned}$$

We then found the lower bound when $\mu < s < \frac{\delta_2}{\mu}$. At μ, $\Pr(Y < s) \geq \frac{\mu-\mu}{\mu} = 0$, and since it is just the natural bound for a cdf, the lower bound for $s < \mu$ is just 0. However, it is possible to use the fact that we know $E(Y^2)$ for the case where $s > \frac{\delta_2}{\mu}$. Again, we use the generalized Markov's inequality to get

$$\begin{aligned}
\Pr\left((Y-t)^2 \geq (s-t)^2\right) &\leq \frac{E\left((Y-t)^2\right)}{(s-t)^2} \\
&= \frac{E\left(Y^2 - 2Yt + t^2\right)}{(s-t)^2} \\
&= \frac{E\left(Y^2\right) - 2E(Y)\,t + t^2}{(s-t)^2} \\
&= \frac{\delta_2 - 2\mu t + t^2}{(s-t)^2}. \qquad (59)
\end{aligned}$$

Now, if we minimize this function with respect to t, the inequality will still be true. We then ønd the derivative and set it equal to 0

$$\begin{aligned}
0 &= \frac{2(t-\mu)(s-t)^2 + 2\left(\delta_2 - 2\mu t + t^2\right)(s-t)}{(s-t)^4} \\
&= \frac{2(t-\mu)}{(s-t)^2} + \frac{2\left(\delta_2 - 2\mu t + t^2\right)}{(s-t)^3}.
\end{aligned}$$

If we solve for t, we get

$$\begin{aligned}
0 &= (s-t)(t-\mu) + \left(\delta_2 - 2\mu t + t^2\right) \\
&= st - s\mu - t^2 + t\mu + \delta_2 - 2\mu t + t^2 \\
&= st - \mu s - \mu t + \delta_2,
\end{aligned}$$

and thus

$$t = \frac{\mu s - \delta_2}{s - \mu}. \qquad (60)$$

Replacing (60) in (59) yields

$$\begin{aligned}
&\frac{\delta_2 - 2\mu\left(\frac{\mu s-\delta_2}{s-\mu}\right) + \left(\frac{\mu s-\delta_2}{s-\mu}\right)^2}{\left(s - \left(\frac{\mu s-\delta_2}{s-\mu}\right)\right)^2} \\
= \quad & \frac{\delta_2 (s-\mu)^2 - 2\mu(\mu s - \delta_2)(s-\mu) + (\mu s - \delta_2)^2}{(s(s-\mu) - (\mu s - \delta_2))^2} \\
= \quad & \frac{\delta_2 s^2 - \mu^2 s^2 + 2\mu^3 s - \delta_2 \mu^2 - 2\mu s \delta_2 + \delta_2^2}{(s^2 - 2s\mu + \delta_2)^2} \\
= \quad & \frac{\sigma^2 s^2 + 2\mu^3 s - \delta_2 \mu^2 - 2\mu s(\sigma^2 + \mu^2) + \delta_2^2}{(s^2 - 2s\mu + \delta_2)^2} \\
= \quad & \frac{\sigma^2 s^2 + \delta_2 \sigma^2 - 2\sigma^2 \mu s}{(s^2 - 2s\mu + \delta_2)^2}.
\end{aligned}$$

By expressing δ_2 as $\sigma^2 + \mu^2$, it follows that

$$\begin{aligned}
\frac{E\left((Y-t)^2\right)}{(s-t)^2} &= \frac{\sigma^2 (s^2 + (\sigma^2 + \mu^2) - 2\mu s)}{(s^2 - 2s\mu + (\sigma^2 + \mu^2))^2} \\
&= \frac{\sigma^2 \left((s-\mu)^2 + \sigma^2\right)}{\left((s-\mu)^2 + \sigma^2\right)^2} \\
&= \frac{\sigma^2}{\left((s-\mu)^2 + \sigma^2\right)}.
\end{aligned}$$

We then have

$$\begin{aligned}
\Pr(Y \geq s) &= \Pr\left((Y-t)^2 \geq (s-t)^2\right) \\
&\leq \frac{\sigma^2}{\left((s-\mu)^2 + \sigma^2\right)}.
\end{aligned}$$

By looking at the expression we obtain for t at the minimum, we can see that this inequality is valid only for $s > \frac{\delta_2}{\mu}$, the value of t being either null or negative otherwise. It then follows

$$\begin{aligned}
\Pr(Y < s) &= 1 - \Pr(Y \geq s) \\
&= 1 - \Pr\left((Y-t)^2 \geq (s-t)^2\right) \\
&\geq 1 - \frac{\sigma^2}{\left((s-\mu)^2 + \sigma^2\right)} \\
&= \frac{(s-\mu)^2}{\left((s-\mu)^2 + \sigma^2\right)}.
\end{aligned}$$

When we also are aware that Y is subject to an upper bound b, i.e. $\Pr(Y \leq b) = 1$, it is possible to reformulate the extremal distributions in (58) as

$$M^{(\mu,\sigma,b)}(s) \leq F_Y(s) \leq W^{(\mu,\sigma,b)}(s), \ s \geq 0. \tag{61}$$

The explicit expressions for this case arise from a slight modiøcation of the results in Table 1, and are presented in Cossette et al. (2000).

When the skewness γ of Y is known, besides of μ and σ, then tighter bounds $M^{(\mu,\sigma,\gamma)}$ and $W^{(\mu,\sigma,\gamma)}$ can be derived such that

$$M^{(\mu,\sigma,\gamma)}(s) \leq F_Y(s) \leq W^{(\mu,\sigma,\gamma)}(s), \ s \geq 0. \tag{62}$$

We can ønd explicit expressions for these extremal distributions in Table 2, where δ_3 stands for $E(Y^3)$, the third moment of Y. The following expressions are also used:

$$\begin{aligned} \beta_1(s) &= \frac{\gamma + 3\sigma^2 + \mu^3 - s\delta_2}{\delta_2 - s\mu}, \\ \beta_2(s) &= \frac{\delta_2 - \mu s}{\mu - s}, \end{aligned}$$

and

$$\alpha_{\pm} = \frac{\delta_3 - \mu\delta_2 \pm \sqrt{(\delta_3 - \mu\delta_2)^2 - 4\sigma^2\left(\mu\delta_3 - \delta_2^2\right)}}{2\sigma^2}.$$

Value of s	$M^{(\mu,\sigma,\gamma)}(s)$	$W^{(\mu,\sigma,\gamma)}(s)$ $-M^{(\mu,\sigma,\gamma)}(s)$
$0 < s < \alpha_-$	0	$\frac{\mu-\beta_2(s)}{s-\beta_2(s)}$
$\alpha_- < s < \frac{\delta_2}{\mu}$	$\frac{\sigma^2+(\mu-s)(\mu-\beta_1(s))}{s\beta_1(s)}$	$\frac{\sigma^2+(\mu-\beta_1(s))\mu}{s(s-\beta_1(s))}$
$\frac{\delta_2}{\mu} < s < \alpha_+$	$\frac{(\mu-s)}{\beta_2(s)-s}$	$\frac{\mu-\beta_2(s)}{s-\beta_2(s)}$
$s > \alpha_+$	$\frac{\sigma^2+(\mu-s)(\mu-\beta_1(s))}{s\beta_1(s)} + \frac{\sigma^2+(\mu-s)\mu}{(\beta_1(s)-s)\beta_1(s)}$	$\frac{\sigma^2+(\mu-\beta_1(s))\mu}{s(s-\beta_1(s))}$

Table 2. Extremal distributions in (62), three moments known, inønite spectrum

As in (61), when Y is known to be bounded above, then we can improve the extremal distributions by ønding $M^{(\mu,\sigma,\gamma,b)}$ and $W^{(\mu,\sigma,\gamma,b)}$ such that

$$M^{(\mu,\sigma,\gamma,b)}(s) \leq F_Y(s) \leq W^{(\mu,\sigma,\gamma,b)}(s), \ s \geq 0. \tag{63}$$

The explicit expressions for this case arise from a slight modiøcation of the results in Table 2, and are presented in Cossette et al. (2000).

26. Losses X_i

Now, let $\mu^{(j)}, \sigma^{(j)}$ and $\gamma^{(j)}$ be the mean, the standard deviation and the skewness corresponding to $F_{X^{(j)}}$, $j = 1, 2, ..., m$. With Tables 1 and 2, we can ønd the best bounds on $F_{X^{(j)}}$, $j = 1, 2, ..., m$:

$$M_j(s) \leq F_{X^{(j)}}(s) \leq W_j(s), \tag{64}$$

where $M_j(s)$ stands for either $M^{\left(\mu^{(j)},\sigma^{(j)}\right)}$ or $M^{\left(\mu^{(j)},\sigma^{(j)},\gamma^{(j)}\right)}$ depending if the skewness is known or not. Similar notation is used for $W_j(s)$.

From these results, we can now say that

$$\widetilde{F}_{\min}(s) \leq F_X(s) \leq \widetilde{F}_{\max}(s), \;\; s \geq 0, \tag{65}$$

where

$$\widetilde{F}_{\min}(s) = \sup_{(x_1,x_2,...,x_m)\in\Sigma(s)} \max\left\{\sum_{j=1}^{m} \lim_{n\to\infty} M_j\left(x_j - \frac{1}{n}\right) - (m-1), 0\right\},$$

and

$$\widetilde{F}_{\max}(s) = \inf_{(x_1,x_2,...,x_m)\in\Sigma(s)} \min\left\{\sum_{j=1}^{m} W_j(x_j), 1\right\}.$$

If the $\left(X_i^{(1)}, X_i^{(2)}, ..., X_i^{(m)}\right)$'s are POD, then (65) still works, but we the improved bounds:

$$\widetilde{F}_{\min}(s) = \sup_{(x_1,x_2,...,x_m)\in\Sigma(s)} \prod_{j=1}^{m} M_j(x_j),$$

and

$$\widetilde{F}_{\max}(s) = 1 - \sup_{(x_1,x_2,...,x_m)\in\Sigma(s)} \prod_{j=1}^{m} (1 - W_j(x_j)).$$

If an upper bound b_j is also available such that i.e. $\Pr\left(X_i^{(j)} \leq b_j\right) = 1$, $j = 1, ..., m$, then it is easy to adjust the previous developments to handle this supplementary information.

Stochastic Bounds on the Total Amount of Loss

When N is independent of the X_i's, it is possible to use the convolution formula to determine the cdf of the aggregate claim S deøned in (38):

$$F_S(s) = \sum_{n=0}^{\infty} \Pr(N = n) F_X^{*n}(s), \ s \geq 0, \tag{66}$$

where $F_X^{*n}(s)$ is the n-fold convolution of F_X. The random variable N is discrete and has often either a Poisson or a Negative Binomial distribution. The choice of this discrete distribution may be based on the relation between the variance and the expectation. When the variance is rather close from the expectation a Poisson distribution may be appropriate, while if the variance is bigger than the expectation, the Negative Binomial distribution may be a better choice.

From (66), and from (40), it is easy to constrain F_S in the case where the marginals are known:

$$F_{S_{\min}}(s) \leq F_S(s) \leq F_{S_{\max}}(s), \ s \geq 0,$$

where

$$F_{S_{\min}}(s) = \sum_{n=0}^{\infty} \Pr(N = n) F_{\min}^{*n}(s),$$

and

$$F_{S_{\max}}(s) = \sum_{n=0}^{\infty} \Pr(N = n) F_{\max}^{*n}(s).$$

Usually, we use the Fast Fourier Transform (FFT) to approximate numerically the cdf of S. As usual, FFT can be used to approximate numerically $F_{S_{\min}}$ and $F_{S_{\max}}$.

When only the ørst moments (the mean, the variance and the skewness) are known, we can use (66) and (62) to derive bounds on F_S:

$$\widetilde{F}_{S_{\min}}(s) \leq F_S(s) \leq \widetilde{F}_{S_{\max}}(s), \ s \geq 0,$$

where

$$\widetilde{F}_{S_{\min}}(s) = \sum_{n=0}^{\infty} \Pr(N = n)\, \widetilde{F}_{\min}^{*n}(s),$$

and

$$\widetilde{F}_{S_{\max}}(s) = \sum_{n=0}^{\infty} \Pr(N = n)\, \widetilde{F}_{\max}^{*n}(s).$$

Note that the results presented for the extremal distributions of F_S are easily reøned to the POD cases (from the marginals or the moments).

Numerical Illustration

This example reproduces the illustration from Cossette et al. (2000). In order to illustrate the results presented, we consider $m = 2$, $X_1^{(1)}$ distributed according to the Exponential distribution with mean 1 and $X_1^{(2)}$ according to a Pareto with parameters 4 and 3, i.e.

$$\Pr\left(X_1^{(2)} \leq x\right) = 1 - \left(\frac{3}{3+x}\right)^4, \ x \geq 0.$$

In Figure 1, we plotted the cdf of $X_1^{(1)}$, together with its two-moment and three-moment approximations M_1 and W_1. We divided the support in three and four parts respectively, as in Tables 1 and 2, used $\mu^{(1)} = 1, \sigma^{2(1)} = 1$ and $\gamma^{(1)} = 2$ and computed the extremal distributions presented in the tables. We should notice that the authors specify that the skewness is 3, but in fact it is really 2. The graph we obtain is identical to their when $\gamma^{(1)} = 2$ is used.

The skewness is deøned to be

$$\begin{aligned} \gamma &= \frac{E\left((X-\mu)^3\right)}{\sigma^3} \\ &= \frac{E\left(X^3\right) - 3E\left(X^2\right)\mu + 2\mu^3}{\sigma^3}. \end{aligned}$$

For an Exponential(1), we know that $E\left(X^3\right) = 6$ and $E\left(X^2\right) = 2$, so

$$\gamma^{(1)} = \frac{6 - 3\,(2) + 2\,(1)}{\left(2 - 1^2\right)^{3/2}} = 2.$$

We repeated the same operations for Figure 2, which is similar to Figure 1, except that we now consider $X_1^{(2)}$, $\left(\mu^{(2)} = 1, \sigma^{2(2)} = 2, \gamma^{(2)} = \frac{20}{2^{3/2}}\right)$, which has a Pareto distribution instead of an Exponential. The graph we obtain presents some diœerences with that in the paper. This is due, again, to the skewness. The authors state that $\gamma^{(2)} = 23$. They assumed that $E\left(X^2\right) = 2$ in their calculations, while it really is $E\left(X^2\right) = 3$.

Since $E\left(X^3\right) = 27$, we have

$$\gamma^{(2)} = \frac{27 - 3\,(3) + 2\,(1)}{\left(3 - 1^2\right)^{3/2}} = \frac{20}{2^{3/2}}.$$

In spite of this diœerence, both øgures lead to the same conclusion: the knowledge of three moments gives tightest bounds.

Figure 3 presents the bounds for F_X, in particular the upper and lower bounds for the case where the marginals are known, and also the bounds for the three-moment case. We did not get a step function as the authors, and our ȷformatȷ for the graphs is not the same, because we did not use the same method to compute the bounds. For each s between 0 and 50, we determined $F_{\min}\left(s\right)$ and $F_{\max}\left(s\right)$ for numerous values of x and then chose the maximum and the minimum values within the vectors obtained. We used the same method to add the three-moment bounds. Figure 4 is similar to Figure 3, except that we added the lower POD bounds for the case where the marginals are known are for the three-moment case. We should precise that the method used reproduces the results obtained by the authors with a certain level of ødelity. However, it would probably be possible to increase the speed of the programs by using the built-in function nlmin in SPLUS, to get the inømums and the supremums needed. Since this function was unknown for us at the moment of writing the programs, we just overcame the problem of minimizing and maximizing by another method.

Finally, Figures 5 and 6 show the bounds on F_S (with and without the POD assumption). The random variable N is assumed to be distributed as a Poisson with mean 1. We then convolved the $F_{\min}\left(s\right)$, $F_{\max}\left(s\right)$, $\widetilde{F}_{\min}\left(s\right)$ and $\widetilde{F}_{\max}\left(s\right)$ and then multiplied them with the probability function of the Poisson to get these bounds. We should note that there is now a probability mass at 0 of $e^{-1} = 0.368$, because

$$\Pr\left(N = 0\right) = \frac{1^0 e^{-1}}{0!} = e^{-1} = 0.368.$$

Figure 5. Graph of $F_{X^{(1)}}$ together with $M_1^{(1,1)}$, $M_1^{(1,1,2)}$, $W_1^{(1,1)}$, and $W_1^{(1,1,2)}$.

Figure 6. Graph of $F_{X^{(2)}}$ together with $M_2^{(1,2)}$, $M_2^{\left(1,2,\frac{20}{2^{3/2}}\right)}$, $W_2^{(1,2)}$, and $W_2^{\left(1,2,\frac{20}{2^{3/2}}\right)}$.

Figure 3. Bounds on F_X.

Figure 4. Bounds on F_X in case of POD.

Figure 5. Bounds on F_S.

Figure 6. Bound on F_S in case of POD.

Conclusion

We presented some results for the upper and lower bounds of F_X. We derived the expression for the common uniform, exponential and Pareto cases, and we also implemented some SPLUS functions that allow the calculations for more complicated cases. We also presented how to derive bounds on F_S. These method have been tested in the numerical example reproduces from Cossette et al. (2000).

Bibliography

[1] Cossette, H., Denuit, M., Marceau, E. (2000). Impact of Dependence Among Multiple Claims in a Single Loss. Insurance: Mathematics and Economics 26, 213-222.

[2] Denuit, M., Genest, C., Marceau, E. (1999). Stochastic Bounds on Sums of Dependent Risks. Insurance: Mathematics and Economics 25, 85-104.

[3] Dhaene, J., Denuit, M. (1999). The Safest Dependence Structure Among Risks. Insurance: Mathematics and Economics 25, 11-21.

[4] Joe, H., (1997). Multivariate Models and Dependence Concepts. Chapman and Hall, London.

[5] Klugman, S.A., Panjer, H. H. and G. E. Willmot, (1998). Loss Models: From Data to Decisions, John Wiley, New York.

[6] Wang, S., Dhaene, J. (1998). Comonotonicity, Correlation Order and Premium Principles. Insurance: Mathematics and Economics 22, 235-242.

Appendix

26.1. Function e3m. This function calculates the necessary expressions to get the two-moment and the three-moment approximations.

```
e3m <- function(s, cond = ȷAȷ, mu = 1, sig = 1, skew = 2)
–
    e2 <- sig + mu^2
    e3 <- skew * (sqrt(sig))^3 + 3 * e2 * mu - 2 * mu^3
    x <- (0:(s * 1000)) * 0.001
    ap <- ((e3 - mu * e2) + sqrt((e3 - mu * e2)^2 - 4 * sig * (mu
            * e3 - e2^2)))/(2 * sig)
    am <- ((e3 - mu * e2) - sqrt((e3 - mu * e2)^2 - 4 * sig * (mu
            * e3 - e2^2)))/(2 * sig)
    xa <- x[1:sum(x < mu)]
    xb <- x[sum(x <= mu):sum(x < e2/mu)]
    xc <- x[sum(x <= e2/mu):length(x)]
    M2 <- c(0 * xa, ((xb - mu)/xb), ((xc - mu)^2/((xc - mu)^2 +
            sig)))[1:length(x)]
    W2 <- c((sig/((xa - mu)^2 + sig)), rep(1, length(xb) +
            length(xc)))[1:length(x)]
    x1 <- x[1:sum(x < am)]
    x2 <- x[sum(x <= am):sum(x < e2/mu)]
    x3 <- x[sum(x <= e2/mu):sum(x < ap)]
    x4 <- x[sum(x <= ap):length(x)]
    b12 <- (skew + 3 * sig + mu^3 - x2 * e2)/(e2 - x2 * mu)
    b14 <- (skew + 3 * sig + mu^3 - x4 * e2)/(e2 - x4 * mu)
    b21 <- (e2 - mu * x1)/(mu - x1)
    b23 <- (e2 - mu * x3)/(mu - x3)
    M3 <- c(0 * x1, (sig + (mu - x2) * (mu - b12))/(x2 * b12),
            (mu - x3)/(b23 - x3), (((sig + (mu - x4) * (mu - b14))
            /(x4 * b14)) + (sig + (mu - x4) * mu)/((b14 - x4) *
            b14)))[1:length(x)]
    W3 <- M3 + c((mu - b21)/(x1 - b21), (sig + (mu - b12) * mu)
            /(x2 * (x2 - b12)), (mu - b23)/(x3 - b23), (sig +
```

```
            (mu - b14) * mu)/(x4 * (x4 - b14)))[1:length(x)]
    if(cond == jAj)
        return(M3)
    if(cond == jBj)
        return(W3)
    if(cond == jCj)
        return(M2)
    if(cond == jDj)
        return(W2)
”
```

26.2. Function expo2. This function calculates the two-moment and the three-moment approximations. It returns either the plot of the exponential or that of the Pareto.

```
expo2 <- function(s = 30, cond = jPj, mu = 1, sig = 2,
                skew = 20/2^(3/2))
–
x <- (0:(s * 1000)) * 0.001
if(cond == jEj)
ex <- pexp(x, 1/mu)
if(cond == jPj)
ex <- 1 - (3/(3 + x))^4
M2 <- e3m(s, jCj, mu, sig, skew)
M3 <- e3m(s, jAj, mu, sig, skew)
W2 <- e3m(s, jDj, mu, sig, skew)
W3 <- e3m(s, jBj, mu, sig, skew)
plot(x, ex, type = jlj, ylab = jFx1j)
lines(x, M2, lty = 2)
lines(x, W2, lty = 3)
lines(x, M3, lty = 4)
lines(x, W3, lty = 5)
legend(5, 0.5, legend = c(jFx1j, jlower 2-moment approxj,
                jupper 2-moment approxj, jlower 3-moment
                approxj, jupper 3-moment approxj), lty = 1:5)
”
```

26.3. Function bds. This function calculates the two-moment and the three-moment approximations for X. It returns either the plot (with or without POD) or the numerical values of the desired bound.

```
bds <- function(cond = jAj)
–
    Fmin <- c()
    Fmax <- c()
```

```
F3min <- c()
F3max <- c()
Hx <- c()
H3x <- c()
for(s in 0:50) –
    x <- (0:(s * 1000)) * 0.001
    M1 <- e3m(s, ȷAȷ, 1, 1, 2)
    M2 <- e3m(s, ȷAȷ, 1, 2, 20/2^(3/2))
    W1 <- e3m(s, ȷBȷ, 1, 1, 2)
    W2 <- e3m(s, ȷBȷ, 1, 2, 20/2^(3/2))
    Fmin <- c(Fmin, max(c(0, 1 - exp( - x) - (3/(3 + s -
                x))^4)))
    Fmax <- c(Fmax, min(c(2 - exp( - x) - (3/(3 + s -
                x))^4, 1)))
    F3min <- c(F3min, max(c(0, M1 + M2[length(M2):1] -
                1)))
    F3max <- c(F3max, min(c(W1 + W2[length(W2):1],
                1)))
    Hx <- c(Hx, max((1 - exp( - x)) * (1 - (3/(3 + s - x))^4)))
    H3x <- c(H3x, max(M1 * M2[length(M2):1]))
”
if(cond == ȷAȷ) –
    plot(0:50, Fmin, type = ȷlȷ, xlab = ȷsȷ, ylab = ȷFxȷ)
    lines(0:50, Fmax, lty = 2)
    lines(0:50, F3min, lty = 3)
    lines(0:50, F3max, lty = 4)
    legend(10, 0.4, legend = c(ȷlower bd on Fxȷ, ȷupper bd
            on Fxȷ, ȷlower bd on Fx (3-moment approx)ȷ,
            ȷupper bd on Fx (3-moment approx)ȷ),
            lty = 1:4)
”
if(cond == ȷBȷ) –
    plot(0:50, Fmin, type = ȷlȷ, xlab = ȷsȷ, ylab = ȷFxȷ)
    lines(0:50, Fmax, lty = 2)
    lines(0:50, F3min, lty = 3)
    lines(0:50, F3max, lty = 4)
    lines(0:50, Hx, lty = 5)
    lines(0:50, H3x, lty = 6)
    legend(10, 0.4, legend = c(ȷlower bd on Fxȷ, ȷupper bd
            on Fxȷ, ȷlower bd on Fx (3-moment approx)ȷ,
            ȷupper bd on Fx (3-moment approx)ȷ, ȷlower bd
            on Fx with PODȷ, ȷlower bd on Fx with POD
```

```
                (3-moment approx)ᴊ), lty = 1:6
    )
    ″
    if(cond == ᴊCᴊ)
        return(Fmin)
    if(cond == ᴊDᴊ)
        return(Fmax)
    if(cond == ᴊEᴊ)
        return(F3min)
    if(cond == ᴊFᴊ)
        return(F3max)
    if(cond == ᴊGᴊ)
        return(Hx)
    if(cond == ᴊHᴊ)
        return(H3x)
″
```

26.4. Function foutr. This function convolves a distribution with the number of claims distributed as a Poisson with parameter 1.

```
foutr <- function(cdf)
–
    long <- 2ˆ12
    fx <- c(diœ(cdf), rep(0, long - length(cdf) + 1))
    mx <- œt(fx)
    len <- length(mx)
    mx2 <- matrix(mx, 11, len, byrow = T)ˆmatrix(rep(0:10, each
                = len), 11, len, byrow = T)
    pn <- matrix(rep(dpois(0:10, 1), each = len), 11, len, byrow
                = T)
    mx3 <- mx2 * pn
    ms <- apply(mx3, 2, sum)
    fs <- Re(œt(ms, inverse = T))[1:long]
    fs <- (fs >= 0) * fs
    fs <- fs/sum(fs)
    Fs <- cumsum(fs)
    return(Fs)
″
```

26.5. Function Fs2. This function calculates the bounds of Fs and plots them (either with or without POD).

```
Fs2 <- function(cond = ᴊAᴊ)
–
    Fmin <- bds(ᴊCᴊ)
```

```
Fmax <- bds(jDj)
F3min <- bds(jEj)
F3max <- bds(jFj)
Hx <- bds(jGj)
H3x <- bds(jHj)
Fsmin <- foutr(Fmin)
Fsmax <- foutr(Fmax)
Fs3min <- foutr(F3min)
Fs3max <- foutr(F3max)
Hs <- foutr(Hx)
H3s <- foutr(H3x)
if(cond == jAj) –
     plot(0:80, Fsmin[1:81], type = jlj, xlab = jsj,
          ylab = jFsj)
     lines(0:80, Fsmax[1:81], lty = 2)
     lines(0:80, Fs3min[1:81], lty = 3)
     lines(0:80, Fs3max[1:81], lty = 4)
     legend(15, 0.6, legend = c(jlower bd on Fsj, jupper bd
               on Fsj, jlower bd on Fs (3-moment approx)j,
               jupper bd on Fs (3-moment approx)j),
               lty = 1:4)
”
if(cond == jBj) –
     plot(0:80, Fsmin[1:81], type = jlj, xlab = jsj,
          ylab = jFsj)
     lines(0:80, Fsmax[1:81], lty = 2)
     lines(0:80, Fs3min[1:81], lty = 3)
     lines(0:80, Fs3max[1:81], lty = 4)
     lines(0:80, Hs[1:81], lty = 5)
     lines(0:80, H3s[1:81], lty = 6)
     legend(15, 0.6, legend = c(jlower bd on Fsj, jupper bd
               on Fsj, jlower bd on Fs (3-moment approx)j,
               jupper bd on Fs (3-moment approx)j, jlower
               bd on Fs with PODj, jlower bd on Fs with
               POD (3-moment approx)j), lty = 1:6)
”
”
```

Introduction

This report aims to complete what have been done about stochastic orders and ruin probabilities in the previous report. We ørst treat some properties of stop-loss order that are very useful in risk theory. Then, in the second section, we discuss ruin probabilities: we present a method to compute them theoretically, and we also discuss ordering in ruin probabilities. We close this section by reproducing the numerical example of Cossette and Marceau (2000). We ønally include in the appendix the proofs of some results quoted in the previous report, along with an example treating of bounds for PCD risks. The SPLUS programs used in the numerical example of the ruin probabilities are also presented in this appendix.

More on Stop-Loss Order

We presented previously the concept of stop-loss order, and its relation with other concepts, as the correlation order and the Frchet bounds. However, this notion also has pleasant invariance properties, which may be particularly useful in risk theory. This means that there exist many operations that we can perform on two ordered risks such that the result remains ordered in the same way. Among these operations, we ønd the convolution, the compounding and the mixing. The rest of this section aims to present some of those.

27. Invariance Properties of Stop-Loss Order

We ørst present a result stating that stop-loss order is maintained under convolution of two random variables.

Theorem 27.1. Let Z be a risk independent of risks X and Y. Then

$$X \leq_{sl} Y \Rightarrow X + Z \leq_{sl} Y + Z.$$

Proof. We consider conditional stop-loss premiums given the realization of the random variable Z, i.e. given $Z = z$. For all d, we have

$$\begin{aligned} E\left(X+Z-d\right)_{+} &= E_Z\left(E\left(\left(X+Z-d\right)_{+}|Z\right)\right) \\ &= \int_0^{\infty} E\left(\left(X+z-d\right)_{+}|Z=z\right)dF_Z\left(z\right) \\ &= \int_0^{\infty} E\left(\left(X-\left(d-z\right)\right)_{+}|Z=z\right)dF_Z\left(z\right) \\ &= \int_0^{\infty} E\left(X-\left(d-z\right)\right)_{+}dF_Z\left(z\right). \end{aligned}$$

By doing the same development for $Y + Z$, we ønd that

$$E\left(Y+Z-d\right)_{+} = \int_0^{\infty} E\left(Y-\left(d-z\right)\right)_{+}dF_Z\left(z\right).$$

For $X \leq_{sl} Y$, we have by deønition that for all retention level d,

$$E(X-d)_+ \leq E(Y-d)_+,$$

and it directly follows that

$$\int_0^\infty E(X-(d-z))_+ \, dF_Z(z) \leq \int_0^\infty E(Y-(d-z))_+ \, dF_Z(z).$$

This then implies that

$$E(X+Z-d)_+ \leq E(Y+Z-d)_+,$$

which means that $X+Z \leq_{sl} Y+Z$, and the result is veriøed. ∎

We present a simple example that veriøes this theorem.

Example 27.1. Let the random variables X and Y be exponentially distributed with parameter $\lambda_1 = 2$ and $\lambda_2 = 1$, respectively. For the general case where the parameter is $\lambda > 0$, the stop-loss premium for a retention level $d \geq 0$ is

$$\begin{aligned} E(S-d)_+ &= \int_d^\infty S_S(s)\, ds \\ &= \int_d^\infty e^{-\lambda s} ds \\ &= -\frac{e^{-\lambda s}}{\lambda}\bigg|_d^\infty \\ &= \frac{e^{-\lambda d}}{\lambda}. \end{aligned}$$

We can then say that $E(S-d)_+$ increases (decreases) as λ decreases (increases), since the derivative of $E(S-d)_+$ with respect to λ is negative:

$$\begin{aligned} \frac{dE(S-d)_+}{d\lambda} &= \frac{-d\lambda e^{-\lambda d} - e^{-\lambda d}}{\lambda^2} \\ &= \frac{-e^{-\lambda d}(1+d\lambda)}{\lambda^2}, \end{aligned}$$

which is necessarily negative since $d, \lambda \geq 0$. We now have for X and Y that

$$E(X-d)_+ = \frac{e^{-2d}}{2} \leq e^{-d} = E(Y-d)_+,$$

so $X \leq_{sl} Y$.

Now, we consider a random variable Z independent of X and Y, and having an Exponential distribution with parameter $\lambda_3 = 0.5$. This is a well-known result that a sum of n independent Exponential distributions with parameter λ_i is distributed as a Gamma with parameters $\alpha = n$ and $\lambda = \sum_{i=1}^{n} \lambda_i$. We then know that the sum $X+Z$ is distributed as a Gamma$(\alpha = 2, \lambda = 2 + 0.5 = 2.5)$, and similarly, the sum $Y + Z$ has a Gamma distribution with parameters $\alpha = 2$ and $\lambda = 1+0.5 = 1.5$.

For the general case, the survival distribution of a Gamma distribution with parameters $\alpha = 2$ and λ is

$$\begin{aligned} S_S(s) &= \int_s^\infty f_S(z)\, dz \\ &= \int_s^\infty \lambda^2 z e^{-\lambda z} dz \\ &= -\lambda z e^{-\lambda z} - e^{-\lambda z}\Big|_s^\infty \\ &= e^{-\lambda s}(1+\lambda s), \end{aligned}$$

for $s \geq 0$, and the stop-loss premium for a level of retention of $d \geq 0$ is

$$\begin{aligned} E(S-d)_+ &= \int_d^\infty S_S(s)\, ds \\ &= \int_d^\infty (1+\lambda s) e^{-\lambda s} ds \\ &= -\frac{2e^{-\lambda s}}{\lambda} - s e^{-\lambda s}\Big|_d^\infty \\ &= \left(\frac{2}{\lambda} + d\right) e^{-\lambda d}. \end{aligned}$$

Again, $E(S-d)_+$ is decreasing in the parameter λ:

$$\begin{aligned} \frac{dE(S-d)_+}{d\lambda} &= \frac{-2\lambda d e^{-\lambda d} - 2e^{-\lambda d}}{\lambda^2} - d^2 e^{-\lambda d} \\ &= \frac{-2e^{-\lambda d}(1+\lambda d)}{\lambda^2} - d^2 e^{-\lambda d}, \end{aligned}$$

which is necessarily negative since $d, \lambda \geq 0$.

For $X + Z$ and $Y + Z$, we then have

$$\begin{aligned} E(X+Z-d)_+ &= \left(\frac{2}{2.5} + d\right) e^{-2.5d} \\ &\leq \left(\frac{2}{1.5} + d\right) e^{-1.5d} = E(Y+Z-d)_+, \end{aligned}$$

so $X + Z \leq_{sl} Y + Z$, which veriøes Theorem 1.1.

We now present a generalization of the previous theorem, valid for a convolution of n independent risks.

Theorem 27.2. If $X_1, ..., X_n$ and $Y_1, ..., Y_n$ are sequences of independent risks, then

$$X_i \leq_{sl} Y_i \ \forall i \Rightarrow \sum_{i=1}^{n} X_i \leq_{sl} \sum_{i=1}^{n} Y_i.$$

Proof. We apply Theorem 1.1 with $X = X_1$, $Y = Y_1$ and $Z = X_2 + ... + X_n$. We then get

$$X_1 + X_2 + ... + X_n \leq_{sl} Y_1 + X_2 + ... + X_n. \tag{67}$$

If we apply Theorem 1.1 again, but with $X = X_2$, $Y = Y_2$, and $Z = Y_1 + X_3 + ... + X_n$, we get

$$Y_1 + X_2 + X_3 + ... + X_n \leq_{sl} Y_1 + Y_2 + X_3 + ... + X_n.$$

From (67), we can then write

$$X_1 + X_2 + X_3 + ... + X_n \leq_{sl} Y_1 + Y_2 + X_3 + ... + X_n.$$

If we continue to apply this method until every X_i, $i = 1, ..., n$ is replaced in the right-hand side, then the theorem is proved. ∎

We now discuss the case where a risk is produced by one of n sources. The index i for which $I_i = 1$ indicates which source produces the risk, and the other I_i are necessarily 0. This is called a mixed distribution, and its invariance property in stop-loss order is presented in the next result.

Theorem 27.3. Let $X_1, ..., X_n$ and $Y_1, ..., Y_n$ be two sequences of risks with $X_i \leq_{sl} Y_i$, for all $i = 1, ..., n$. If $I_1, ..., I_n$ have a joint distribution such that $I_1 + ... + I_n \equiv 1$ and, marginally, $P(I_i = 1) = p_i = 1 - P(I_i = 0)$, such that $\sum p_i = 1$, then stop-loss order is retained for the mixed random variables:

$$\sum_{i=1}^{n} I_i X_i \leq_{sl} \sum_{i=1}^{n} I_i Y_i. \tag{68}$$

Let F_i and G_i be the distribution functions of X_i and Y_i. Then for the cdf's of the mixed random variables in (68) we have

$$\sum_{i=1}^{n} p_i F_i \leq_{sl} \sum_{i=1}^{n} p_i G_i.$$

Proof. Since there is only one source at a time that can produce the risks, we have by conditioning on the I_i's:

$$\begin{aligned} E\left(\sum_{i=1}^{n} I_i X_i - d\right)_+ &= E\left(I_1X_1 + ... + I_nX_n - d\right)_+ \\ &= E\left(I_1X_1 + ... + I_nX_n - d \,|I_1 = 1\right)_+ \Pr\left(I_1 = 1\right) + ... \\ &\quad + E\left(I_1X_1 + ... + I_nX_n - d \,|I_n = 1\right)_+ \Pr\left(I_n = 1\right) \\ &= p_1 E\left(X_1 - d\right)_+ + ... + p_n E\left(X_n - d\right)_+ \\ &= \sum_{i=1}^{n} p_i E\left(X_i - d\right)_+ , \end{aligned}$$

and similarly

$$E\left(\sum_{i=1}^{n} I_i Y_i - d\right)_+ = \sum_{i=1}^{n} p_i E\left(Y_i - d\right)_+ .$$

Since $X_i \leq_{sl} Y_i$ for all $i = 1, ..., n$, it immediately follows from Theorem 1.2 that

$$\begin{aligned} E\left(\sum_{i=1}^{n} I_i X_i - d\right)_+ &= \sum_{i=1}^{n} p_i E\left(X_i - d\right)_+ \\ &\leq \sum_{i=1}^{n} p_i E\left(Y_i - d\right)_+ = E\left(\sum_{i=1}^{n} I_i Y_i - d\right)_+ , \end{aligned}$$

and hence by deønition

$$\sum_{i=1}^{n} I_i X_i \leq_{sl} \sum_{i=1}^{n} I_i Y_i .$$

∎

We present an example illustrating this result.

Example 27.2. Suppose that $X_1, X_2,$ and X_3 have Exponential distributions with parameter $\lambda_1 = 1, \lambda_2 = 4,$ and $\lambda_3 = 6$ respectively. Similarly, $Y_1, Y_2,$ and Y_3 have Exponential distributions with parameter $\lambda_1' = 5, \lambda_2' = 5,$ and $\lambda_3' = 8$. The probabilities that the risk arises from the ith distribution are $p_1 = 0.1$, $p_2 = 0.5$, and $p_3 = 0.4$, and the I_i,

$i = 1, 2, 3$, are of course mutually exclusive. We then have

$$\begin{aligned} E\left(\sum_{i=1}^{3} I_i X_i - d\right)_+ &= \sum_{i=1}^{3} p_i E\left(X_i - d\right)_+ \\ &= 0.1e^{-d} + 0.5\left(\frac{e^{-4d}}{4}\right) + 0.4\left(\frac{e^{-6d}}{6}\right), \end{aligned}$$

and

$$\begin{aligned} E\left(\sum_{i=1}^{3} I_i Y_i - d\right)_+ &= \sum_{i=1}^{3} p_i E\left(Y_i - d\right)_+ \\ &= 0.1\left(\frac{e^{-5d}}{5}\right) + 0.5\left(\frac{e^{-5d}}{5}\right) + 0.4\left(\frac{e^{-8d}}{8}\right). \end{aligned}$$

From Example 1.1, we know that $X_i \leq_{sl} Y_i$, $i = 1, 2, 3$. It then follows from Theorem 1.2 that a weighted sum of these X_i's is automatically smaller in stop-loss order than a weighted sum of these Y_i's, provided that the weights are the same. The invariance property for mixed random variables is then veriøed.

A very popular type of model in risk theory is the compound distribution. This model assumes that the total claim of a portfolio is a sum of a random number N of independent and identically distributed (iid) claims X_i. We present a result stating that if either X_i or N is replaced by a riskier variable, then the resulting compound distribution is also riskier.

Theorem 27.4. Let $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ be sequences of iid risks. Let N and M be counting variables independent of X_i and Y_i. If $X_i \leq_{sl} Y_i$ for all i, and $N \leq_{sl} M$, we have

$$\sum_{i=1}^{N} X_i \leq_{sl} \sum_{i=1}^{N} Y_i,$$

as well as

$$\sum_{i=1}^{N} X_i \leq_{sl} \sum_{i=1}^{M} X_i.$$

Proof. Let $S_n = X_1 + \ldots + X_n$, $n = 1, 2, \ldots$ and $S_0 = 0$. If F and G are the distribution functions of the risks X_i and Y_i, then the n-fold convolution of F with itself, written F^{*n}, is the distribution function of S_n. We can then write

$$\Pr\left(\sum_{i=1}^{N} X_i \leq x\right) = \sum_{n=0}^{\infty} P\left(N = n\right) F^{*n}\left(x\right).$$

By Theorem 1.2, we know that $\sum_{i=1}^{n} X_i \leq \sum_{i=1}^{n} Y_i$, and thus that $F^{*n} \leq_{sl} G^{*n}$. From Theorem 1.3 for mixed distributions, we then obtain

$$\sum_{n=0}^{\infty} P\left(N = n\right) F^{*n}\left(x\right) \leq_{sl} \sum_{n=0}^{\infty} P\left(N = n\right) G^{*n}\left(x\right).$$

The ørst part of the theorem is proved. The second part follows immediately from the fact that for a retention level $d \geq 0$, $E\left(S_n - d\right)$, $n = 0, 1, 2, \ldots$ is a non-decreasing convex function. For more details, see Kaas et al. (1994). ∎

The last invariance property we present is the conditional compound Poisson distribution.

Theorem 27.5. Let Λ_j be a non-negative structure variable, and N_j be an integer valued non-negative random variable. The joint distribution of $\left(\Lambda_j, N_j\right)$ is such that given $\Lambda_j = \lambda$, N_j is Poisson (λ) distributed, $j = 1, 2$. Let $X_1, X_2, \ldots$ be a sequence of iid risks, and let

$$S_j = \sum_{i=1}^{N_j} X_i.$$

Then, $\Lambda_1 \leq_{sl} \Lambda_2$ implies $S_1 \leq_{sl} S_2$.

Proof. See Kaas et al. (1994). ∎

It immediately follows from this theorem that a Poisson(λ_1) is smaller in stop-loss order than a Poisson(λ_2) when $\lambda_1 \leq \lambda_2$. We also present in an example a result that will be useful thereafter in our numerical example of ruin probabilities. We want to show that the Negative Binomial distribution is riskier than the Poisson distribution when the means are equal.

Example 27.3. Consider a Poisson(μ) and a Negative Binomial(r, p) distributions with equal means, so

$$\mu = \frac{rp}{1-p} \Rightarrow \mu - \mu p = rp \Rightarrow \mu = rp + \mu p,$$

and then $p = r/\left(\mu + r\right)$. Since a Negative Binomial is just a mixture of Poisson distributions with a Gamma used as mixing distribution, we can express both distributions in this example as a mixture of Poisson distributions. By Theorem 1.5, all we have to show is that the mixing distribution used for the Negative Binomial is riskier than that for the Poisson.

Let Λ_1 have a degenerate distribution such that $\Pr(\Lambda_1 = \mu) = 1$, and let Λ_2 have a Gamma(r, β) distribution with $\beta = p/(1-p) = r/\mu$. Note that we used the well-known result of a mixture Poisson-Gamma resulting in a Negative Binomial, and øxed the parameters of the Gamma so that the Negative Binomial has parameters r and $p = r/(r+\mu)$. Λ_2 then has a Gamma distribution, $M_2 | \Lambda_2 = \lambda$ has a Poisson distribution, and the unconditional distribution of M_2 is Negative Binomial.

To show that $\Lambda_1 \leq_{sl} \Lambda_2$, we have to consider the stop-loss premiums of these random variables. For Λ_1, we have

$$E(\Lambda_1 - d)_+ = \max(0, \mu - d),$$

where $d \geq 0$ is the retention level. For Λ_2, we have to use the relation

$$E(\Lambda_2 - d) = \frac{E(\Lambda_2) - E(\Lambda_2 \wedge d)}{\Pr(\Lambda_2 > d)},$$

where $E(\Lambda_2 \wedge d)$ is the expectation of Λ_2 limited at d. From Klugman et al. (1998), we have for $\Lambda_2 \sim$ Gamma$\left(r, \frac{r}{\mu}\right)$

$$E(\Lambda_2 \wedge d) = r\left(\frac{\mu}{r}\right)\Gamma\left(r+1; \frac{dr}{\mu}\right) + d\left(1 - \Gamma\left(r; \frac{dr}{\mu}\right)\right),$$

where

$$P(\Lambda_2 \leq x) = \Gamma\left(r; \frac{xr}{\mu}\right) = 1 - \sum_{j=0}^{r-1} \frac{(xr/\mu)^j e^{-(xr/\mu)}}{j!}, \qquad (69)$$

is the cdf of Λ_2 for an integer r. Note that when r is not an integer, $\Gamma\left(r; \frac{xr}{\mu}\right)$ is given by the incomplete gamma function. We then have that

$$\begin{aligned} E(\Lambda_2 - d) &= \frac{r\left(\frac{\mu}{r}\right) - r\left(\frac{\mu}{r}\right)\Gamma\left(r+1; \frac{dr}{\mu}\right) - d\left(1 - \Gamma\left(r; \frac{dr}{\mu}\right)\right)}{1 - \Gamma\left(r; \frac{dr}{\mu}\right)} \\ &= \frac{\mu\left(1 - \Gamma\left(r+1; \frac{dr}{\mu}\right)\right)}{1 - \Gamma\left(r; \frac{dr}{\mu}\right)} - d, \end{aligned}$$

and hence

$$E(\Lambda_2 - d)_+ = \max\left(0, \frac{\mu\left(1-\Gamma\left(r+1; \frac{dr}{\mu}\right)\right)}{1-\Gamma\left(r; \frac{dr}{\mu}\right)} - d\right).$$

From (69), we know that for r being an integer,

$$\Gamma\left(r; \frac{xr}{\mu}\right) \geq \Gamma\left(r+1; \frac{xr}{\mu}\right) \Rightarrow 1-\Gamma\left(r; \frac{xr}{\mu}\right) \leq 1-\Gamma\left(r+1; \frac{xr}{\mu}\right).$$

Note that the same relation holds when r is not an integer, see Klugman et al (1998) p.570. From this inequality, we can aŒrm that

$$\frac{1-\Gamma\left(r+1; \frac{dr}{\mu}\right)}{1-\Gamma\left(r; \frac{dr}{\mu}\right)} \geq 1,$$

and then

$$\begin{aligned} E(\Lambda_1 - d)_+ &= \max(0, \mu - d) \\ &\leq \max\left(0, \frac{\mu\left(1-\Gamma\left(r+1; \frac{dr}{\mu}\right)\right)}{1-\Gamma\left(r; \frac{dr}{\mu}\right)} - d\right) = E(\Lambda_2 - d)_+, \end{aligned}$$

so $\Lambda_1 \leq_{sl} \Lambda_2$.

Now, since M_j, $j = 1, 2$, are such that $M_j | \Lambda_j = \lambda$ has a Poisson (λ) distribution, we have by Theorem 1.5 that $M_1 \leq_{sl} M_2$, and then the Negative Binomial is riskier than the Poisson.

Ruin Probabilities

We presented previously the basics of ruin theory, and simulated the numerical example presented in Cossette and Marceau (2000). We now deepen this subject in some way, as we discuss the numerical methods to approximate ruin probabilities. We also present a ruin order valid in speciøc situations and quote its relation with stop-loss order. We close this section by reproducing the numerical results obtained in Cossette and Marceau (2000), but this time in a theoretical way rather than using simulation.

28. Evaluation of Ruin Probabilities

In the previous report, we presented the discrete-ruin model, but we only went over the deønitions of the ruin and non-ruin probabilities, without discussing the way of computing these probabilities in practice. The following theorem presents a recursive method allowing to compute the non-ruin probabilities over the periods 1 to n.

Theorem 28.1. Let $\{W_i, i = 1, 2, ...\}$ be a sequence of iid random variables and c the annual premium income constant over each period. Then,

$$\phi(u, 1, n) = \int_0^{u+c} \phi(u + c - w, 1, n - 1)\, dF_W(w). \tag{70}$$

Proof. See Cossette and Marceau (2000). ∎

Unfortunately, the calculation of exact values of $\phi(u, 1, n)$ from (70) is rarely possible. However, there exists an algorithm approximating $\phi(u, 1, n)$, but this algorithm requires the discretization of the distribution function F_W. We assume that $F_{\widetilde{W}}$ is the discrete distribution obtained after the process of discretization, and $\widetilde{W}$ the corresponding discrete random variable. If

$$P\left(\widetilde{W} = k\right) = f_k,$$

for $k = 0, 1, ..., M$, then

$$F_{\widetilde{W}}(k) = P\left(\widetilde{W} \leq k\right) = \sum_{j=0}^{k} f_j,$$

where f_k are the mass probabilities. We suppose that the premium income is a constant integer p and that the surplus process takes only integer values. We denote by $\psi_{k,1,n}$ and $\phi_{k,1,n}$ the ønite-time ruin and non-ruin probabilities calculated with $F_{\widetilde{W}}$ over the periods 1 to n with an initial surplus k (an integer).

Theorem 28.2. Let k, p, j be integers. Then,

$$\phi_{k,1,n} = \sum_{j=1}^{\min(k+p,M)} \phi_{k+p-j,1,n-1} f_j, \tag{71}$$

for $n = 2, 3, ...,$ where

$$\phi_{k,1,1} = F_{\min(k+p,M)} = \sum_{j=0}^{\min(k+p,M)} f_j. \tag{72}$$

Proof. See Cossette and Marceau (2000). ∎

From this theorem, we can get directly the ruin probability $\psi_{k,1,n}$, since it is the complement of $\phi_{k,1,n}$:

$$\psi_{k,1,n} = 1 - \phi_{k,1,n}.$$

29. Computing the Discrete Distribution $F_{\widetilde{W}}$

Obtaining the distribution function of F_W is not always evident. A popular method is to use the Fast Fourier Transform (FFT) to get the characteristic functions ϕ. For a book of business, the ørst step consists in discretizing the severity distributions of each class of business, $F_{X^{(j)}}$, $j = 1, ..., m$, and take their Fourier transform. From the obtained $\phi_{\widetilde{X}^{(j)}}$'s, we then calculate the characteristic function of $\widetilde{W}$, $\phi_{\widetilde{W}}$. We invert the function $\phi_{\widetilde{W}}$ with the FFT method, and this produces the vector of mass probabilities deøning the probability distribution function $f_{\widetilde{W}}$, from which it is easy to get $F_{\widetilde{W}}$. We can see that F_W is not directly discretized, but it is rather the distribution of the $X^{(j)}$'s that are discretized in order to determine $F_{\widetilde{W}}$. We can ønally use this approximation of F_W and insert it in (71) and (72) to evaluate $\phi_{k,1,n}$.

We now present the relation between ϕ_W and the $\phi_{X^{(j)}}$'s in order to determine the characteristic function of W. Note that we can also apply this relation to the discrete versions of W and $X^{(j)}$, $j = 1, ..., m$, that is $\widetilde{W}$ and $\widetilde{X}^{(j)}$, $j = 1, ..., m$:

$$\begin{aligned}\phi_W(t) &= \phi_{W^{(1)},W^{(2)},W^{(3)}}(t,t,t) \\ &= P_{N^{(1)},N^{(2)},N^{(3)}}\left(\phi_{X^{(1)}}(t), \phi_{X^{(2)}}(t), \phi_{X^{(3)}}(t)\right),\end{aligned}$$

where $P_N(t)$ is the probability generating function (pgf) of N.

It is now possible to obtain explicit expressions of $\phi_W(t)$ for the Poisson model with common shock and for the Negative Binomial with common component. We consider the case with three classes of business in a book of business, but this can be easily generalized to higher dimensions. For the Poisson model with common shock, we get

$$\begin{aligned}\phi_W(t) &= \phi_{W^{(1)},W^{(2)},W^{(3)}}(t,t,t) \\ &= \exp\left(\lambda\left(\phi_X(t)-1\right)\right), \qquad (73)\end{aligned}$$

where

$$\lambda = \lambda_{11}+\lambda_{22}+\lambda_{33}+\lambda_{12}+\lambda_{13}+\lambda_{23}+\lambda_{123},$$

and

$$\begin{aligned}\phi_X(t) &= \frac{\lambda_{11}}{\lambda}\phi_{X^{(1)}}(t)+\frac{\lambda_{22}}{\lambda}\phi_{X^{(2)}}(t)+\frac{\lambda_{33}}{\lambda}\phi_{X^{(3)}}(t) \\ &\quad+\frac{\lambda_{12}}{\lambda}\phi_{X^{(1)}+X^{(2)}}(t)+\frac{\lambda_{13}}{\lambda}\phi_{X^{(1)}+X^{(3)}}(t) \\ &\quad+\frac{\lambda_{23}}{\lambda}\phi_{X^{(2)}+X^{(3)}}(t)+\frac{\lambda_{123}}{\lambda}\phi_{X^{(1)}+X^{(2)}+X^{(3)}}(t).\end{aligned}$$

Note that the characteristic function of the convolution of two independent random variables, say X and Y, is just the product of their respective characteristic function:

$$\phi_{X+Y}(t) = \phi_X(t)\,\phi_Y(t).$$

For the Negative Binomial model with common component, we use

$$\begin{aligned}\phi_W(t) &= \phi_{W^{(1)},W^{(2)},W^{(3)}}(t,t,t) \\ &= A\times B, \qquad (74)\end{aligned}$$

where

$$A = \prod_{i=1}^{3}\left[1-\beta_j\left(\phi_{X^{(j)}}(t)-1\right)\right]^{-\alpha_{jj}},$$

and

$$B = \left[1-\beta_1\left(\phi_{X^{(1)}}(t)-1\right)-\beta_2\left(\phi_{X^{(2)}}(t)-1\right)-\beta_3\left(\phi_{X^{(3)}}(t)-1\right)\right]^{-\alpha_{j0}}.$$

For more details on how these distributions are found, refer to Cossette and Marceau (2000).

30. Order in Ruin

We now discuss the inÆuence of increasing risk on ruin probabilities. It is sensed to think that for risks with equal means, the least variable risk in that class is the safer risk, that is the risk having minimum ruin probability. Also, it is possible to show that in general, when a risk has uniformly smallest stop-loss premiums, then this risk has uniformly smallest ruin probabilities. From this aŒrmation, we then quote a theorem presenting a link between stop-loss order and ruin probabilities.

Theorem 30.1. Consider two compound Poisson risk processes with equal premium per unit time and also equal Poisson parameter, but with diœerent individual claims X and Y, and ruin probabilities $\psi_X(u)$ and $\psi_Y(u)$, respectively. Then,

$$X \leq_{sl} Y \Rightarrow \psi_X(u) \leq \psi_Y(u) \quad \forall u \geq 0.$$

Proof. See Kaas et al. (1994). ∎

We should notice that stop-loss order implies ordered ruin probabilities, but the converse is not necessarily true. Also, if we consider the variability order (see Kaas et al. (1994)), then we can say that stop-loss order implies variability order, which in turn implies ordered ruin probabilities. So, in many circumstances, we can identify the element of a class of risks providing the largest ruin probability, just by ønding the most variable element in the class. Note that we do not present the deønition of variability order, since it refers to some deønitions we did not present in this report. For more details about this notion, refer to Kaas et al. (1994).

We now consider the example presented in Cossette and Marceau (2000), the same example that we simulated in the previous report. We noticed through these simulations that for the same severity distributions, the Negative Binomial model with common component gives uniformly bigger ruin probabilities than the Poisson model with common shock. It would then be interesting to determine if there exists a stop-loss order between these two models.

When the coeŒcient of correlation is 0, i.e. when there are no possibility of common shock (independent case), this is easily seen from Example 1.3 and from Theorem 1.5 that the Negative Binomial model is riskier than the Poisson model, since the mean of a Negative Binomial with parameters $\alpha = 1, \beta = 4$ is equal to that of a Poisson(4).

However, when the coeŒcient of correlation is non-zero, we cannot make the same aŒrmation even if we are tempted to, according to the results of the simulation. From Theorem 1.5, we know that a random variable X distributed as a Poisson(λ_1) is smaller in stop-loss order than a random variable Y having a Poisson(λ_2) distribution, for $\lambda_1 \leq \lambda_2$. Also, we know that for a Poisson and a Negative Binomial with the same mean, the Negative Binomial is riskier. This then implies that a Poisson with a smaller mean than a Negative Binomial is smaller in stop-loss order. However, the problem comes from the fact that we cannot conclude anything for the case where the mean of the Poisson is bigger than that of the Negative Binomial. For the same mean, the Negative Binomial is riskier, but if it has a smaller mean, is the diœerence big enough to make the Poisson riskier, or does the Negative Binomial keep its title? We are left with this interrogation and thus, we cannot make any conclusion for the Poisson model with common shock versus the Negative Binomial with common component.

To make this more clear, we consider the case where a book of business consists in two classes of business. Either for the Negative Binomial with common component or for the Poisson model with common shock, we have:

$$\begin{aligned} W^{(1)} &= \sum_{i=1}^{N^{(11)}} X_i^{(1)} + \sum_{i=N^{(11)}+1}^{N^{(11)}+N^{(12)}} X_i^{(1)}, \\ W^{(2)} &= \sum_{i=1}^{N^{(22)}} X_i^{(2)} + \sum_{i=N^{(22)}+1}^{N^{(22)}+N^{(12)}} X_i^{(2)}. \end{aligned}$$

The reason why we cannot make any conclusion comes from the fact that for the same coeŒcient of correlation $\rho\left(N^{(1)}, N^{(2)}\right)$, the Negative Binomial and the Poisson models do not have the same mean for the random variable of the number of common shocks, $N^{(12)}$. Hence, when $E\left(N^{(12)}\right)$ for the Poisson is smaller than $E\left(N^{(12)}\right)$ for the Negative Binomial, then $E\left(N^{(11)}\right)$ and $E\left(N^{(22)}\right)$ are bigger for the Poisson than for the Negative Binomial, and vice versa. This thus means that there is always a term of the aggregate claim for which we cannot conclude anything for the stop-loss order.

As an example, consider the numerical example of Cossette and Marceau (2000). For $\rho\left(N^{(1)}, N^{(2)}\right) = 0.25$, we have to compare two sums of risks where the number or risks is distributed as a Poisson(3) and one sum having a Poisson(1) distribution, with two sums of risks where the number of risks is distributed as a Negative Binomial$(0.6875, 4)$ and one sum having a Negative Binomial$(0.3125, 4)$. Since the Poisson(1) has mean 1 and the Negative Binomial$(0.3125, 4)$ has mean 1.25, then it follows from Example 1.3 and Theorem 1.5 that the Negative Binomial is riskier than the Poisson. However, since the Poisson(3) has mean 3 and the Negative Binomial$(0.6875, 4)$ has mean 2.75, then we are not allowed to conclude that one sum of risks is riskier than the other, and then we cannot generalize the previous result for this dependent case. Note that we have only to consider the discrete distributions of the number of claims, since the severity distributions are the same in both cases.

According to the results obtained in the example of Cossette and Marceau (2000), it is evident that the Negative Binomial generates a riskier portfolio, but we did not ønd any generalization of this result in terms of stop-loss order. However, we believe that with a little work, it is probably possible to show this result for isolate cases. For the moment, we can only explain this phenomenon by the fact that the Negative Binomial model produces a bigger correlation coeŒcient $\rho\left(W^{(1)}, W^{(2)}\right)$ than the Poisson model for the same $\rho\left(N^{(1)}, N^{(2)}\right)$.

We are now interested in the biggest stop-loss premiums possible for the Poisson and the Negative Binomial models, for given marginals and for a given retention level $d \geq 0$. We saw in the previous report that every sum of risks is smaller in stop-loss order than the sum of the comonotonic version of these risks. Since stop-loss order is directly related to stop-loss premiums, then we know that the biggest stop-loss premium for a given retention level will arise for comonotonic risks. In a Poisson model with common shock and a Negative Binomial with common component, we know that the severity distributions are independent, and that the relation of dependence is between the random variables for the number of claims. Hence, we expect (even if we do not know how to prove it) that a comonotonic version of these random variables would generate the biggest stop-loss premiums. If the counting variables are comonotonic, this means that all the claims are common to all the classes of business. In other words, the only random variable generating the number of claims would be the same for every class (i.e. the variable of the common shock).

For both models in the example of Cossette and Marceau (2000), we computed the stop-loss premiums for many values of λ_{12} and α_0, for a retention level of $d = 10$. These values appear in Table 1 and 2, and as expected, the biggest stop-loss premium is when the parameter for the counting distribution of the common shock is maximum. Also, all the values obtained for the Negative Binomial model are bigger than the values for the Poisson model.

λ_{12}	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
SLP	1.83	1.85	1.87	1.89	1.91	1.93	1.94	1.96	1.98
λ_{12}	2.25	2.50	2.75	3.00	3.25	3.50	3.75	4.00	
SLP	2.00	2.02	2.03	2.05	2.07	2.09	2.10	2.12	

Table 1. Stop-loss premiums for the Poisson common shock model (deductible of 10)

α_0	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50
SLP	2.78	2.82	2.85	2.88	2.91	2.94	2.97	3.00	3.04	3.07	3.10
α_0	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00	
SLP	3.13	3.16	3.20	3.23	3.26	3.29	3.32	3.35	3.39	3.42	

Table 2. Stop-loss premiums for the NB common component model (deductible of 10)

31. Numerical Illustration

After having simulated the results of Cossette and Marceau (2000), we now reproduce the results using the theoretical methods presented above. We ørst discretize the severity distributions using the method of rounding (mass dispersal) presented in Klugman et al. (1998), with intervals of length 1. With the FFT function, we ønd the characteristic functions of $\phi_{X^{(1)}}$ and $\phi_{X^{(2)}}$, and using (73) and (74), we then compute the characteristic functions $\phi_{\widetilde{W}}$. We invert them with the inverse FFT function to ønd the discrete distribution of W, $F_{\widetilde{W}}$. We ønally use the recursive method given in Theorem 2.2 in order to ønd the ruin probabilities for given initial surplus. The SPLUS function used to perform these operations is presented in appendix.

We present in Table 3 the ruin probabilities over periods 1 to 20 for many values of the initial surplus u, that is $\psi(u, 1, 20)$, for the Poisson model with common shock. The ørst column presents the independent case, while the second and third columns present the case where $\rho\left(N^{(1)}, N^{(2)}\right)$ is equal to 0.25 and 0.75, respectively. The results obtained diœer a little from the results obtained by the authors. This diœerence does not seem too big, either over the diœerent initial surplus or the diœerent correlation structures, and this might be due to the method and the length of the intervals used to discretize the distribution functions of $X^{(1)}$ and $X^{(2)}$.

Similarly, we present in Table 4 the ruin probabilities over periods 1 to 20 for many values of the initial surplus u, that is $\psi(u, 1, 20)$, for the Negative Binomial model with common component. The ørst column presents the independent case, while the second and third columns present the case where $\rho\left(N^{(1)}, N^{(2)}\right)$ is equal to 0.25 and 0.75, respectively. Again, the results obtained diœer a little from the results obtained by the authors. This diœerence does not seem too big, either over the diœerent initial surplus or the diœerent correlation structures, and this might be due to the method and the length of the intervals used to discretize the distribution functions of $X^{(1)}$ and $X^{(2)}$.

u	$\psi(u,1,20,0)$	$\psi(u,1,20,0.25)$	$\psi(u,1,20,0.75)$
0	0.6213	0.6286	0.6418
10	0.3431	0.3564	0.3806
20	0.1782	0.1894	0.2105
30	0.0918	0.0995	0.1147
40	0.0466	0.0515	0.0614
50	0.0234	0.0262	0.0323
60	0.0116	0.0132	0.0167
70	0.0057	0.0066	0.0085
80	0.0028	0.0033	0.0043
90	0.0014	0.0016	0.0022
100	0.0007	0.0008	0.0011
110	0.0003	0.0004	0.0005
120	0.0002	0.0002	0.0003
130	0.0001	0.0001	0.0001
140	0.0000	0.0001	0.0001
150	0.0000	0.0000	0.0000

Table 3. Ruin probabilities $\psi(u,1,20)$ for the Poisson model

u	$\psi(u,1,20,0)$	$\psi(u,1,20,0.25)$	$\psi(u,1,20,0.75)$
0	0.6787	0.6821	0.6905
10	0.4764	0.4944	0.5263
20	0.3159	0.3417	0.3856
30	0.2059	0.2329	0.2790
40	0.1318	0.1564	0.1993
50	0.0829	0.1034	0.1406
60	0.0513	0.0674	0.0979
70	0.0312	0.0433	0.0674
80	0.0187	0.0275	0.0459
90	0.0110	0.0172	0.0309
100	0.0064	0.0106	0.0206
110	0.0037	0.0065	0.0136
120	0.0021	0.0039	0.0088
130	0.0012	0.0024	0.0057
140	0.0007	0.0014	0.0037
150	0.0004	0.0008	0.0023

Table 4. Ruin probabilities $\psi(u,1,20)$ for the NB model

Conclusion

We have presented a set of invariance properties for stop-loss order that reveal to be very useful in risk theory. We also discussed a method for computing ruin probabilities, by another way than simulation. We talked about the link between stop-loss order and ruin probabilities, but we did not ønd any stop-loss order between the Poisson model with common shock and the Negative Binomial model with common component for another case than independence (i.e. when there is no common shock!). However, spending more time on this problem may bring us to ønd what are the necessary conditions to get a stop-loss order between these models, and then justify why we obtained uniformly bigger ruin probabilities for the Negative Binomial model in the numerical example considered. This could be a good subject for ending the term in case we have some time left. Finally, after having simulated the ruin probabilities of the example of Cossette and Marceau (2000), we computed them using the approximation method presented.

Bibliography

[1] Bowers, N.L., Gerber, H.U., Hickman, J.C., Jones, D.A., and C.J. Nesbitt (1997). Actuarial Mathematics. Schaumburg, Ill.: Society of Actuaries.

[2] Cossette, H., Marceau, E. (2000). The Discrete-time Risk Model with Correlated Classes of Business. Insurance: Mathematics and Economics 26, 133-149.

[3] Denuit, M., Dhaene, J., Ribas, C. (2001). Does Positive Dependence Between Individual Risks Increase Stop-Loss Premiums? Insurance: Mathematics and Economics 28, 305-308.

[4] Dhaene, J., Goovaerts, M. (1996). Dependency of risks and stop-loss order. ASTIN Bulletin 26 (2), 201-212.

[5] Dhaene, J., Denuit, M. (1999). The Safest Dependence Structure Among Risks. Insurance: Mathematics and Economics 25, 11-21.

[6] Joe, H. (1997). Multivariate Models and Dependence Concepts. Chapman & Hall, London.

[7] Kaas, R., van Heerwaarden, A.E., Goovaerts, M.J., (1994). Ordering of Actuarial Risks. CAIRE, Brussels.

[8] Klugman, S.A., Panjer, H.H., Willmot, G.E. (1998). Loss Models: From Data to Decisions. Wiley, New York.

[9] Wang, S. (1996). Premium calculation by transforming the layer premium density. ASTIN Bulletin 26, 71-92.

[10] Wang, S., Dhaene, J. (1998). Comonotonicity, Correlation Order and Premium Principles. Insurance: Mathematics and Economics 22, 235-242.

Appendix

32. Proofs and Exercises

We provide in this section the proofs of some results presented in the second report. In a concern of clarity, we inserted the majority of the results for which we present the proofs. However, these results are only quoted and for more details about the notation and the terms used, please refer to the previous report.

Note that $\mathcal{R}_2(F_X, F_Y)$ refers to the set of all possible marginal distributions for (X, Y), and similarly, we use $\mathcal{R}_n(F_{X_1}, ..., F_{X_n})$ to refer to the set of all possible marginal distributions for $(X_1, ..., X_n)$.

32.1. Correlation Order.

Definition 32.1. Let $(X_1, Y_1), (X_2, Y_2)$ be two elements of $\mathcal{R}_2(F_X, F_Y)$. We say that (X_1, Y_1) is less correlated than (X_2, Y_2), written $(X_1, Y_1) \leq_{corr} (X_2, Y_2)$, if either of the following equivalent conditions holds:

1. For all non-decreasing functions f and g for which the covariances exist,

$$Cov(f(X_1), g(Y_1)) \leq Cov(f(X_2), g(Y_2)).$$

2. For all $x, y \geq 0$, the following inequality holds:

$$F_{X_1,Y_1}(x, y) \leq F_{X_2,Y_2}(x, y).$$

Proof. We want to prove that $1 \Leftrightarrow 2$. We ørst consider $1 \Rightarrow 2$, and then $2 \Rightarrow 1$.

We assume that (1) holds. Let f be the following indicator function

$$f(x) = I(x > x_1) = \begin{cases} 0, & x \leq x_1 \\ 1, & x > x_1 \end{cases},$$

and similarly, let the function g be

$$g(y) = I(y > y_1) = \begin{cases} 0, & y \leq y_1 \\ 1, & y > y_1 \end{cases}.$$

Then, by applying the deønition of the covariance, we ønd that

$$\begin{aligned} & E\left(I\left(X_1>x_1, Y_1>y_1\right)\right)-E\left(I\left(X_1>x_1\right)\right) E\left(I\left(Y_1>y_1\right)\right) \\ \leq\ & E\left(I\left(X_2>x_1, Y_2>y_1\right)\right)-E\left(I\left(X_2>x_1\right)\right) E\left(I\left(Y_2>y_1\right)\right). \end{aligned}$$

Since the pairs (X_1, X_2) and (Y_1, Y_2) are identically distributed, the similar terms cancel out, giving

$$E\left(I\left(X_1>x_1, Y_1>y_1\right)\right) \leq E\left(I\left(X_2>x_1, Y_2>y_1\right)\right).$$

Note that this last term does not cancel out, even if (X_1, X_2) and (Y_1, Y_2) have respectively the same marginals, because their relation of dependence is not necessarily the same.

It is well-known that the expectation of indicator functions can be expressed in terms of probabilities:

$$\begin{aligned} E\left(f\left(X\right)\right) &= E\left(I\left(X>x_1\right)\right) \\ &= 0 \operatorname{Pr}\left(X \leq x_1\right)+1 \operatorname{Pr}\left(X>x_1\right) \\ &= \operatorname{Pr}\left(X>x_1\right). \end{aligned}$$

We can then write, equivalently,

$$\operatorname{Pr}\left(X_1>x_1, Y_1>y_1\right) \leq \operatorname{Pr}\left(X_2>x_1, Y_2>y_1\right),$$

or

$$S_{X_1, Y_1}\left(x_1, y_1\right) \leq S_{X_2, Y_2}\left(x_1, y_1\right).$$

From the fact that $S_{X,Y}(x, y)=1-F_X(x)-F_Y(y)+F_{X,Y}(x, y)$, we ønd

$$\begin{aligned} & 1-F_{X_1}\left(x_1\right)-F_{Y_1}\left(y_1\right)+F_{X_1, Y_1}\left(x_1, y_1\right) \\ \leq\ & 1-F_{X_2}\left(x_1\right)-F_{Y_2}\left(y_1\right)+F_{X_2, Y_2}\left(x_1, y_1\right), \end{aligned}$$

and since the marginals are the same for X_1, X_2 and for Y_1, Y_2, then it reduces to

$$F_{X_1, Y_1}\left(x_1, y_1\right) \leq F_{X_2, Y_2}(x, y).$$

Since this relation is veriøed for the indicator functions, it then follows that it is true for any non-decreasing functions f and g for which the covariances exist. This is explained by the fact that all the functions f and g can be approximated by indicator functions.

We now suppose that (2) holds. It necessarily follows that for non-decreasing functions f and g,

$$\operatorname{Pr}\left(f\left(X_1\right) \leq x_1, g\left(Y_1\right) \leq y_1\right) \leq \operatorname{Pr}\left(f\left(X_2\right) \leq x_1, g\left(Y_2\right) \leq y_1\right),$$

for all $x_1, y_1 \geq 0$, since we have the same functions f and g applied respectively to the random variables with the same marginal distributions on both sides of the inequality, which let the relation unchanged (this is an one-to-one transformation).

From Dhaene and Goovaerts (1996), we have the following result:

$$Cov\left(X,Y\right)=\int_0^\infty\int_0^\infty\left(F_{X,Y}\left(u,v\right)-F_X\left(u\right)F_Y\left(v\right)\right)dudv,$$

for any $(X,Y)\in\mathcal{R}_2\left(F_X,F_Y\right)$. We then have

$$\begin{aligned} & \Pr\left(f\left(X_1\right)\leq x_1, g\left(Y_1\right)\leq y_1\right)-\Pr\left(f\left(X_1\right)\leq x_1\right)\Pr\left(g\left(Y_1\right)\leq y_1\right) \\ \leq\ & \Pr\left(f\left(X_2\right)\leq x_1, g\left(Y_2\right)\leq y_1\right)-\Pr\left(f\left(X_2\right)\leq x_1\right)\Pr\left(g\left(Y_2\right)\leq y_1\right), \end{aligned}$$

and if we take the double integral on both sides, we get

$$\begin{aligned} & Cov\left(f\left(X_1\right),g\left(Y_1\right)\right) \\ = & \int_0^\infty\int_0^\infty\left(\Pr\left(f\left(X_1\right)\leq x_1, g\left(Y_1\right)\leq y_1\right)\right. \\ & \left.-\Pr\left(f\left(X_1\right)\leq x_1\right)\Pr\left(g\left(Y_1\right)\leq y_1\right)dx_1dy_1\right. \\ \leq & \int_0^\infty\int_0^\infty\left(\Pr\left(f\left(X_2\right)\leq x_1, g\left(Y_2\right)\leq y_1\right)\right. \\ & \left.-\Pr\left(f\left(X_2\right)\leq x_1\right)\Pr\left(g\left(Y_2\right)\leq y_1\right)dx_1dy_1\right. \\ = & Cov\left(f\left(X_2\right),g\left(Y_2\right)\right), \end{aligned}$$

which proves that

$$Cov\left(f\left(X_1\right),g\left(Y_1\right)\right)\leq Cov\left(f\left(X_2\right),g\left(Y_2\right)\right).$$

∎

32.2. Link Between Stop-Loss Order and Correlation Order.

Theorem 32.1. Let (X_1,Y_1) and (X_2,Y_2) be elements of $\mathcal{R}_2\left(F_X,F_Y\right)$. If

$$\left(X_1,Y_1\right)\leq_{corr}\left(X_2,Y_2\right),$$

then

$$X_1+Y_1\leq_{sl}X_2+Y_2.$$

Proof. We know that X_1+Y_1 smaller in stop-loss order than X_2+Y_2 means that for all $d\geq 0$,

$$E\left(X_1+Y_1-d\right)_+\leq E\left(X_2+Y_2-d\right)_+.$$

We have that

$$E\left(X+Y-d\right)_+=E\left(X\right)+E\left(Y\right)-d+E\left(d-X-Y\right)_+, \qquad (75)$$

since when $E(X) + E(Y) - d \geq 0$, then $E(d - X - Y) \leq 0$, leading to $E(d - X - Y)_+ = 0$ and when $E(X) + E(Y) - d = -\alpha < 0$, then $E(d - X - Y)_+ = \alpha > 0$, resulting in $E(X) + E(Y) - d + E(d - X - Y)_+ = 0$.

For non-negative real numbers x and y, we can express $(d - x - y)_+$ as

$$(d - x - y)_+ = \int_0^d I(x \leq u, y \leq d - u)\, du. \tag{76}$$

For instance, if we have $d = 6$, $x = 2$, $y = 1$, then

$$\begin{aligned}
(6 - 2 - 1)_+ &= \int_0^6 I(1 \leq u, 2 \leq 6 - u)\, du \\
&= \int_0^1 0du + \int_1^4 du + \int_4^6 0du \\
&= 3 \\
&= \max(0, 6 - 2 - 1).
\end{aligned}$$

From (76) we have

$$E(d - X - Y)_+ = \int_0^d E(I(X \leq u, Y \leq d - u))\, du.$$

Since the expectation of the indicator function of an event is equivalent to the probability of this event, we then have

$$\begin{aligned}
E(d - X - Y)_+ &= \int_0^d \Pr(X \leq u, Y \leq d - u)\, du \\
&= \int_0^d F_{X,Y}(u, d - u)\, du,
\end{aligned}$$

leading to

$$E(X + Y - d)_+ = E(X) + E(Y) - d + \int_0^d F_{X,Y}(u, d - u)\, du. \tag{77}$$

From Deønition 5.1, we know that (X_1, Y_1) being less correlated than (X_2, Y_2) is equivalent to say that $F_{X_1,Y_1}(x, y) \leq F_{X_2,Y_2}(x, y)$. The following relation is then veriøed

$$\int_0^d F_{X_1,Y_1}(x, d - x)\, dx \leq \int_0^d F_{X_2,Y_2}(x, d - x)\, dx,$$

and adding some terms on both sides of the inequality leads to

$$\begin{aligned} & E(X_1) + E(Y_1) - d + \int_0^d F_{X_1,Y_1}(x, d-x)\,dx \\ \leq \; & E(X_2) + E(Y_2) - d + \int_0^d F_{X_2,Y_2}(x, d-x)\,dx. \end{aligned}$$

Since X_1 and X_2 have the same marginals, as Y_1 and Y_2, then $E(X_1) + E(Y_1) - d = E(X_2) + E(Y_2) - d$, and from (77), it follows that

$$E(X_1 + Y_1 - d)_+ \leq E(X_2 + Y_2 - d)_+,$$

which means, by deønition, that the sum $X_1 + Y_1$ is smaller in stop-loss order than the sum $X_2 + Y_2$. ∎

32.3. Related Results. Note that the proof of Theorem 1.2 from the previous report directly follows from Theorem 5.1, along with the notion of Frchet bounds. It is a fact that risks attaining the upper Frchet bound are comonotonic, and by the deønition of comonotonicity, a pair of this kind of risks can be written as $\left(F_X^{-1}(U), F_Y^{-1}(U)\right)$, where U is uniformly distributed on $(0, 1)$. It is also a fact that we can express a pair of mutually exclusive risks, i.e. risks attaining the lower Frchet bound, as $\left(F_X^{-1}(U), F_Y^{-1}(1-U)\right)$. From Deønition 5.1, saying that

$$F_{X,Y}^{me}(x, y) \leq F_{X,Y}(x, y) \leq F_{X,Y}^{c}(x, y),$$

where $F_{X,Y}^{c}$ and $F_{X,Y}^{me}$ stand for the comonotonic and the mutually exclusive versions of $F_{X,Y}$ respectively, is equivalent to say that

$$\left(F_X^{-1}(U), F_Y^{-1}(1-U)\right) \leq_{corr} (X, Y) \leq_{corr} \left(F_X^{-1}(U), F_Y^{-1}(U)\right),$$

which in turn, is equivalent by Theorem 5.1 to

$$F_X^{-1}(U) + F_Y^{-1}(1-U) \leq_{sl} X + Y \leq_{sl} F_X^{-1}(U) + F_Y^{-1}(U).$$

This thus proves Theorem 1.2 from previous report.

Theorem 2.1 from previous report is also pretty straight-forward to prove. We know that π is a premium principle that preserves stop-loss order, and we also know that $(X_1, Y_1) \leq_{corr} (X_2, Y_2)$. Then, by Theorem 5.1, we have that $X_1 + Y_1 \leq_{sl} X_2 + Y_2$, which is by deønition

$$E(X_1 + Y_1 - d)_+ \leq E(X_2 + Y_2 - d)_+,$$

for all $d \geq 0$, and since π preserves stop-loss order, then the relation

$$\pi(X_1 + Y_1) \leq \pi(X_2 + Y_2),$$

has to be true, which proves Theorem 2.1 of the previous report.

Corollary 2.2 from the previous report is found by simply inserting the result of Theorem 1.2 from the previous report, i.e. the special cases for Frchet bounds, into Theorem 2.1 from the previous report. This is exactly the same reasoning than for the proof of Theorem 2.1 from the previous report, but in considering comonotonic and mutually exclusive risks instead of general risks.

32.4. Wang's Premium Principle.

Theorem 32.2. Wang's premium principle preserves stop-loss order, i.e.

$$X \leq_{sl} Y \Rightarrow H_g(X) \leq H_g(Y).$$

Moreover, it is additive in the class of comonotonic risks,

$$H_g(X+Y) = H_g(X) + H_g(Y),$$

for comonotonic risks X and Y.

Proof. By deønition, $X \leq_{sl} Y$ means

$$E(X-d)_+ \leq E(Y-d)_+,$$

for all $d \geq 0$. We can express this as

$$E(X-d)_+ = \int_d^\infty S_X(x)\,dx \leq \int_d^\infty S_Y(y)\,dy = E(Y-d)_+.$$

If we consider a non-decreasing and concave function g with $g(0) = 0$ and $g(1) = 1$, then we can apply the function g to S_X and S_Y without altering the relation, because of the non-decreasing property of g (this is an one-to-one transformation). We get

$$\begin{aligned} E(H_g(X)-d)_+ &= \int_d^\infty g(S_X(x))\,dx \\ &\leq \int_d^\infty g(S_Y(y))\,dy = E(H_g(Y)-d)_+, \end{aligned}$$

which proves the ørst part of the theorem.

For the second part, we need the following relation to ønd an expectation

$$E(X) = \int_0^\infty S_X(x)\,dx.$$

We should recall that the expectation is the area under the curve of S_X. The previous formula has been determined by ȷslicing verticallyȷ, i.e. by adding all the vertical slices of width dx under the curve S_X. However, it is possible to modify this formula by considering the quantile axis instead of the x-axis, i.e. by ȷslicing horizontallyȷ. We can see this as adding all the horizontal slices for each diœerential dq under the curve S_X^{-1}. We then have

$$E(X) = \int_0^1 S_X^{-1}(q)\,dq. \tag{78}$$

For a non-decreasing function g with $g(0) = 0$ and $g(1) = 1$, it then follows that

$$H(X) = \int_0^\infty g(S_X(x))\,dx = \int_0^1 S_X^{-1}(q)\,dg(q).$$

This comes from the fact that

$$g(S_X(x)) = g(q) = q', \tag{79}$$

where q represents the quantile of the function $S_X(x)$ and q' represents the quantile of the function $g(S_X(x))$. Note that the range of this last function goes from 0 to 1. Since the survival function takes values between 0 and 1, and since g is non-decreasing with $g(0) = 0$ and $g(1) = 1$, then this function cannot take values outside from $(0, 1)$. We can transform (79) to obtain

$$S_X(x) = g^{-1}(q'),$$

and by isolating x, we have

$$x = S_X^{-1}\left(g^{-1}(q')\right).$$

From (78), we then have

$$\begin{aligned} H_g(X) &= \int_0^\infty g(S_X(x))\,dx \\ &= \int_0^1 S_X^{-1}\left(g^{-1}(q')\right)dq', \end{aligned}$$

and since we know from (79) that $g(q) = q' \Rightarrow q = g^{-1}(q')$, it follows that

$$H_g(X) = \int_0^1 S_X^{-1}(q)\,dg(q).$$

Wang (1996) presents a result stating that for two comonotonic risks X and Y, the following relation holds:

$$S_X^{-1}(q) + S_Y^{-1}(q) = S_{X+Y}^{-1}(q),$$

for $0 \leq q \leq 1$. It thus directly follows that

$$\begin{aligned} H_g(X+Y) &= \int_0^1 S_{X+Y}^{-1}(q)\, dg(q) \\ &= \int_0^1 \left(S_X^{-1}(q) + S_Y^{-1}(q)\right) dg(q) \\ &= H_g(X) + H_g(Y), \end{aligned}$$

proving the second part of the theorem. ▮

32.5. Generalized Frchet Bounds. We now prove the result of generalized Frchet bounds, that is the Frchet bounds for a multivariate joint distribution. From this result, the joint distribution of $\mathbf{X}$ is subject to the following bounds:

$$\begin{aligned} \max\left\{\sum_{k=1}^{n} F_{X_k}(x_k) - (n-1), 0\right\} &\leq F_{\mathbf{X}}(\mathbf{x}) \\ &\leq \min\{F_{X_1}(x_1), ..., F_{X_n}(x_n)\}. \end{aligned}$$

Proof. To prove this result, we will use the set theory. For instance, we consider the marginal cdf $F_{X_k}(x_k)$ as the probability of the event $E_k = \{X_k \leq x_k\}$. In a similar reasoning, $F_{\mathbf{X}}(\mathbf{x})$ is the probability of an intersection of n events, $E_1 \cap ... \cap E_n$.

We ørst prove the right-hand side of the inequality. The probability of an intersection of events being always smaller than or equal to the probability of each of the individual events, we have

$$\Pr(E_1 \cap ... \cap E_n) \leq \Pr(E_k),$$

for $k = 1, 2, ..., n$. The probability of the intersection must then be smaller than or equal to the smallest probability among the individual events. It then follows that

$$\begin{aligned} F_{\mathbf{X}}(\mathbf{x}) &= \Pr(E_1 \cap ... \cap E_n) \\ &\leq \min(\Pr(E_1), ..., \Pr(E_n)) = \min\{F_{X_1}(x_1), ..., F_{X_n}(x_n)\}, \end{aligned}$$

and the result for the upper bound is veriøed.

We now consider the left-hand side of the inequality. Using the DeMorgan's law of sets, the events

$$(E_1 \cap ... \cap E_n)^c = E_1^c \cup ... \cup E_n^c,$$

are equivalent, where the superscript c denotes the complement of the event. Thus, we see that

$$\begin{aligned} \Pr(E_1 \cap ... \cap E_n) &= 1 - \Pr(E_1 \cap ... \cap E_n)^c \\ &= 1 - \Pr(E_1^c \cup ... \cup E_n^c), \end{aligned}$$

and since the sum of the probabilities of individual events has a probability at least as big as the probability of the union of these events, we can write

$$\begin{aligned} 1-\Pr\left(E_1^c\cup...\cup E_n^c\right) &\geq 1-\sum_{k=1}^{n}\Pr\left(E_k^c\right) \\ &= 1-\sum_{k=1}^{n}\left(1-\Pr\left(E_k\right)\right) \\ &= 1-n+\sum_{k=1}^{n}\Pr\left(E_k\right). \end{aligned}$$

We then have

$$\begin{aligned} F_{\mathbf{X}}\left(\mathbf{x}\right) &= \Pr\left(E_1\cap...\cap E_n\right) \\ &\geq 1-n+\sum_{k=1}^{n}\Pr\left(E_k\right)=\sum_{k=1}^{n}F_{X_k}\left(x_k\right)-\left(n-1\right), \end{aligned}$$

and because probabilities have to be non-negative, we have

$$F_{\mathbf{X}}\left(\mathbf{x}\right)\geq\max\left\{\sum_{k=1}^{n}F_{X_k}\left(x_k\right)-\left(n-1\right),0\right\}. \tag{80}$$

∎

32.6. Cdf for Lower Frchet Bound.

Theorem 32.3. A necessary and suŒcient condition for the lower Frchet bound to be a cdf in $\mathcal{R}_n\left(F_{X_1},...,F_{X_n}\right)$ is that either

1. $\sum_{j=1}^{n}F_{X_j}\left(x_j\right)\leq 1$ for all $\mathbf{x}\in\mathbb{R}^n$ with $0<F_{X_j}\left(x_j\right)<1, j=1,...,n$; or
2. $\sum_{j=1}^{n}F_{X_j}\left(x_j\right)\geq n-1$ for all $\mathbf{x}\in\mathbb{R}^n$ with $0<F_{X_j}\left(x_j\right)<1, j=1,...,n$.

Proof. We want to prove that condition (1) and (2) are suŒcient conditions for the lower Frchet bound to be a proper multivariate cdf. However, the conditions that a function must satisfy in order to be a multivariate cdf diœer from the univariate case. According to Joe (1997), we can prove the suŒcient condition by verifying the rectangle inequality, i.e. for all $\left(a_1,...,a_m\right),\left(b_1,...,b_m\right)$ with $a_i<b_i$, $i=1,...,m$,

$$\sum_{i_1=1}^{2}...\sum_{i_m=1}^{2}\left(-1\right)^{i_1+...+i_m}F\left(x_{1i_1},...,x_{mi_m}\right)\geq 0, \tag{81}$$

where $x_{j1}=a_j$ and $x_{j2}=b_j$.

Let $x_{j1} < x_{j2}$, $p_{j1} = F(x_{j1})$, and $p_{j2} = F(x_{j2})$, $j = 1, ..., m$. Let $(y)_+ = \max\{0, y\}$, as usual. The rectangle condition for the lower Frchet bound leads to

$$\sum_{i_1=1}^{2} ... \sum_{i_m=1}^{2} (-1)^{i_1+...+i_m} \left[\sum_{j=1}^{m} p_{ji_j} - (m-1)\right] \geq 0. \qquad (82)$$

This equation means that we add the term $\pm\left[\sum_{j=1}^{m} p_{ji_j} - (m-1)\right]$ for all possible combinations of the p_{j1} and p_{j2}, $j = 1, ..., m$, the sign of each term being positive when the number of p_{j1} in it is even, and negative when the number of p_{j1} is odd.

To prove the ørst case, assume that $(x_{11}, ..., x_{m1})$ satisøes condition (1). If a term contains two probabilities that are less than one together, then this term takes the value 0. For instance, consider the case of m components in a term and two of these components are less than 1 together. Since the sum of the $m-2$ remaining terms is at most $m-2$, then the sum of all m components cannot be greater than $m-1$, and then the term $\left[\sum_{j=1}^{m} p_{ji_j} - (m-1)\right]$ is 0. This means that the only non-zero terms contain at most one of the p_{j1}'s, $j = 1, ..., m$. Then, by eliminating the zero terms in (82), we get:

$$\begin{aligned}
&(p_{12} + p_{22} + ... + p_{m2} - (m-1))_+ \\
&- (p_{11} + p_{22} + ... + p_{m2} - (m-1))_+ \\
&- (p_{12} + p_{21} + p_{32} + ... + p_{m2} - (m-1))_+ \\
&- ... - (p_{12} + p_{22} + ... + p_{m-1,2} + p_{m1} - (m-1))_+ .
\end{aligned} \qquad (83)$$

If $p_{12} = p_{22} = ... = p_{m2} = 1$, then (83) becomes $1 - p_{11} - p_{21} - ... - p_{m1} \geq 0$, since $\sum_{j=1}^{m} p_{j1} \leq 1$. If $p_{12} = ... = p_{j-1,2} = p_{j+1,2} = ... = p_{m2} = 1$, and $p_{j2} < 1$, then (83) becomes $p_{j2} - p_{j1} \geq 0$, since $\sum_{j=1}^{m} p_{j1} \leq 1$. Finally, if at least two of the m probability in a term are less than one, then (83) is 0. The rectangle inequality is thus satisøed and condition (1) is suŒcient to have a proper cdf.

We now assume that $(x_{12}, ..., x_{m2})$ satisøes condition (2), in order to prove the second case. If $p_{j1} > 0$ for all j, then (82) becomes 0 since all of the terms are non-negative. If at most $m-2$ of the p_{j1} are zero, then (82) is zero because the signs $(-1)^{i_1+...+i_m}$ of p_{j1}, p_{j2} for the non-zero terms balance out for all j. If $p_{11} = ... = p_{j-1,1} = p_{j+1,1} = ... = p_{m1} = 0$ and $p_{j1} > 0$, for $j = 1, ..., m$, then (82) becomes $p_{j2} - p_{j1} \geq 0$. If $p_{11} = ... = p_{m1} = 0$, then (82) becomes $p_{12} + ... + p_{m2} - (m-1) \geq 0$. Hence, condition (2) is suŒcient to have a proper cdf. ∎

32.7. PCD Risks Versus PQD Risks. We now want to show that for $n = 2$, positive quadrant dependence is equivalent to positive cumulative dependence, that is $PQD \Leftrightarrow PCD$.

Suppose that the risks X_1 and X_2 are PQD. By deønition of PCD risks, we have $\mathcal{I} \subset \{1, 2\}$, and $S_{\mathcal{I}} = X_1 + X_2$. Risks are said to be PCD if for any $\mathcal{I}$ and $j \notin \mathcal{I}$, $S_{\mathcal{I}}$ and X_j are PQD. In our case, $S_1 = X_1$ and X_2 are PQD by assumption, as well as $S_2 = X_2$ and X_1, and the PCD condition is then veriøed.

Now assume that the risks X_1 and X_2 are PCD. By deønition, it means that $S_1 = X_1$ and X_2, as well as $S_2 = X_2$ and X_1, are PQD, which proves the necessary condition.

Alternatively, we can use Theorem 9.1 from the previous report to prove the suŒcient condition..

Suppose that the risks X_1 and X_2 are PQD. By deønition, we have

$$F_{X_1,X_2}(x_1, x_2) \geq F_{X_1}(x_1) F_{X_2}(x_2),$$

that is the risks are more correlated than independent risks, which is equivalent to

$$X_1, X_2 \geq_{corr} X_1^{\perp}, X_2^{\perp}.$$

From Theorem 5.1, we can write

$$X_1 + X_2 \geq_{sl} X_1^{\perp} + X_2^{\perp},$$

and from Theorem 9.1 from the previous report, it follows that the risks X_1 and X_2 are PCD.

Note that it is normal that $PQD \Leftrightarrow PCD$ for $n = 2$, since PCD is a notion that allows to extend PQD to more than just 2 dimensions!

32.8. Bounds of General Risks Versus PCD Risks. Consider the risks $X_1, ..., X_n$ with marginal distributions $F_{X_1}, ..., F_{X_n}$. We consider four versions of these risks, that is the mutually exclusive, independent, PCD and comonotonic versions.

For the comonotonic risks, let X_i be described by

$$X_i = F_{X_i}^{-1}(U), \ i = 1, ..., n,$$

where U is uniformly distributed on the interval $(0, 1)$.

For the PCD risks, let X_i be described by

$$X_i = B_i I, \ i = 1, ..., n,$$

where B_i, $i = 1, ..., n$, is distributed as an Exponential with mean 1, and where I has the following distribution:

$$\Pr(I = j) = \begin{cases} \frac{n}{n-1}, & j = 0 \\ \frac{1}{n}, & j = 1 \end{cases}.$$

For the independent and the mutually exclusive risks, let X_i be described by

$$X_i = B_i I_i, \ i = 1, ..., n,$$

where B_i, $i = 1, ..., n$, is distributed as an Exponential with mean 1, and where I_i, $i = 1, ..., n$, has the following distribution:

$$\Pr(I_i = j) = \begin{cases} \frac{n}{n-1}, & j = 0 \\ \frac{1}{n}, & j = 1 \end{cases}.$$

However, for the mutually exclusive version, we have that $I_1 + ... + I_n \equiv 1$.

We want to show that for mutually exclusive risks, condition (1) on p.15 of the previous report is satisøed. We have that

$$\begin{aligned} q_i &= 1 - F_{X_i}(0) = 1 - \Pr(B_i I_i \leq 0) \\ &= 1 - \Pr(B_i I_i \leq 0 \,|I_i = 0) \Pr(I_i = 0) \\ &\quad - \Pr(B_i I_i \leq 0 \,|I_i = 1) \Pr(I_i = 1) \\ &= 1 - (1)\left(\frac{n-1}{n}\right) - (0)\left(\frac{1}{n}\right) \\ &= 1 - \frac{n-1}{n} = \frac{n-n+1}{n} = \frac{1}{n}, \end{aligned}$$

for $i = 1, ..., n$, which leads to

$$\sum_{i=1}^{n} q_i = \sum_{i=1}^{n} \frac{1}{n} = 1 \leq 1,$$

satisfying the condition to have mutually exclusive risks.

We now compute the cdf of X_i, $i = 1, ..., n$. For $x_i \geq 0$, $i = 1, ..., n$, we have

$$\begin{aligned} \Pr(B_i I_i \leq x_i) &= \Pr(B_i I_i \leq x_i \,|I_i = 0) \Pr(I_i = 0) \\ &\quad + \Pr(B_i I_i \leq x_i \,|I_i = 1) \Pr(I_i = 1) \\ &= 1\left(\frac{n-1}{n}\right) + F_{B_i}(x_i)\left(\frac{1}{n}\right) \\ &= \frac{n-1}{n} + \frac{1-e^{-x_i}}{n} \\ &= \frac{n-e^{-x_i}}{n}. \end{aligned}$$

The distribution of X_i, $i = 1, ..., n$, is then given by

$$F_{X_i}(x_i) = \begin{cases} 0, & \frac{n-1}{n} \\ x_i, & \frac{n-e^{-x_i}}{n} \end{cases},$$

for $x_i \geq 0$, $i = 1, ..., n$.

The sum of the marginals is

$$\sum_{i=1}^{n} F_{X_i}(x_i) = \sum_{i=1}^{n} \frac{n-e^{-x_i}}{n} = \sum_{i=1}^{n}\left(1-\frac{e^{-x_i}}{n}\right)$$
$$= n - \frac{1}{n}\sum_{i=1}^{n} e^{-x_i},$$

and we obtain for the lower Frchet bound

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = \max\left\{\sum_{i=1}^{n} F_{X_i}(x_i) - (n-1), 0\right\}$$
$$= \max\left\{n - \frac{1}{n}\sum_{i=1}^{n} e^{-x_i} - (n-1), 0\right\}$$
$$= \max\left\{1 - \frac{1}{n}\sum_{i=1}^{n} e^{-x_i}, 0\right\}.$$

Since the expression $1-\frac{1}{n}\sum_{i=1}^{n} e^{-x_i}$ cannot be negative, we are left with

$$F_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = 1 - \frac{1}{n}\sum_{i=1}^{n} e^{-x_i},$$

for $x_i \geq 0$, $i = 1, \ldots, n$.

We now compute the joint cdf of $X_1^{\perp}, \ldots, X_n^{\perp}$, the independent version of the risks

$$F_{X_1^{\perp},\ldots,X_n^{\perp}}(x_1,\ldots,x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$$
$$= \prod_{i=1}^{n}\left(\frac{n-e^{-x_i}}{n}\right)$$
$$= \prod_{i=1}^{n}\left(1-\frac{e^{-x_i}}{n}\right).$$

It is diŒcult to quantify how much of an improvement represents the independent version of the risks over the lower Frchet bound in terms of lower bound. However, the existence of such an improvement is evident by looking at the cdf 's obtained.

Let

$$a_i = \frac{e^{-x_i}}{n}.$$

We then have for the independent case

$$\begin{aligned} F_{X_1^{\perp},...,X_n^{\perp}}(x_1,...,x_n) &= \prod_{i=1}^{n}\left(1-\frac{e^{-x_i}}{n}\right) \\ &= \prod_{i=1}^{n}(1-a_i). \end{aligned}$$

For $n=2$, we have

$$\begin{aligned} F_{X_1^{\perp},X_2^{\perp}}(x_1,x_2) &= (1-a_1)(1-a_2) \\ &= 1-a_1-a_2+a_1a_2, \end{aligned}$$

and for $n=3$, we obtain

$$\begin{aligned} F_{X_1^{\perp},...,X_n^{\perp}}(x_1,...,x_n) &= (1-a_1)(1-a_2)(1-a_3) \\ &= (1-a_1-a_2+a_1a_2)(1-a_3) \\ &= 1-a_1-a_2-a_3+a_1a_2+a_2a_3+a_1a_3-a_1a_2a_3. \end{aligned}$$

For the mutually exclusive version, we have respectively

$$F_{X_1,X_2}(x_1,x_2) = 1-a_1-a_2,$$

and

$$F_{X_1,X_2,X_3}(x_1,x_2,x_3) = 1-a_1-a_2-a_3.$$

Since the terms a_1a_2 and $(a_1a_2+a_2a_3+a_1a_3-a_1a_2a_3)$ are clearly bigger than 0 when $0 \leq a_i \leq 1$ for all i, then it is evident that

$$F_{X_1^{\perp},...,X_n^{\perp}}(x_1,...,x_n) \geq F_{X_1,...,X_n}(x_1,...,x_n),$$

and so the independent case is a tighter bound for PCD risks.

We can generalize this result for the case where we have n PCD risks, by using the proof by induction. We know that for $n=3$,

$$1-a_1-a_2-a_3 \leq (1-a_1)(1-a_2)(1-a_3).$$

Let assume that this is true for $n=k$:

$$1-a_1-...-a_k \leq (1-a_1)...(1-a_k).$$

We verify that it is true for $n=k+1$:

$$\begin{aligned} 1-a_1-...-a_k-a_{k+1} &\leq (1-a_1)...(1-a_k)-\alpha_{k+1} \\ &\leq (1-a_1)...(1-a_k)(1-\alpha_{k+1}), \end{aligned}$$

since if we let $\beta=(1-a_1)...(1-a_k)$, we necessarily have that

$$\beta-\alpha_{k+1} \leq \beta(1-\alpha_{k+1}) = \beta-\beta\alpha_{k+1},$$

since $0 \leq a_i, \beta \leq 1$, for all i. Then, it follows that

$$\begin{aligned} F_{X_1,...,X_n}(x_1, ..., x_n) &= 1 - a_1 - ... - a_n \\ &\leq (1 - a_1) ... (1 - a_n) = F_{X_1^{\perp},...,X_n^{\perp}}(x_1, ..., x_n), \end{aligned}$$

for all $n \geq 1$.

33. SPLUS Functions

33.1. Stop-Loss Premiums.

This function calculates the stop-loss(d) premiums

```
stoploss <- function(fs, d1, pas = 1)–
    psld <- c()
    s <- c(1:length(fs) - 1) * pas
    for(i in 1:length(d1)) –
        s1 <- ((s > d1[i]) * (d1[i] >= 0)) * (s - d1[i])
        psld <- c(psld, sum(s1 * fs))
    ”
    return(psld)
”
```

33.2. Ruin Probabilities.

This function calculates ruin probabilities for many values of the initial surplus, either for the Poisson with common shock or for the NB with common component. It takes as arguments the model (¡P¡ or ¡NB¡), the length of the interval for discretizing (pas), the parameters of each class of business (l1,l2,l0), the security margin (theta), and the periods of the ruin probabilities (p).

```
pruin <- function(cond = ¡P¡, pas = 1, l1 = 4, l2 = 4, l0 = 0,
theta = 0.15, p = 20)–
    l11 <- l1 - l0
    l22 <- l2 - l0
    l <- l11 + l22 + l0
    long <- 2ˆ15
#Discretize the severity distributions and calculate their characteristic function
    Fx1 <- pweibull((pas * (2 * (0:(2ˆ12 - 1)) + 1))/2, 0.5, 0.5625)
    fx1 <- diœ(c(0, Fx1, 1))
    fx1l <- c(fx1, rep(0, (long - length(fx1))))
    Mx1 <- œt(fx1l)
    Fx2 <- pexp((pas * (2 * (0:(2ˆ12 - 1)) + 1))/2, 1/1.125)
    fx2 <- diœ(c(0, Fx2, 1))
    fx2l <- c(fx2, rep(0, (long - length(fx2))))
    Mx2 <- œt(fx2l)
```

```
    #Compute the characteristic function of W
        if(cond == ȷPȷ) –
            Mx <- (l11 * Mx1 + l22 * Mx2 + l0 * Mx1 * Mx2)/l
            Ms <- exp(l * (Mx - 1))
        ″
        if(cond == ȷNBȷ) –
            A <- (1 - 4 * (Mx1 - 1))^( - l11) * (1 - 4 * (Mx2 - 1))^( -
l22)
            B <- (1 - 4 * (Mx1 - 1) - 4 * (Mx2 - 1))^( - l0)
            Ms <- A * B
        ″
    #Compute the premiums received by the insurer at each period
        EW <- (0.5625 * gamma(1 + 1/0.5) * 4 + 1.125 * 4) * (1 +
theta)
    #Calculate the discrete distribution of W
        fs <- Re(œt(Ms, inverse = T))[1:long]
        fs <- (fs >= 0) * fs
        fs <- fs/sum(fs)
        Fs <- cumsum(fs)
        Fs1 <- c()
        for(i in 0:((long - 2) * pas)) –
            Fs1 <- c(Fs1, Fs[i/pas + 1])
        ″
        fs1 <- diœ(c(0, Fs1))
    #Calculate the ruin probabilities over p periods for many values of
the initial surplus
        v <- 5000
        b11 <- c()
        for(k in 0:v) –
            b11 <- c(b11, sum(fs[1:((k + EW)/pas + 1)]))
        ″
        im1 <- b11
        b1i <- c()
        for(j in 2:20) –
            for(k in 0:v) –
                b1i <- c(b1i, sum(im1[1:(k + EW + 1)] * fs1[(k +
EW + 1):1]))
            ″
            im1 <- b1i
            b1i <- c()
        ″
        return(1 - im1[1:150])
```

"