Aggregation of Correlated Risk Portfolios: Models & Algorithms [1]

by Shaun S. Wang, Ph.D.

ABSTRACT

This report presents a set of tools for modeling and combining correlated risks. Various correlation structures are generated using copula, common mixture, component and distortion models. These correlation structures are specified in terms of (i) the joint cumulative distribution function or (ii) the joint characteristic function, and lend themselves to efficient methods of aggregation by using Monte Carlo simulation or direct Fourier inversion. For a set of correlated risks with arbitrary marginals and any (positive definite) matrix of correlation coefficients (or Kendall's tau), simple yet general methods are proposed for combining the correlated risks.

Acknowledgment: The author wishes to thank the CAS Committee on Theory of Risk for initiating and supporting this fascinating research project. A special thanks goes to Phil Heckman and Glenn Meyers for providing numerous stimulating comments while overseeing the progress of this project. The author also benefited from discussions with Ole Hesselager and editorial comments by Julia Wirch and Ken Seng Tan.

---

[1] Work performed under a research contract with the CAS Committee on Theory of Risk.

Aggregation of Correlated Risk Portfolios: Models & Algorithms

Table of Contents

PART I. BASIC CONCEPTS & TOOLS

PART II. COPULA & MONTE CARLO SIMULATION

PART III. COMMON MIXTURE & COMPONENT MODELS

PART IV. JOINT PGF CONSTRUCTION & FFT METHODS

Aggregation of Correlated Risk Portfolios: Models & Algorithms

by Shaun S. Wang, Ph.D.

1 Introduction

A good introduction for this research report is the original Request For Proposal (RFP) drafted by the CAS Committee on Theory of Risk. In the following paragraph, the original RFP is restated with minor modification.

> Aggregate loss distributions are probability distributions of the total dollar amount of loss under one or a block of insurance policies. They combine the separate effects of the underlying frequency and severity distributions. In the actuarial literature, a number of methods have been developed for modeling and computing the aggregate loss distributions, see Heckman & Meyers (1983), Panjer (1981) and Robertson (1992). The main issue underlying this research project is how to combine aggregate loss distributions for separate but correlated classes of business.
>
> Assume a book of business is the union of disjoint classes of business each of which has an aggregate distribution. These distributions may be given in many different ways. Among other ways, they may be specified parametrically, e.g. lognormal or transformed beta with given parameters; they may be given by specifying separate frequency and severity distributions; e.g. negative binomial frequency and pareto severity with given parameters. The classes of business are NOT independent. For this project, assume that we are given a correlation matrix (or some other easily obtainable measure of dependency) and that the correlation coefficients vary among different pairs of classes. The problem is how to calculate the aggregate loss distribution for the whole book.

In the traditional actuarial theory, individual risks are usually assumed to be independent, mainly because the mathematics for correlated risks is less tractable. The CAS recognizes the importance of modeling and combining correlated risks, and wishes to enhance the development of tools and models that improve the accuracy of the estimation of aggregate loss distributions for blocks of insurance risks. The modeling of dependent risks has special relevance to the current on-going project of Dynamic Financial Analysis.

In general, combining correlated loss variables requires knowledge of their joint (multivariate) distribution. However, the available data regarding the association between loss variables is often limited to some summary statistics (e.g. correlation matrix). In the special case of multivariate normal distributions, the covariance matrix and the mean vector, as summary statistics, completely specify the joint distribution. For general loss frequency or severity distributions, specific dependency models have to be used in conjunction with summary statistics.

When we move into the high dimension world of multivariate distributions, the variety of dependency structures dramatically increases. Given fixed marginal distributions and correlation matrix, we can construct infinitely many joint distributions. Ideally, models for dependency structure should be easy to implement and require relatively few input parameters. As well, the choice of the dependency model and its parameter values should reflect the underlying correlation-generating mechanism.

In developing dependency models, we are aiming at simple implementation by Monte Carlo simulation or by direct Fourier inversion. To this end, we will take the following approaches to modeling and combining correlated risks, which are organized in four parts:

Part I. As a starting point, we first review some basic measures of dependency including correlation coefficients and Kendall's tau. As well, we introduce some basic concepts including the joint cumulative distribution function and the joint probability generating function, which will form the basis of the whole paper.

Part II. In recognizing that the key to a simulation of correlated risks lies in the generation of multivariate uniform numbers, we investigate various correlation structures by using the concept of copulas (i.e. multivariate uniform distributions). In particular, we suggest the use of Cook-Johnson copula and the normal copula, as they lead to efficient simulation techniques.

Part III. With due consideration to the underlying correlation-generating mechanism, we will generate a variety of dependency models by using common mixtures and common shocks. These dependency models allow simple methods of aggregation by Monte Carlo simulation or by direct Fourier inversion.

Part IV. In some situations, although the marginal distributions and their correlation matrix are given, the joint distribution of correlated risks is not fully specified. This is mainly due to lack of multivariate data. For a wide variety of marginal distributions with a given correlation matrix, we propose a simple method of combining the correlated risks by Fast Fourier Transforms.

For the reader's convenience, an inventory of commonly used univariate distributions is given in Appendix A, this includes both discrete and continuous distributions. As a convention, we use X, Y and Z to represent any random variables (discrete, continuous or mixed), while using N and K to represent only discrete variables defined on non-negative integers.

PART I. BASIC CONCEPTS & TOOLS

2 Some Basics of Monte Carlo Simulation

Assume that X has a cumulative distribution function (cdf) F_X and a survivor function (s.f.) $S_X(x) = 1 - F_X(x)$. We define F_X^{-1} and S_X^{-1} as follows:

$$\begin{aligned} F_X^{-1}(q) &= \inf\{x : F_X(x) \geq q\}, \qquad 0 < q < 1 \\ S_X^{-1}(q) &= \inf\{x : S_X(x) \leq q\}, \qquad 0 < q < 1. \end{aligned}$$

Remark that F_X^{-1} is non-decreasing, S_X^{-1} is non-increasing and $S_X^{-1}(q) = F_X^{-1}(1-q)$.

The traditional Monte Carlo simulation method is based on the following result.

Lemma 2.1 *For any random variable X and any random variable U which is uniformly distributed on $(0,1)$, we have that X and $F_X^{-1}(U)$ have the same cdf.*

Proof: $\mathrm{P}\{F_X^{-1}(U) \leq x\} = \mathrm{P}\{U \leq F_X(x)\} = F_X(x)$. □

A Monte Carlo simulation of a random variable X can be achieved by first drawing a random uniform number u from $U \sim \text{Uniform}(0,1)$, and then inverting u by $x = F_X^{-1}(u)$.

In a similar way, a Monte Carlo simulation of k variables, $(X_1, \cdots, X_k)$, usually starts with k uniform random variables, $(U_1, \cdots, U_k)$. If the variables $(X_1, \cdots, X_k)$ are independent (correlated), then we need k independent (correlated) uniform random variables $(U_1, \cdots, U_k)$. For a set of given marginals, the correlation structure of the variables $(X_1, \cdots, X_k)$ is completely determined by the correlation structure of the uniform random variables, $(U_1, \cdots, U_k)$.

Definition 2.1 *A* **copula** *is defined as the joint cdf of k uniform random variables*

$$C(u_1, \cdots, u_k) = Pr\{U_1 \leq u_1, \cdots, U_k \leq u_k\}.$$

For any set of arbitrary marginal distributions, the formula

$$F_{X_1,\cdots,X_k}(x_1, \cdots, x_k) = C(F_{X_1}(x_1), \cdots, F_{X_k}(x_k)) \tag{2.1}$$

defines a joint cdf with marginal cdf's $F_{X_1}, \cdots, F_{X_k}$; The formula

$$S_{X_1,\cdots,X_k}(x_1, \cdots, x_k) = C(S_{X_1}(x_1), \cdots, S_{X_k}(x_k)) \tag{2.2}$$

defines a joint s.f. with marginal s.f. $S_{X_1}, \cdots, S_{X_k}$.

The multivariate distributions given by eq. (2.1) and eq. (2.2) are usually different, although they both have the same Kendall's tau as defined in section 3.4.

3 Measures of Association

3.1 Pearson correlation coefficients

For random variables X and Y, the **Pearson correlation coefficient**, defined by

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

always lies in the range $[-1,1]$. Note that $\rho(X,Y)=1$, if, and only if, $X = aY + b$ for some constants $a > 0$ and b. If the there is no linear relationship between X and Y, the permissible range of $\rho(X,Y)$ is further restricted.

Example 3.1 Consider the case that $\log X \sim N(\mu, 1)$ and $\log Y \sim N(\mu\,\sigma, \sigma^2)$. The maximum correlation between X and Y is obtained when the deterministic relation $Y = X^{\sigma}$ holds. Thus, for these fixed marginals we have [see Appendix A.4.2]

$$\max\{\rho(X,Y)\} = \frac{\exp(\sigma)-1}{\sqrt{\exp(\sigma^2)-1}\,\sqrt{e-1}}.$$

Observe that

- $\max\{\rho(X,Y)\} = 1$ when $\sigma = 1$ (i.e. $X = Y$)
- $\max\{\rho(X,Y)\}$ decreases to zero as σ increases to ∞
- $\max\{\rho(X,Y)\}$ decreases to $1/\sqrt{e-1}$ as σ decreases to 0.

For a set of k random variables $X_1, \cdots, X_k$, the correlation matrix

$$\begin{pmatrix} \rho(X_1,X_1) & \cdots & \rho(X_1,X_k) \\ \vdots & \vdots & \vdots \\ \rho(X_k,X_1) & \cdots & \rho(X_k,X_k) \end{pmatrix}, \qquad -1 \le \rho(X_i,X_j) \le 1,$$

is always positive definite, as it is symmetric and diagonally dominant.

3.2 Covariance coefficients

For non-negative random variables X and Y, we define the **covariance coefficient** as

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

Note that the permissible range of $\omega(X,Y)$ depends on the shape of the marginal distributions.

Example 3.2 Reconsider the variables X and Y in Example 3.1, it can be shown that

$$\max\{\omega(X,Y)\} = e^{\sigma} - 1.$$

Observe that

- $\max\{\omega(X,Y)\} = e-1$ when $\sigma = 1$ (i.e. $X=Y$)
- $\max\{\omega(X,Y)\}$ increases to infinity as σ increases to infinity
- $\max\{\omega(X,Y)\}$ decreases to zero as σ decreases to zero.

Example 3.3 Consider two negative binomial variables (see Appendix A.1) $N_1 \sim \mathrm{NB}(r_1, q_1)$ and $N_2 \sim \mathrm{NB}(r_2, q_2)$. It can be verified that

$$\omega(N_1, N_2) = \rho(N_1, N_2)\,\frac{1}{\sqrt{r_1 r_2}}\sqrt{\frac{1+q_1}{q_1}}\sqrt{\frac{1+q_2}{q_2}},$$

which decreases to zero as the product $r_1 r_2$ increases to infinity, while q_1 and q_2 are kept fixed.

For k non-negative random variables, $X_1, \cdots, X_K$, we define the matrix of covariance coefficients as

$$\begin{pmatrix} \omega(X_1,X_1) & \cdots & \omega(X_1,X_k) \\ \vdots & \vdots & \vdots \\ \omega(X_k,X_1) & \cdots & \omega(X_k,X_k) \end{pmatrix}.$$

Remark: One should exercise caution when choosing a parameter value for $\omega(X,Y)$, as its permissible range is sensitive to the marginal distributions. A practical method for obtaining the maximal positive and negative covariances between risks X and Y are given in the next sub-section by eq. (3.1) and eq. (3.2).

3.3 Fréchet bounds, co-monotonicity and maximal correlation

Now consider the bivariate random variables (X,Y). Let

$$F_{X,Y}(x,y) = \mathrm{P}\{X \le x,\ Y \le y\}, \quad S_{X,Y}(x,y) = \mathrm{P}\{X > x,\ Y > y\}$$

be the joint cdf and the joint s.f. of (X,Y), respectively. Note that

$$F_{X,Y}(x,\infty) = F_X(x), \qquad F_{X,Y}(\infty,y) = F_Y(y), \qquad \text{for } -\infty < x,\ y < \infty.$$

$$S_{X,Y}(x,y) = 1 - F_X(x) - F_Y(y) + F_{X,Y}(x,y) \neq 1 - F_{X,Y}(x,y).$$

If X and Y are independent, then $F_{X,Y}(x,y) = F_X(x){\cdot}F_Y(y)$ and $S_{X,Y}(x,y) = S_X(x){\cdot}S_Y(y)$. In general, the joint cdf $F(x,y)$ is constrained from above and below.

Lemma 3.1 *For any bivariate cdf $F_{X,Y}$ with given marginals F_X and F_Y, we have*

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

Proof: The first inequality results from the fact that $S(x,y) \ge 0$, and the second inequality can be proven using $P(A \cap B) \le \min[P(A), P(B)]$. □

The upper bound

$$F_u(x,y) = \min[F_X(x),\ F_Y(y)],$$

and the lower bound

$$F_l(x,y) = \max[F_X(x) + F_Y(y) - 1, 0],$$

are called the Fréchet bounds.

Closely associated with the Fréchet bounds is the concept of comonotonicity. The upper Fréchet bound is reached if X and Y are comonotonic. The lower Fréchet bound is reached if X and $-Y$ are comonotonic.

Definition 3.1 *Two random variables X and Y are* **comonotonic** *if there exists a random variable Z such that*

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

where the functions u, v are non-decreasing.

Recall that X and Y are positively perfectly correlated if, and only if, $Y = aX + b$, $a > 0$. This linear condition is quite restrictive. Co-monotonicity is an extension of the concept of perfect correlation to random variables with any arbitrary distributions. Consider

$$X = \begin{cases} Z, & Z \leq d \\ d, & Z > d, \end{cases} \qquad Y = \begin{cases} 0, & Z \leq d \\ Z - d, & Z > d. \end{cases}$$

Note that X and Y are *not* perfectly correlated since one cannot be written as a function of the other. However, since X and Y are always non-decreasing functions of the original risk Z, they are comonotonic. They are bets on the same event and neither of them is a *hedge* against the other.

The concept of co-monotonicity can also be explained in terms of Monte Carlo simulation by inversion of random uniform numbers. In order to simulate comonotonic risks, X and Y, the same sample of random uniform numbers can be used in an inversion by F_X and F_Y, respectively. By contrast, if X and Y are independent, two independent samples of random uniform numbers have to be used in an inversion by F_X and F_Y, respectively.

For given marginals F_X and F_Y, the maximal possible correlation exists when X and Y are comonotonic, in which case an approximation of the covariance can be obtained from

$$\mathrm{Cov}[X,Y] \approx \sum_{j=1}^{n} F_X^{-1}\left(\frac{j}{n+1}\right) F_Y^{-1}\left(\frac{j}{n+1}\right) - \mathrm{E}[X]\mathrm{E}[Y], \tag{3.1}$$

for some large number n. The maximal negative correlation exists when X and $-Y$ are comonotonic, in which case an approximation of the covariance can be obtained from

$$\mathrm{Cov}[X,Y] \approx \sum_{j=1}^{n} F_X^{-1}\left(\frac{j}{n+1}\right) F_Y^{-1}\left(1-\frac{j}{n+1}\right) - \mathrm{E}[X]\mathrm{E}[Y], \tag{3.2}$$

for some large number n.

3.4 Kendall's tau

As a measure of association, **Kendall's tau** is defined as

$$\tau = \tau(X,Y) = \Pr\{(X_2 - X_1)(Y_2 - Y_1) \geq 0\} - \Pr\{(X_2 - X_1)(Y_2 - Y_1) < 0\},$$

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

Kendall's tau has the following properties (e.g. Genest and Mackay, 1986):

- $-1 \leq \tau \leq 1$.
- If X and Y are comonotonic, then $\tau = 1$.
- If X and $-Y$ are comonotonic, then $\tau = -1$.
- If X and Y are independent, then $\tau = 0$.
- τ is invariant under strictly monotone transforms. That is, if f and g are strictly increasing (or decreasing) functions, then $\tau(f(X), g(Y)) = \tau(X,Y)$.
- $\tau(F_X(X), F_Y(Y)) = \tau(S_X(X), S_Y(Y)) = \tau(X,Y)$. Thus, Kendall's tau is often measured in terms of uniform random variables over $[0,1]\times[0,1]$.

Assume that we have available a random sample of bivariate observations, (X_i, Y_i), $i = 1, \cdots, k$. A non-parametric estimate of the Kendall's tau is

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

4 Probability Generating Functions & Characteristic Functions

4.1 Univariate case

Let X be a non-negative random variable of discrete, continuous or mixed type. Let $f_X(x)$ be the probability (density) function of X, i.e.

$$f_X(x) = \begin{cases} \Pr\{X = x\}, & \text{if } X \text{ is discrete} \\ \frac{d}{dx}F_X(x), & \text{if } X \text{ is continuous} \end{cases}$$

- The **probability generating function (pgf)** of X is defined by

$$P_X(t) = \mathrm{E}[t^X] = \begin{cases} \sum f_X(x)t^x & \text{if } X \text{ is discrete} \\ \int f_X(x)t^x dx & \text{if } X \text{ is continuous} \end{cases}$$

- The **moment generating function (mgf)** of X is defined by

$$M_X(t) = \mathrm{E}[e^{tX}] = P_X(e^t)$$

- The **characteristic function (ch.f.)**, also called Fourier transform, is defined by

$$\phi_X(t) = \mathrm{E}[e^{\mathrm{i}tX}] = P_X(e^{\mathrm{i}t}) = M_X(\mathrm{i}t),$$

  where $\mathrm{i} = \sqrt{-1}$ is the imaginary unit.
  Fast Fourier Transform (FFT) can be viewed as a discrete Fourier transform. For more details, one can consult the text of Klugman et. al. (1998).

- It holds that $P_X(1) = M_X(0) = \phi_X(0) = 1$ and

$$\mathrm{E}[X] = \left[\frac{d}{dt}P_X(t)\right]_{t=1} = \left[\frac{d}{dt}M_X(t)\right]_{t=0} = -\mathrm{i}\left[\frac{d}{dt}\phi_X(t)\right]_{t=0}.$$

4.2 Multivariate framework

For a set of random variables $(X_1, \cdots, X_k)$, let $f_{X_1,\cdots,X_k}$ be their **joint** probability (density) function, i.e.

$$f_{X_1,\cdots,X_k}(x_1,\cdots,x_k) = \begin{cases} \Pr\{X_1 = x_1, \cdots, X_k = x_k\}, & \text{if the } X_j \text{ are discrete} \\ \frac{\partial^k}{\partial x_1 \cdots \partial x_k} F_{X_1,\cdots,X_k}(x_1,\cdots,x_k), & \text{if the } X_j \text{ are continuous.} \end{cases}$$

As standard tools for multivariate random variables $(X_1, \cdots, X_k)$, the **joint pgf**, **joint mgf**, and **joint ch.f.** are defined as follows (see Johnson *et al.*, 1997, pages 2-12):

$$
\begin{aligned}
P_{X_1,\cdots,X_k}(t_1,\cdots,t_k) &= \mathrm{E}[t_1^{X_1}\cdots t_k^{X_k}] \\
M_{X_1,\cdots,X_k}(t_1,\cdots,t_k) &= \mathrm{E}[e^{t_1X_1+\cdots+t_kX_k}] \quad = P_{X_1,\cdots,X_k}(e^{t_1},\cdots,e^{t_k}) \\
\phi_{X_1,\cdots,X_k}(t_1,\cdots,t_k) &= \mathrm{E}[e^{\mathrm{i}(t_1X_1+\cdots+t_kX_k)}] \quad = P_{X_1,\cdots,X_k}(e^{\mathrm{i}t_1},\cdots,e^{\mathrm{i}t_k}).
\end{aligned}
$$

Note that in terms of the probability (density) function we have

$$
P_{X_1,\cdots,X_k}(t_1,\cdots,t_k) = \begin{cases} \sum_{(x_1,\cdots,x_k)} f_{X_1,\cdots,X_k}(x_1,\cdots,x_k)\, t_1^{x_1}\cdots t_k^{x_k}, & \text{discrete case} \\ \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f_{X_1,\cdots,X_k}(u_1,\cdots,u_k) t_1^{u_1}\cdots t_k^{u_k} du_1\cdots du_k, & \text{continuous case.} \end{cases}
$$

The joint pgf $P_{X_1,\cdots,X_k}$ or the joint ch.f. $\phi_{X_1,\cdots,X_k}$ completely specifies the joint probability distribution. Equivalent results are obtained either in terms of a pgf or in terms of a ch.f.

- The marginal pgf or ch.f. can be obtained by

$$
\begin{aligned}
P_{X_j}(t_j) &= P_{X_1,\cdots,X_j,\cdots,X_k}(1,\cdots,1,t_j,1,\cdots,1), \\
\phi_{X_j}(t_j) &= \phi_{X_1,\cdots,X_j,\cdots,X_k}(0,\cdots,0,t_j,0,\cdots,0).
\end{aligned}
$$

- If the variables $X_1, \cdots, X_k$ are mutually independent, then

$$
P_{X_1,\cdots,X_k}(t_1,\cdots,t_k) = \prod_{j=1}^{k} P_{X_j}(t_j).
$$

- If two sets of variables $\{X_1, \cdots, X_m\}$ and $\{Y_1, \cdots, Y_n\}$ are independent, then

$$
P_{X_1,\cdots,X_m,Y_1,\cdots,Y_n}(t_1,\cdots,t_m,s_1,\cdots,s_n) = P_{X_1,\cdots,X_m}(t_1,\cdots,t_m)\, P_{Y_1,\cdots,Y_n}(s_1,\cdots,s_n).
$$

- The covariances can be evaluated by $\mathrm{Cov}[X_i,\, X_j] = \mathrm{E}[X_iX_j] - \mathrm{E}[X_i]\mathrm{E}[X_j]$ with

$$
\begin{aligned}
\mathrm{E}[X_iX_j] &= \frac{\partial^2}{\partial t_i\,\partial t_j} P_{X_1,\cdots,X_m}(1,\cdots,1) \\
&= -\frac{\partial^2}{\partial t_i\,\partial t_j} \phi_{X_1,\cdots,X_m}(0,\cdots,0).
\end{aligned}
$$

  This can be seen from the expression

$$
\frac{\partial^2}{\partial t_i\,\partial t_j} P_{X_1,\cdots,X_k}(t_1,\cdots,t_k) = \sum x_i x_j f_{X_1,\cdots,X_k}(x_1,\cdots,x_k)\, t_1^{x_1}\cdots t_i^{x_i-1}\cdots t_j^{x_j-1}\cdots t_k^{x_k}.
$$

- For a discrete multivariate distribution, the joint probability function is

$$
f_{X_1,\cdots,X_k}(x_1,\cdots,x_k) = \frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1}\cdots(\partial t_k)^{x_k}} P_{X_1,\cdots,X_k}(0,\cdots,0) \prod_{i=1}^{k} \frac{1}{x_i!}.
$$

4.3 Aggregation of correlated variables

Theorem 4.1 *For any k correlated variables $X_1, \cdots, X_k$ with joint pgf $P_{X_1,\cdots,X_k}$ and joint ch.f. $\phi_{X_1,\cdots,X_k}$, the sum $Z = X_1 + \cdots + X_k$ has a pgf and a ch.f.*

$$P_Z(t) = P_{X_1,\cdots,X_k}(t, \cdots, t), \qquad \phi_Z(t) = \phi_{X_1,\cdots,X_k}(t, \cdots, t).$$

Proof: $P_Z(t) = \mathrm{E}[t^{X_1+\cdots+X_k}] = \mathrm{E}[t^{X_1} \cdots t^{X_K}] = P_{X_1,\cdots,X_k}(t, \cdots, t)$. □

If we know the joint ch.f. of the k correlated variables, it is straightforward to get the ch.f. for their sum $\phi_Z(t) = \phi_{X_1,\cdots,X_k}(t, \cdots, t)$. Then the probability distribution of Z can be obtained by the inversion method of Heckman/Meyers or via Fast Fourier Transform (FFT) (see Robertson, 1992). The relation $\phi_{X_1+\cdots+X_k}(t) = \phi_{X_1,\cdots,X_k}(t, \cdots, t)$ can be used to

- combine correlated risk portfolios if we let X_i represent the aggregate loss distributions for each individual risk portfolio.
- evaluate the total claim number distribution if we let X_i represent the claim frequency for each individual risk portfolio.
- combine individual claims if we let X_i represent the claim size for each individual risk.

4.4 Aggregation of risk portfolios with correlated frequencies

Consider the aggregation of two correlated risk portfolios:

$$Z = (X_1 + \cdots + X_N) + (Y_1 + \cdots + Y_K),$$

where N and K are correlated, while the pair (N, K) is independent of the claim sizes X and Y, and the X_i's and Y_j's are mutually independent. We have

$$\begin{aligned}
P_Z(t) &= \mathrm{E}[t^Z] = \mathrm{E}[t^{(X_1+\cdots+X_N)+(Y_1+\cdots+Y_K)}] \\
&= \mathrm{E}_{N,K}\mathrm{E}[t^{(X_1+\cdots+X_n)+(Y_1+\cdots+Y_m)} \,|N = n, K = m] \\
&= \mathrm{E}_{N,K}[P_X(t)^N P_Y(t)^K] \\
&= P_{N,K}(P_X(t), P_Y(t)).
\end{aligned}$$

In terms of ch.f. we have

$$\phi_Z(t) = P_{N,K}(\phi_X(t), \phi_Y(t)).$$

PART II. COPULA & MONTE CARLO SIMULATION

5 The Cook-Johnson Family of Multivariate Uniform Distributions

Let $(U_1, \cdots, U_k)$ be a k-dimensional uniform distribution with support on the hypercube $(0,1)^k$ and having the joint cdf

$$F^{(\alpha)}_{U_1,\cdots,U_k}(u_1, \cdots, u_k) = \left\{ \sum_{j=1}^{k} u_j^{-1/\alpha} - k + 1 \right\}^{-\alpha}, \tag{5.1}$$

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

Cook and Johnson (1981) studied the family of multivariate uniform distributions given by eq. (5.1). They showed that

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

and

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

Thus, the correlation approaches to its maximum (i.e. comonotonicity) when α decreases to zero; the correlation approaches zero when α increases to infinity.

Cook and Johnson (1981) also gave the following simple simulation algorithm for the multivariate uniform distribution given by eq. (5.1):

Step 1. Let $Y_1, \cdots, Y_k$ be independent and each has an Exponential(1) distribution.

Step 2. Let Z have a Gamma$(\alpha, 1)$ distribution.

Step 3. Then the variables

$$U_j = [1 + Y_j/Z]^{-\alpha}, \quad j = 1, \cdots, k, \tag{5.2}$$

have a joint cdf given by eq. (5.1).

For a set of arbitrary marginals distributions, $F_{X_1}, \cdots, F_{X_k}$, we can define a joint cdf by

$$F_{X_1,\cdots,X_k}(x_1, \cdots, x_k) = \left\{ \sum_{j=1}^{k} F_{X_j}(x_j)^{-1/\alpha} - k + 1 \right\}^{-\alpha}. \tag{5.3}$$

Alternatively, we can define a joint s.f. by

$$S_{X_1,\cdots,X_k}(x_1,\cdots,x_k) = \left\{ \sum_{j=1}^{k} S_{X_j}(x_j)^{-1/\alpha} - k + 1 \right\}^{-\alpha}. \tag{5.4}$$

For both of the multivariate distributions given by eq. (5.3) and eq. (5.4), the Kendall's tau is

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

Consider the task of aggregating k risk portfolios, $(X_1, \cdots, X_k)$, where each X_j may represent the aggregate loss amount for the j^{th} risk portfolio. If we assume that $(X_1, \cdots, X_k)$ have a multivariate distribution given by eq. (5.3), a simulation of $X_1, \cdots, X_k$ can be easily implemented by:

Step 4. Invert the $(U_1, \cdots, U_k)$ in eq. (5.2) using $(F_{X_1}^{-1}, \cdots, F_{X_k}^{-1})$.

Alternatively, if we assume that $(X_1, \cdots, X_k)$ have a multivariate distribution given by eq. (5.4), a simulation of $X_1, \cdots, X_k$ can be easily implemented by:

Step 4*. Invert the $(U_1, \cdots, U_k)$ in eq. (5.2) using $(S_{X_1}^{-1}, \cdots, S_{X_k}^{-1})$.

In the multivariate uniform distribution given by eq. (5.1), all correlations are positive. Negative correlations can be accommodated by applying the transforms $U_i^* = 1 - U_i$ to some, but not all, uniform variables in eq. (5.2).

In this dependency model, no restriction is imposed on the marginal distributions, F_{X_j} or S_{X_j}, $j = 1, \cdots, k$. However, the correlation parameters are quite restricted in the sense that the Kendall's tau have to be the same for any pair of risks. To overcome this restriction in the correlation parameters, we shall introduce the *normal copula* which permits arbitrary correlation parameters, $\tau_{ij} = \tau(X_i, X_j)$, in the next section.

6 The Normal Copula & Monte Carlo Simulation

In general, the modeling and combining of correlated risks are most straight-forward if the correlated risks have a multivariate normal distribution. In this section, we will use the multivariate normal distribution to construct the *normal copula*, and then use it to generate multivariate distributions with arbitrary marginal distributions. The normal copula enjoys much flexibility in the selection of correlation parameters. As well, it lends itself to simple Monte Carlo simulation techniques.

Assume that $(Z_1, \cdots, Z_k)$ have a multivariate normal distribution with standard normal marginals $Z_j \sim N(0,1)$ and a positive definite correlation matrix

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \cdots & \rho_{1k} \\ \rho_{21} & 1 & \cdots & \rho_{2k} \\ \vdots & \vdots & & \vdots \\ \rho_{k1} & \rho_{k2} & \cdots & 1 \end{pmatrix},$$

where $\rho_{ij} = \rho_{ji}$ is the correlation coefficient between Z_i and Z_j. Note that ρ_{ij} is further constrained by eq. (3.1) and eq. (3.2). Then $(Z_1, \cdots, Z_k)$ have a joint pdf:

$$f(z_1, \cdots, z_k) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left\{-\frac{1}{2}\boldsymbol{z}'\Sigma^{-1}\boldsymbol{z}\right\}, \quad \boldsymbol{z} = (z_1, \cdots, z_k). \tag{6.5}$$

From the correlation matrix Σ we can construct a lower triangular matrix

$$B = \begin{pmatrix} b_{11} & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kk} \end{pmatrix},$$

such that $\Sigma = BB'$. In other words, the correlation matrix Σ equals the matrix product of B and its transpose B'. The elements of the matrix B can be calculated from the following Choleski's algorithm [see Burden and Faires (1989, section 6.6); Johnson (1987, section 4.1)]:

$$b_{ij} = \frac{\rho_{ij} - \sum_{s=1}^{j-1} b_{is} b_{js}}{\sqrt{1 - \sum_{s=1}^{j-1} b_{js}^2}}, \quad 1 \le j \le i \le n, \tag{6.6}$$

with the convention that $\sum_{s=1}^{0}(.) = 0$. It is noted that:

- For $i > j$, the denominator of eq. (6.6) equals b_{jj}.
- The elements of B should be calculated from top to bottom and from left to right.

The following simulation algorithm can be used to generate multivariate normal variables with a joint pdf given by eq. (6.5). [see Fishman (1996, pp. 223-224)]

Step 1. Construct the lower triangular matrix $B = (b_{ij})$ by eq. (6.6).

Step 2. Generate a column vector of independent standard normal variables $\boldsymbol{Y} = (Y_1, \cdots, Y_k)'$.

Step 3. Take the matrix product $\boldsymbol{Z} = B\boldsymbol{Y}$ of B and $\boldsymbol{Y}$. Then $\boldsymbol{Z} = (Z_1, \cdots, Z_k)'$ has the required joint pdf given by eq. (6.5).

Let $\Phi(.)$ represent the cdf of the standard normal distribution:

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-t^2/2}\, dt.$$

Then $\Phi(Z_1), \cdots, \Phi(Z_k)$ have a multivariate uniform distribution with Kendall's tau (e.g. Frees and Valdez, 1997):

$$\tau(\Phi(Z_i), \Phi(Z_j)) = \tau(Z_i, Z_j) = \frac{2}{\pi} \arcsin(\rho_{ij}),$$

where $\arcsin(x)$ is an inverse trigonometric function such that $\sin(\arcsin(x)) = x$.

The following result can be easily verified but nevertheless is stated as a theorem due to its importance.

Theorem 6.1 *Assume that $(Z_1, \cdots, Z_k)$ have a joint pdf given by eq. (6.5) and let $H(z_1, \cdots, z_k)$ be their joint cumulative distribution function. Then*

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

defines a multivariate uniform cdf — called the **normal copula**.

For any set of given marginal cdfs $F_1, \cdots, F_k$, the variables

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

have a joint cdf

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

with marginal cdfs $F_1, \cdots, F_k$ and Kendall's tau

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

Although the normal copula does not have a simple analytical expression, it lends itself to a very simple Monte Carlo simulation algorithm.

Suppose that we are given a set of correlated risks $(X_1, \cdots, X_k)$ with marginal cdfs $F_{X_1}, \cdots, F_{X_k}$ and Kendall's tau $\tau_{ij} = \tau(X_i, X_j)$. If we assume that $(X_1, \cdots, X_k)$ can be described by the normal copula in Theorem 6.1, then the following Monte Carlo simulation procedures can be used:

Step 1. Let $\rho_{ij} = \sin(\frac{\pi}{2}\tau_{ij})$ and construct the lower triangular matrix $B = (b_{ij})$ by eq. (6.6).

Step 2. Generate a column vector of independent standard normal variables $\boldsymbol{Y} = (Y_1, \cdots, Y_k)'$.

Step 3. Take the matrix product of B and $\boldsymbol{Y}$: $\boldsymbol{Z} = (Z_1, \cdots, Z_k)' = B\boldsymbol{Y}$.
Set $u_i = \Phi(Z_i)$ for $i = 1, \cdots, k$.

Step 4. Set $X_i = F_{X_i}^{-1}(u_i)$ for $i = 1, \cdots, k$.

Remark: In Appendix B we give an overview of various other families of copulas and the associated Monte Carlo simulation techniques.

PART III. COMMON MIXTURE & COMPONENT MODELS

7 Common Mixture Models

In many situations, individual risks are correlated since they are subject to the same claim generating mechanism or are influenced by changes in the common underlying economic/legal environment. For instance, in property insurance, risk portfolios in the same geographic location are correlated, where individual claims are contingent on the occurrence and severity of a natural disaster (hurricane, tornado, earthquake, or severe weather condition). In liability insurance, new court rulings or social inflation may set new trends which affect the settlement of all liability claims for one line of business.

One way of modeling situations where the individual risks $\{X_1, X_2, \cdots, X_n\}$ are subject to the same external mechanism is to use a secondary mixing distribution. The uncertainty about the external mechanism is then described by a structure parameter, θ, which can be viewed as a realization of a random variable Θ. The aggregate losses of the risk portfolio can then be seen as a two-stage process: First the external parameter $\Theta = \theta$ is drawn from the distribution function, F_Θ, of Θ. Next, the claim frequency (or severity) of each individual risk X_i, $(i = 1, 2, \cdots, n)$, is obtained as a realization from the conditional distribution function, $F_{X_i|\Theta}(x_i|\theta)$, of $X_i|\Theta$.

7.1 Common Poisson mixtures

Consider k discrete random variables $N_1, \cdots, N_k$. Assume that there exists a random parameter Θ such that

$$(N_j|\Theta = \theta) \sim \text{Poisson}(\theta\lambda_j), \quad j = 1, \cdots, k,$$

where the variable Θ has a pdf $\pi(\theta)$ and a mgf M_Θ. For any given $\Theta = \theta$, the variables $(N_j|\theta)$ are independent and Poisson$(\lambda_j\theta)$ distributed with a conditional joint pgf

$$P_{N_1,\cdots,N_k|\Theta}(t_1, \cdots, t_k|\theta) = \mathrm{E}[t_1^{N_1} \cdots t_k^{N_k}|\Theta = \theta] = e^{\theta[\lambda_1(t_1-1)+\cdots\lambda_k(t_k-1)]}.$$

However, unconditionally, $N_1, \cdots, N_k$ are correlated as they depend upon the same random parameter Θ. The unconditional joint pgf for $N_1, \cdots, N_k$ is

$$\begin{aligned} P_{N_1,\cdots,N_k}(t_1, \cdots, t_k) &= \mathrm{E}_\Theta[\mathrm{E}[t_1^{N_1} \cdots t_k^{N_k}|\Theta]] = \textstyle\int_0^\infty e^{\theta[\lambda_1(t_1-1)+\cdots+\lambda_k(t_k-1)]}\,\pi(\theta)d\theta \\ &= M_\Theta(\lambda_1(t_1-1) + \cdots + \lambda_k(t_k-1)). \end{aligned}$$

It has marginal pgfs $P_{N_j}(t_j) = M_\Theta(\lambda_j(t_j-1))$ with $\mathrm{E}[N_j] = \lambda_j\mathrm{E}[\Theta]$.

Note that

$$\begin{aligned}\mathrm{Cov}[N_i, N_j] &= \mathrm{E}_\Theta \mathrm{Cov}[N_i|\Theta, N_j|\Theta] + \mathrm{Cov}[\mathrm{E}[N_i|\Theta], \mathrm{E}[N_j|\Theta]] \\ &= \mathrm{Cov}[\Theta\lambda_i, \Theta\lambda_j] = \lambda_i\lambda_j \mathrm{Var}[\Theta].\end{aligned}$$

The covariance coefficient between N_i and N_j, $(i \neq j)$, is

$$\omega(N_i, N_j) = \frac{\mathrm{Cov}[N_i, N_j]}{\mathrm{E}[N_i]\,\mathrm{E}[N_j]} = \frac{\mathrm{Var}[\Theta]}{\{\mathrm{E}[\Theta]\}^2}.$$

Example 7.1 If Θ has a Gamma$(\alpha, 1)$ distribution with mgf $M_\Theta(z) = (1-z)^{-\alpha}$, then

$$P_{N_1,\cdots,N_k}(t_1, \cdots, t_k) = [1 - \lambda_1(t_1 - 1) - \cdots - \lambda_k(t_k - 1)]^{-\alpha} \tag{7.1}$$

defines a multivariate negative binomial with marginals NB(α, λ_j) and covariance coefficients $\omega(N_i, N_j) = 1/\alpha$.

Example 7.2 If Θ has an Inverse-Gaussian distribution, IG$(\beta, 1)$, with a mgf $M_\Theta(z) = e^{\frac{1}{\beta}[1-\sqrt{1-2\beta z}]}$, then

$$P_{N_1,\cdots,N_k}(t_1, \cdots, t_k) = \exp\left\{\frac{1}{\beta} - \frac{1}{\beta}\sqrt{1 - 2\beta[\lambda_1(t_1 - 1) + \cdots + \lambda_k(t_k - 1)]}\right\}$$

defines a multivariate Poisson-Inverse-Gaussian with marginals P-IG$(\beta\lambda_j, \lambda_j)$ and covariance coefficients $\omega(N_i, N_j) = \beta$.

Consider combining k risk portfolios. Assume that the frequencies N_j, $j = 1, \cdots, k$, are correlated via a common Poisson-Gamma mixture and have a joint pgf given by eq. (7.1). If the severities X_j, $j = 1, \cdots, k$, are mutually independent and independent of the frequencies, there is a simple method of combining the aggregate loss distributions. Given $\lambda = \lambda_1 + \cdots + \lambda_k$ and $P_X(t) = \frac{\lambda_1}{\lambda} P_{X_1}(t) + \cdots \frac{\lambda_k}{\lambda} P_{X_k}(t)$, then

$$P_{N_1,\cdots,N_k}(P_{X_1}(t), \cdots, P_{X_k}(t)) = [1 - \lambda(P_X(t) - 1)]^{-\alpha}.$$

In other words, the total loss amount for the combined risk portfolios has a compound NB(α, λ) distribution with the severity distribution being a weighted average of individual severity distributions. In this case, dependency does not complicate the computation; in fact, it simplifies the calculation. It is simpler than combining independent compound negative binomial distributions.

In this multivariate Poisson-Gamma mixture model, the k marginals, NB(α, λ_j), are required to have the same parameter α. This requirement limits its applicability in combining risk portfolios; in many practical cases the negative binomial frequencies, NB(α_j, λ_j), have different parameter values, α_j. Later in this paper we will overcome this limitation by extending the Poisson-Gamma mixture model to allow arbitrary negative binomial frequencies, NB(α_j, λ_j).

Similar arguments can be made about the Poisson-Inverse-Gaussian distributions.

7.2 Common exponential mixtures

Consider k continuous random variables $X_1, \cdots, X_k$. Assume that there exists a random parameter Θ such that $(X_j|\Theta = \theta)$ is exponentially distributed with parameter $\lambda_j\theta$ and survivor function

$$S_{X_j|\Theta}(t_j|\theta) = \Pr\{X_j > t_j|\Theta = \theta\} = e^{-\theta\lambda_j t_j}, \quad j = 1, \cdots, k,$$

where the variable Θ has a probability density function $\pi(\theta)$ and a mgf M_Θ.

For any given $\Theta = \theta$, the variables $(X_j|\theta)$, $j = 1, \cdots, k$, are conditionally independent and have a conditional joint survivor function

$$S_{X_1,\cdots,X_k|\Theta}(t_1, \cdots, t_k|\theta) = \Pr\{X_1 > t_1, \cdots, X_k > t_k\,|\Theta = \theta\} = e^{-\theta[\lambda_1 t_1 + \cdots + \lambda_k t_k]}.$$

However, unconditionally, $X_1, \cdots, X_k$ are correlated as they depend upon the same random parameter Θ. The unconditional joint survivor function for $X_1, \cdots, X_k$ is

$$\begin{aligned} S_{X_1,\cdots,X_k}(t_1, \cdots, t_k) &= \textstyle\int_0^\infty e^{-\theta[\lambda_1 t_1 + \cdots + \lambda_k t_k]}\,\pi(\theta)d\theta \\ &= M_\Theta(-\lambda_1 t_1 - \cdots - \lambda_k t_k). \end{aligned}$$

Example 7.3 If Θ has a Gamma$(\alpha, 1)$ distribution with mgf $M_\Theta(z) = (1-z)^{-\alpha}$, this defines a family of multivariate Pareto distributions

$$S_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = [1 + \lambda_1 t_1 + \cdots + \lambda_k t_k]^{-\alpha},$$

with marginals Pareto$(\alpha, 1/\lambda_j)$.

Example 7.4 If Θ has an inverse Gaussian distribution with mgf $M_\Theta(z) = e^{\frac{1}{\beta}\left[1-\sqrt{1-2\beta z}\,\right]}$, this defines a family of multivariate E-IG distribution

$$S_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \exp\left[\frac{1}{\beta} - \frac{1}{\beta}\sqrt{1 + 2\beta(\lambda_1 t_1 + \cdots + \lambda_k t_k)}\right],$$

with marginals E-IG$(\beta\lambda_j, \lambda_j)$.

Now we consider the aggregation of k individual claim amounts. Suppose that the k individual claim amounts $X_1, \cdots, X_k$ are identically distributed with $X_i \sim$ Pareto(α, β). But they are correlated by a common Exponential-Gamma mixture with a joint survivor function

$$S_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \left[1 + \frac{1}{\beta}(t_1 + \cdots + t_k)\right]^{-\alpha}.$$

Then the sum $X_1 + \cdots + X_k$ has a Pareto$(\alpha, n\beta)$ distribution. This is because, for any given $\Theta = \theta$, $(X_1 + \cdots + X_k|\theta) \sim$ Exponential(θ/n).

Alternatively, this common exponential mixture model can be obtained by applying the Cook-Johnson copula to k identical marginal survivor functions, Pareto(α, β). In other words, the Cook-Johnson copula can be viewed as an extension of the common exponential mixture model.

8 Extended Common Poisson Mixture Models

The common Poisson mixture models in the previous section have very simple correlation structures and are very easy to use. However, they are quite restricted in the sense that it does not permit arbitrary parameter values in the marginal distributions. In this section we extend the common Poisson mixture model so that the marginal distributions may have arbitrary parameter values. This extended model permits simple implementation by Monte Carlo simulations.

Suppose that there exist random variables $(\Theta_1, \cdots, \Theta_k)$ such that the conditional variables $(N_1, \cdots, N_k)|(\Theta_1 = \theta_1, \cdots, \Theta_k = \theta_k)$ are independent Poisson(θ_j) variables with

$$P_{N_1,\cdots,N_k|(\Theta_1,\cdots,\Theta_k)}(t_1, \cdots, t_k|\theta_1, \cdots, \theta_k) = \prod_{j=1}^{k} P_{N_j}(t_j|\theta_j) = \prod_{j=1}^{k} e^{\theta_j(t_j-1)},$$

where $M_{\Theta_1,\cdots,\Theta_k}(t_1, \cdots, t_k) = \mathrm{E}_{\Theta_1,\cdots,\Theta_k}[e^{t_1\Theta_1+\cdots+t_k\Theta_k}]$ is the joint mgf of $(\Theta_1, \cdots, \Theta_k)$.

The unconditional joint pgf is

$$\begin{aligned} P_{N_1,\cdots,N_k}(t_1, \cdots, t_k) &= \mathrm{E}_{(\Theta_1,\cdots,\Theta_k)} P_{N_1,\cdots,N_k}(t_1, \cdots, t_k|\Theta_1, \cdots, \Theta_k) \\ &= M_{\Theta_1,\cdots,\Theta_k}((t_1 - 1), \cdots, (t_k - 1)). \end{aligned}$$

By taking the first and second order partial derivatives of this joint pgf at $(1, \cdots, 1)$, we obtain

$$\mathrm{E}[N_i] = \mathrm{E}[\Theta_i], \quad \text{and} \quad \mathrm{Cov}[N_i, N_j] = \mathrm{Cov}[\Theta_i, \Theta_j].$$

We observe a one-to-one correspondence between the correlation structures of the variables $(N_1, \cdots, N_k)$ and the mixing parameters $(\Theta_1, \cdots, \Theta_k)$.

Now consider the case that $\Theta_j \sim \mathrm{Gamma}(\alpha_j, \beta_j)$, and thus $N_j \sim \mathrm{NB}(\alpha_j, \beta_j)$, with arbitrary parameter values, $\alpha_j, \beta_j > 0$. We further assume that the variables Θ_j, $j = 1, \cdots, k$, are comonotonic, thus can be simulated by using the same set of uniform random numbers. For $i \neq j$, the covariance $\mathrm{Cov}[\Theta_i, \Theta_j]$ can be numerically calculated by using eq. (3.1). For this dependency model, we have a simple Monte Carlo simulation algorithm:

Step 1. Generate a uniform number, u, from $U \sim \mathrm{Uniform}(0, 1)$.

Step 2. Let $\theta_j = F_{\Theta_j}^{-1}(u)$, where $\Theta_j \sim \mathrm{Gamma}(\alpha_j, \beta_j)$, $j = 1, \cdots, k$.

Step 3. Simulate $(N_1, \cdots, N_k)$ from k independent Poisson(θ_j) variables, $j = 1, \cdots, k$.

If the α_j's are the same, we get the common Poisson mixture model in Example 7.1.

9 Component Models

Consider the aggregation of different lines of business. For a multi-line insurer, the correlation between lines of business may differ from one region to another. Therefore, it may be more appropriate to divide each line into components and model the correlation separately for each component (e.g. by geographic region). There may exist higher correlations between lines in a high catastrophe risk region where the presense of the catastrophe risk may generate a common shock or a common mixture.

Note that many families of frequency and severity distributions are infinitely divisible. Let $X \oplus Y$ represent the sum of two independent random variables, and $F_Y \oplus F_Y$ represent the convolution of two probability distributions. We have

- $\text{Poisson}(\lambda_1) \oplus \text{Poisson}(\lambda_2) = \text{Poisson}(\lambda_1 + \lambda_2)$
- $\text{NB}(\alpha_1, \beta) \oplus \text{NB}(\alpha_2, \beta) = \text{NB}(\alpha_1 + \alpha_2, \beta)$
- $\text{P-IG}(\beta, \mu_1) \oplus \text{P-IG}(\beta, \mu_2) = \text{P-IG}(\beta, \mu_1 + \mu_2)$
- $\text{Gamma}(\alpha_1, \beta) \oplus \text{Gamma}(\alpha_2, \beta) = \text{Gamma}(\alpha_1 + \alpha_2, \beta)$
- Inverse-Gaussian: $\text{IG}(\beta, \mu_1) \oplus \text{IG}(\beta, \mu_2) = \text{IG}(\beta, \mu_1 + \mu_2)$

Infinitely divisible distributions are especially useful for dividing risks into independent components. Consider k infinitely divisible risks $X_j(\alpha_j)$, $j = 1, \cdots, k$, with α_j as the divisible parameter.

Consider a decomposition:

$$\begin{array}{rclcl} X_1(\alpha_1) & = & X_{11}(\alpha_{11}) & \oplus \cdots \oplus & X_{1n}(\alpha_{1n}) \\ \vdots & & \vdots & & \vdots \\ X_k(\alpha_k) & = & X_{k1}(\alpha_{k1}) & \oplus \cdots \oplus & X_{kn}(\alpha_{kn}) \end{array}, \quad \alpha_{js} \geq 0. \tag{9.1}$$

Then we can generate correlation structures component by component:

$$P_{X_1, \cdots, X_k} = \prod_{s=1}^{n} Q_{X_{1s}, \cdots, X_{ks}},$$

where the joint pgf $Q_{X_{1s}, \cdots, X_{ks}}$ for the s^{th} components can be modeled by using a common mixture, a common shock, or by assuming independence, as appropriate. It can be verified that for the component model in eq. (9.1) we have

$$\text{Cov}[X_i, X_j] = \sum_{s=1}^{n} \text{Cov}[X_{is}, X_{js}].$$

9.1 Common shocks models

Let $X_j = X_{ja} \oplus X_{jb}$, $j = 1, \cdots, k$, be a decomposition into two *independent* components.

$$P_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \mathrm{E}[t_1^{X_{1a}} \cdots t_k^{X_{ka}}]\,\mathrm{E}[t_1^{X_{1b}} \cdots t_k^{X_{kb}}].$$

If $X_{1a} = \cdots = X_{ka} = X_0$, we obtain

$$P_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \mathrm{E}[(t_1 \cdots t_k)^{X_0}]\,\mathrm{E}[t_1^{X_{1b}} \cdots t_k^{X_{kb}}].$$

In particular, if the X_{ib}'s are independent, we have $\mathrm{Cov}[X_i, X_j] = \mathrm{Var}[X_0]$. The only source of correlation comes from the common shock variable X_0.

Example 9.1 Consider the aggregation of two correlated compound Poisson distributions:

- Portfolio 1. The claim frequency N_1 has a Poisson(λ_1) distribution and the claim severity X has a probability function $f_1(x)$.
- Portfolio 2. The claim frequency N_2 has a Poisson(λ_2) distribution and the claim severity Y has a probability function $f_2(y)$.
- Assume that X, Y are independent and both are independent of (N_1, N_2). However, N_1 and N_2 are correlated via a common shock model

$$N_1 = N_0 \oplus N_{1b}, \quad N_2 = N_0 \oplus N_{2b}$$

where $N_0 \sim$ Poisson(λ_0), $N_{1b} \sim$ Poisson$(\lambda_1 - \lambda_0)$, and $N_{2b} \sim$ Poisson$(\lambda_2 - \lambda_0)$.

In this common shock model (N_1, N_2) have a joint pgf:

$$P_{N_1,N_2}(t_1, t_2) = E[t_1^{N_1} t_2^{N_2}] = \exp[\lambda_1(t_1 - 1) + \lambda_2(t_2 - 1) + \lambda_0(t_1 - 1)(t_2 - 1)],$$

with $\mathrm{Cov}[N_1, N_2] = \mathrm{Var}[X_0] = \lambda_0$. It can be shown that the aggregate losses for the combined risk portfolio,

$$S = (X_1 + \cdots + X_{N_1}) + (Y_1 + \cdots + Y_{N_2}),$$

has a compound Poisson$(\lambda_1 + \lambda_2 - \lambda_0)$ distribution with a severity probability function

$$f(z) = \frac{\lambda_1 - \lambda_0}{\lambda_1 + \lambda_2 - \lambda_0} f_1(z) + \frac{\lambda_2 - \lambda_0}{\lambda_1 + \lambda_2 - \lambda_0} f_2(z) + \frac{\lambda_0}{\lambda_1 + \lambda_2 - \lambda_0} f_{1*2}(z),$$

where f_{1*2} represents the convolution of f_1 and f_2. Thus existing methods can be applied.

This common shock model can be easily extended to any higher dimension $(k > 2)$. For illustrative purposes, now we give an example involving three frequency variables.

Example 9.2 The joint pgf

$$P_{N_1,N_2,N_3}(t_1, t_2, t_3) = \exp\left\{\sum_{i=1}^{3} \lambda_{ii}(t_i - 1) + \sum_{i<j} \lambda_{ij}(t_i t_j - 1) + \lambda_{123}(t_1 t_2 t_3 - 1)\right\}, \tag{9.2}$$

defines a multivariate Poisson distribution with marginals

$$N_j \sim \text{Poisson}(\lambda_{123} + \sum_{i=1}^{3} \lambda_{ij}), \quad j = 1, 2, 3,$$

and for $i \neq j$, $\text{Cov}[N_i, N_j] = \lambda_{ij} + \lambda_{123}$.

We let

- $K_{ii} \sim \text{Poisson}(\lambda_{ii})$, for $i = 1, 2, 3$.
- $K_{ij} \sim \text{Poisson}(\lambda_{ij})$, for $1 \leq i < j \leq 3$.
- $K_{ij} = K_{ji}$, for $1 \leq i, j \leq 3$.
- $K_{123} \sim \text{Poisson}(\lambda_{123})$.
- $N_j = K_{1j} \oplus K_{2j} \oplus K_{3j} \oplus K_{123}$, for $j = 1, 2, 3$.

Then the so-constructed (N_1, N_2, N_3) have a joint pgf given by (9.2). In this model, K_{123} represents the common shock among all three variables (N_1, N_2, N_3). In addition, for $i \neq j$, $K_{ij} = K_{ji}$ represents the extra common shock between N_i and N_j.

Note that we can easily simulate the correlated frequencies, (N_1, N_2, N_3), component by component.

Subject to scale transforms, the common shock multivariate Poisson model can be extended to gamma variables.

Example 9.3 Consider two variables $X_1 \sim \text{Gamma}(\alpha_1, \beta_1)$ and $X_2 \sim \text{Gamma}(\alpha_2, \beta_2)$. Suppose there is a decomposition

$$X_1 = \beta_1(X_0 \oplus X_{1b}), \quad X_2 = \beta_2(X_0 \oplus X_{2b}),$$

where $X_0 \sim \text{Gamma}(\alpha_0, 1)$ with $\alpha_0 \leq \min\{\alpha_1, \alpha_2\}$, $X_{1b} \sim \text{Gamma}(\alpha_1 - \alpha_0, 1)$ and $X_{2b} \sim \text{Gamma}(\alpha_2 - \alpha_0, 1)$. Then $\text{Cov}[X_1, X_2] = \beta_1\beta_2\text{Var}[X_0] = \alpha_0\beta_1\beta_2$ and

$$X_1 + X_2 = (\beta_1 + \beta_2)X_0 \oplus \beta_1 X_{1b} \oplus \beta_2 X_{2b}.$$

9.2 Peeling method

Recall that the common Poisson-Gamma mixture requires that the marginals $N_j \sim \text{NB}(\alpha, \lambda_j)$ must have the same parameter value α. Now we shall illustrate that, by using the component method, we can construct correlated multivariate negative binomials with arbitrary parameters (α_j, λ_j).

Suppose that we are given k marginal negative binomial distributions:

$$N_1 \sim \text{NB}(\alpha_1, \lambda_1), \quad \cdots, \quad N_k \sim \text{NB}(\alpha_k, \lambda_k).$$

Model 1. Let $\alpha_0 \leq \min\{\alpha_1, \cdots, \alpha_k\}$ and let each N_j $(j = 1, \cdots, k)$ have a decomposition:

$$N_j = N_{ja} \oplus N_{jb}, \quad N_{ja} \sim \text{NB}(\alpha_0, \lambda_j), \quad N_{jb} \sim \text{NB}(\alpha_j - \alpha_0, \lambda_j).$$

Note that the N_{ja}'s have the same parameter α_0, thus can be modeled by a common Poisson-Gamma mixture

$$P_{N_{1a},\cdots,N_{ka}}(t_1,\cdots,t_k) = \{1-\lambda_1(t_1-1)-\cdots-\lambda_k(t_k-1)\}^{-\alpha_0}.$$

If we assume that the N_{jb}'s are independent, then $(N_1,\cdots,N_k)$ have a joint pgf

$$P_{N_1,\cdots,N_k}(t_1,\cdots,t_k) = \{1-\lambda_1(t_1-1)-\cdots-\lambda_k(t_k-1)\}^{-\alpha_0}\prod_{j=1}^{k}\{1-\lambda_j(t_j-1)\}^{\alpha_0-\alpha_j}.$$

Note that $\mathrm{Cov}[N_i,N_j] = \alpha_0\lambda_i\lambda_j = \frac{\alpha_0}{\alpha_i\alpha_j}\mathrm{E}[N_i]\mathrm{E}[N_j]$. Simple methods exist for combining the individual aggregate loss distributions, provided that the severities are mutually independent, and independent of $(N_1,\cdots,N_k)$.

Model 2. Assume that the α_j are in an ascending order, $\alpha_1\le\cdots\le\alpha_k$. The decomposition

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

can be used in conjunction with common mixture models to generate the following joint pgf:

$$\begin{aligned}P_{N_1,\cdots,N_k}(t_1,\cdots,t_k) &= \{1-\lambda_1(t_1-1)-\cdots-\lambda_k(t_k-1)\}^{-\alpha_1}\\ &\quad\times\{1-\lambda_2(t_2-1)-\cdots-\lambda_k(t_k-1)\}^{\alpha_1-\alpha_2}\\ &\quad\times\cdots\times\{1-\lambda_k(t_k-1)\}^{\alpha_{k-1}-\alpha_k}.\end{aligned}$$

It can be verified that the marginal univariate pgf is $P_{N_j}(t_j) = [1-\lambda_j(t_j-1)]^{-\alpha_j}$ and the marginal bivariate pgf is

$$P_{N_i,N_j}(t_i,t_j) = \{1-\lambda_i(t_i-1)-\lambda_j(t_j-1)\}^{-\alpha_i}\times\{1-\lambda_j(t_j-1)\}^{\alpha_i-\alpha_j},\quad i<j,$$

with $\mathrm{Cov}[N_i,N_j] = \alpha_i\lambda_i\lambda_j = \frac{1}{\alpha_j}\mathrm{E}[N_i]\mathrm{E}[N_j]$.

9.3 Mixed correlation models

Assume that the joint pgfs $P_{X_1,\cdots,X_k}$ and $Q_{X_1,\cdots,X_k}$ have the same marginals $P_{X_1},\cdots,P_{X_k}$. Then the mixed joint pgf

$$q\ P_{X_1,\cdots,X_k}(t_1,\cdots,t_k)+(1-q)\ Q_{X_1,\cdots,X_k}(t_1,\cdots,t_k),\quad(0<q<1),$$

also has marginal pgfs $P_{X_1},\cdots,P_{X_k}$. For this mixed joint pgf, we have

$$\mathrm{Cov}[X_i,X_j] = (1-q)\mathrm{Cov}^P[X_i,X_j]+q\mathrm{Cov}^Q[X_i,X_j],$$

where Cov^P and Cov^Q represent the covariances implied by the joint pgfs P and Q, respectively.

The mixture of joint pgfs can be used to adjust, up or down, the covariance coefficients. For example, if we feel that a common mixture joint pgf P would give too strong of a correlation, then we can mix it with an independent joint pgf Q.

PART IV. JOINT PGF CONSTRUCTION & FFT METHODS

10 Distortion of Joint Probability Generating Functions

Let $X_1, \cdots, X_k$ be k random variables (discrete, continuous, or multivariate variables) with pgfs $P_{X_1}(t_1), \cdots, P_{X_k}(t_k)$, respectively. If the X_j's are mutually independent, we have

$$P_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \prod_{j=1}^{k} P_{X_j}(t_j).$$

Let g be a strictly increasing function over $[0, 1]$ with $g(1) = 1$ and whose inverse function is g^{-1}. *In a quite loose sense*, we assume that $g \circ P_{X_1,\cdots,X_k}$ specifies a joint pgf with marginal pgfs $g \circ P_{X_j}$, $(j = 1, \cdots, k)$. By assuming that the distorted joint pgf $g \circ P_{X_1,\cdots,X_k}$ has uncorrelated marginals, namely,

$$g \circ P_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \prod_{j=1}^{k} g \circ P_{X_j}(t_j),$$

a correlation structure is introduced to the original joint pgf:

$$P_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = g^{-1}\left\{\prod_{j=1}^{k} g \circ P_{X_j}(t_j)\right\}.$$

For mathematical convenience we introduce $h(x) = \ln g(x)$ which is a strictly increasing function over $[0, 1]$ with $h(1) = 0$. In terms of h, the above equation can be expressed as

$$P_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = h^{-1}\left\{\sum_{j=1}^{k} h \circ P_{X_j}(t_j)\right\}. \tag{10.1}$$

Note that eq. (10.1) may not define a proper multivariate distribution, as the only constraint on the joint probability (density) function is that they sum to one. It defines a proper multivariate distribution if, and only if, the joint probability (density) function, $f_{X_1,\cdots,X_k}$, is non-negative everywhere. Recall that for discrete distributions,

$$f_{X_1,\cdots,X_k}(x_1, \cdots, x_k) = \frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1}\cdots(\partial t_k)^{x_k}} P_{X_1,\cdots,X_k}(0, \cdots, 0) \prod_{i=1}^{k} \frac{1}{x_i!}.$$

This expression can also be verified by using the Taylor series expansion. Thus it will suffice to verify that all partial derivatives of $P_{X_1,\cdots,X_k}$ at $t_1 = \cdots = t_k = 0$ are non-negative.

Theorem 10.1 *If eq. (10.1) does define a joint pgf, then we have*

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

Proof: We will take the second order partial derivative, $\frac{\partial^2}{\partial t_i \partial t_j}$, $(i \neq j)$, on both sides of the equation

$$h \circ P_{X_1,\cdots,X_k}(t_1,\cdots,t_k) = \sum_{j=1}^{k} h \circ P_{X_j}(t_j).$$

We obtain zero by taking the second order partial derivative, $\frac{\partial^2}{\partial t_i \partial t_j}$, $(i \neq j)$, on the right-hand side. Thus we should also get zero for the second order partial derivative on the left-hand side:

$$0 = \frac{\partial^2}{\partial t_i \partial t_j} \{h \circ P_{X_1,\cdots,X_k}\} = \frac{\partial}{\partial t_i} \left\{ h'(P_{X_1,\cdots,X_k}) \frac{\partial P_{X_1,\cdots,X_k}}{\partial t_j} \right\},$$

which further yields that

$$h''(P_{X_1,\cdots,X_k}) \frac{\partial P_{X_1,\cdots,X_k}}{\partial t_i} \frac{\partial P_{X_1,\cdots,X_k}}{\partial t_j} + h'(P_{X_1,\cdots,X_k}) \frac{\partial^2 P_{X_1,\cdots,X_k}}{\partial t_i \partial t_j} = 0.$$

Setting the values $t_s = 1$ for $s = 1,\cdots,k$, we get

$$h''(1)\mathrm{E}[X_i]\mathrm{E}[X_j] + h'(1)\mathrm{E}[X_i X_j] = 0.$$

□

This family of multivariate distributions has a symmetric structure in the sense that ω_{ij} is the same for all $i \neq j$. It would be suitable for combining risks in the same class, where any two individual risks share the same covariance coefficient.

Questions remain as to which distortion function to use, and whether the distortion method by eq. (10.1) defines a proper multivariate distribution. In general, the feasibility of the distortion method depends on the marginal distributions.

The next section shows how the distortion method is inherently connected to the common Poisson-mixture models.

10.1 Links with the common Poisson mixtures

Reconsider the common Poisson mixture model in Section 7.1: For any given θ, $(N_j|\Theta = \theta)$, $j = 1,\cdots,k$, are conditionally independent Poisson variables with mean $\lambda_j\theta$. If the random parameter Θ has a mgf M_Θ, then $(N_1,\cdots,N_k)$ has an unconditional joint pgf

$$P_{N_1,\cdots,N_k}(t_1,\cdots,t_k) = M_\Theta(\lambda_1(t_1-1)+\cdots+\lambda_k(t_k-1)),$$

with marginal pgf $P_{N_j}(t_j) = M_\Theta(\lambda_j(t_j-1))$.

Lemma 10.1 *For a non-negative random variable Θ, the inverse of the moment generating function, M_Θ^{-1}, is well defined over the range $[0,1]$ with $\frac{d}{du}M_\Theta^{-1}(u) > 0$, $M_\Theta^{-1}(0) = -\infty$ and $M_\Theta^{-1}(1) = 0$.*

If we define $h(y) = M_\Theta^{-1}(y)$, then the joint pgf for the common Poisson mixture model satisfies

$$P_{N_1,\cdots,N_k}(t_1,\cdots,t_k) = h^{-1}\left\{\sum_{j=1}^{k} h \circ P_{N_j}(t_j)\right\}.$$

Example 10.1 If Θ has a Gamma$(1/\omega, 1)$ distribution with mgf $M_\Theta(z) = (1-z)^{-1/\omega}$, then $h(y) = 1 - y^{-\omega}$ and we get the following **joint** pgf:

$$P^{(\omega)}_{X_1,\cdots,X_k}(t_1,\cdots,t_k) = \{\ P_{X_1}(t_1)^{-\omega} + \cdots + P_{X_k}(t_k)^{-\omega} - k + 1\ \}^{-\frac{1}{\omega}}, \quad \omega \neq 0,$$

with $\mathrm{Cov}[X_i, X_j] = \omega\,\mathrm{E}[X_i]\mathrm{E}[X_j]$ and $\lim_{\omega\to 0} P^{(\omega)}_{X_1,\cdots,X_k} = P_{X_1}(t_1)\cdots P_{X_k}(t_k)$.

Example 10.2 If Θ has an inverse Gaussian distribution, IG$(\omega, 1)$, with a mgf $M_\Theta(z) = \exp\{\frac{1}{\omega}[1-\sqrt{1-2\omega z}]\}$, then $h(y) = \ln y - \frac{\omega}{2}(\ln y)^2$ and we get the following **joint** pgf:

$$P^{(\omega)}_{X_1,\cdots,X_k}(t_1,\cdots,t_k) = \exp\left\{\frac{1}{\omega} - \sqrt{\frac{1}{\omega^2} - \sum_{j=1}^{k}[\frac{2}{\omega}\ln P_{X_j}(t_j) - (\ln P_{X_j}(t_j))^2]}\right\},$$

with $\mathrm{Cov}[X_i, X_j] = \omega\,\mathrm{E}[X_i]\mathrm{E}[X_j]$ and $\lim_{\omega\to 0} P^{(\omega)}_{X_1,\cdots,X_k} = P_{X_1}(t_1)\cdots P_{X_k}(t_k)$.

10.2 A family of multivariate negative binomial distributions

As an example of the distortion method, now we give a family of multivariate distributions with arbitrary negative binomial marginal distributions, NB(α_j, β_j), $j = 1, \cdots, k$.

Theorem 10.2 *The joint pgf*

$$P_{N_1,\cdots,N_k}(t_1,\cdots,t_k) = \left\{\sum_{j=1}^{k}[1-\beta_j(t_j-1)]^{\alpha_j\omega} - k + 1\right\}^{-\frac{1}{\omega}}, \quad \omega \neq 0, \tag{10.2}$$

defines a multivariate negative binomial distribution with marginals NB(α_j, β_j) when either of the following conditions holds:

- $0 < \omega < \min\{1/\alpha_j,\ j = 1, \cdots, k\}$,
- $\omega < 0$ *such that* $P_{N_1,\cdots,N_k}(0,\cdots,0) > 0$ *and* $1/\omega$ *is a negative integer.*

Proof: The eq. (10.2) can be rewritten as

$$P_{N_1,\cdots,N_k}(t_1,\cdots,t_k) = Q(t_1,\cdots,t_k)^{-\frac{1}{\omega}},$$

where

$$Q(t_1,\cdots,t_k) = \sum_{j=1}^{k}[1+\beta_j - \beta_j t_j]^{\alpha_j\omega} - k + 1.$$

(i) For $0 < \omega < \min\{1/\alpha_j,\ j = 1, \cdots, k\}$ we have $\alpha_j \omega \leq 1$ and the partial derivatives $\frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1}\cdots(\partial t_k)^{x_k}} P_{N_1,\cdots,N_k}$ are the sum of terms of the following form:

$$a\, Q(t_1, \cdots, t_k)^{-b} \prod_{j=1}^{k} [1 + \beta_j - \beta_j t_j]^{-c_j}, \qquad a, b, c_j \geq 0.$$

Thus, the joint probability function

$$f_{N_1,\cdots,N_k}(x_1, \cdots, x_k) = \frac{\partial^{x_1+\cdots+x_k}}{(\partial t_1)^{x_1} \cdots (\partial t_k)^{x_k}} P_{N_1,\cdots,N_k}(0, \cdots, 0) \prod_{i=1}^{k} \frac{1}{x_i!}$$

is always non-negative. Therefore eq. (10.2) does define a proper joint distribution.

(ii) When $\omega < 0$ such that $P_{N_1,\cdots,N_k}(0, \cdots, 0) > 0$ and $1/\omega$ is a negative integer, we have

$$P(t_1, \cdots, t_k) = Q(t_1, \cdots, t_k)^n, \qquad \text{where } n = -1/\omega \text{ is a positive integer.}$$

which can be viewed as the n-fold convolutions of $Q(t_1, \cdots, t_k)$. Note that $[1 + \beta_j - \beta_j t_j]^{\alpha_j \omega}$ represents the pgf of $\mathrm{NB}(-\alpha_j \omega, \beta_j)$. Thus, $Q(t_1, \cdots, t_k)$ defines a proper multivariate distribution as long as $P_{N_1,\cdots,N_k}(0, \cdots, 0) > 0$. □

Note that eq. (10.2) allows arbitrary marginal negative binomial distributions, $\mathrm{NB}(\alpha_j, \lambda_j)$, thus is more general than the common Poisson-Gamma mixture model. In the special case that all α_j are the same, $\alpha_j = \alpha$, the family of joint distributions in eq. (10.2) return to the common Poisson-Gamma mixture model with $\omega = 1/\alpha$.

Remark: Assume that k individual risk portfolios are specified by their frequencies and severities: (N_j, X_j), $j = 1, \cdots, k$. If $(N_1, \cdots, N_k)$ has a joint pgf as in eq. (10.2), and the only correlation exists between the frequencies, then the aggregate loss, Z, for the combined risk portfolios has a ch.f.

$$\phi_Z(t) = \left\{ \sum_{j=1}^{k} [1 - \beta_j(\phi_{X_j}(t) - 1)]^{\alpha_j \omega} - k + 1 \right\}^{-\frac{1}{\omega}}, \quad \omega \neq 0.$$

Thus FFT can be used to evaluate the aggregate loss distribution.

11 A Family of Multivariate Distributions with Given Marginals and Covariance Matrix

We define a family of joint ch.f. with parameters ω_{ij} $(1 \leq i < j \leq k)$ as follows:

$$\phi_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \phi_{X_1}(t_1) \cdots \phi_{X_k}(t_k) \left\{ 1 + \sum_{i<j} \omega_{ij} \left[1 - \phi_{X_i}(t_i)\right] \left[1 - \phi_{X_j}(t_j)\right] \right\}. \qquad (11.1)$$

Theorem 11.1 *If $(X_1, \cdots, X_k)$ has a joint ch.f. as in eq. (11.1), then we have*

$$\mathrm{Cov}[X_i, X_j] = \omega_{ij}\, \mathrm{E}[X_i]\, \mathrm{E}[X_j], \quad i \neq j.$$

Proof: Rewrite the joint ch.f. as $\phi = U \cdot W$ where

$$U = U(t_1, \cdots, t_k) = \phi_{X_1}(t_1) \cdots \phi_{X_k}(t_k)$$

and

$$W = W(t_1, \cdots, t_k) = 1 + \sum_{i<j} \omega_{ij} [1 - \phi_{X_i}(t_i)]\, [1 - \phi_{X_j}(t_j)].$$

For $i \neq j$, we have

$$\frac{\partial U}{\partial t_i} = \frac{U}{\phi_{X_i}(t_i)}\, \phi'_{X_i}(t_i), \qquad \frac{\partial^2 U}{\partial t_i \partial t_j} = \frac{U}{\phi_{X_i}(t_i)\phi_{X_j}(t_j)}\, \phi'_{X_i}(t_i)\, \phi'_{X_j}(t_j),$$

and

$$\frac{\partial W}{\partial t_i} = -\left\{ \sum_{j \neq i} \omega_{ij} [1 - \phi_{X_j}(t_j)] \right\} \phi'_{X_i}(t_i), \qquad \frac{\partial^2 W}{\partial t_i \partial t_j} = \omega_{ij}\, \phi'_{X_i}(t_i)\, \phi'_{X_j}(t_j).$$

Note that

$$\frac{\partial^2 \phi}{\partial t_i \partial t_j} = \frac{\partial^2 U}{\partial t_i \partial t_j} W + \frac{\partial U}{\partial t_i} \frac{\partial W}{\partial t_j} + \frac{\partial U}{\partial t_j} \frac{\partial W}{\partial t_i} + U \frac{\partial^2 W}{\partial t_i \partial t_j}.$$

Setting $t_s = 0$ for $s = 1, \cdots, k$, we have $U(0, \cdots, 0) = W(0, \cdots, 0) = 1$ and $\frac{\partial W}{\partial t_s}(0, \cdots, 0) = 0$. Thus, we have

$$\mathrm{E}[X_i X_j] = -\frac{\partial^2 \phi}{\partial t_i \partial t_j}(0, \cdots, 0) = -(1 + \omega_{ij})\, \phi'_{X_i}(0)\, \phi'_{X_j}(0) = (1 + \omega_{ij}) \mathrm{E}[X_i] \mathrm{E}[X_j].$$

□

For simplicity, we consider the bivariate case with joint ch.f.

$$\phi_{X,Y}(t, s) = \phi_X(t)\, \phi_Y(s)\, \left\{1 + \omega [1 - \phi_X(t)][1 - \phi_Y(s)]\right\}.$$

Expanding this expression we obtain

$$\phi_{X,Y}(t, s) = (1 + \omega)\phi_X(t)\, \phi_Y(s) - \omega \phi_X(t)^2\, \phi_Y(s) - \omega \phi_X(t)\, \phi_Y(s)^2 + \omega \phi_X(t)^2\, \phi_Y(s)^2.$$

from this we can easily identify the joint pdf

$$f_{X,Y}(x, y) = f_X(x)\, f_Y(y) \left[1 + \omega \left(1 - \frac{f_X^{*2}(x)}{f_X(x)}\right) \left(1 - \frac{f_Y^{*2}(y)}{f_Y(y)}\right) \right]. \tag{11.2}$$

If the ratios $f_X^{*2}(x)/f_X(x)$ and $f_Y^{*2}(y)/f_Y(y)$ are bounded from above, there exist $a, b > 0$, such that $f_{X,Y}$ in eq. (11.2) defines a proper bivariate pdf for all $\omega \in (-a, b)$. To see this, let $A = \max[f_X^{*2}(x)/f_X(x)] \geq 1$ and $B = \max[f_Y^{*2}(y)/f_Y(y)] \geq 1$, then we have $a = b = (A-1)^{-1}(B-1)^{-1}$.

The ratio $f_X^{*2}(x)/f_X(x)$ is very closely related to the right-tail behavior of f_X. In fact, there exists a branch of mathematics, called sub-exponential distributions, which studies specifically the behavior of this ratio for some distributions. Embrecht and Veraverbeke (1982) give a good review of the sub-exponential theory in relation to insurance claims modeling. Here I summarize the relevant results in the literature on sub-exponential distributions. For details, the readers can consult Chapter 10 of Panjer and Willmot (1992).

Heavy-tailed If the probability (density) function f_X satisfies

$$\lim_{x\to\infty} \frac{f_X^{*2}(x)}{f_X(x)} = 2,$$

then the distribution with pdf f_X is said to be **sub-exponential**. The sub-exponential family include the following members:

Transformed Beta (Venter, 1983)	Burr
Pareto	Loglogistic
Log-normal	Exponential-Inverse-Gaussian (E-IG)
Weibull $(0 < c < 1)$	Inverse Gamma

Moderate-tailed If the probability (density) function f_X satisfies

$$\lim_{x\to\infty} \frac{f_X^{*2}(x)}{f_X(x)} = C, \qquad \text{where } C \text{ is some constant greater than 2,}$$

then the distribution with pdf f_X is said to be **moderate-tailed**. Examples include:

Generalized Inverse Gaussian $(\lambda < 0)$	Poisson-GIG $(\lambda < 0)$
Inverse Gaussian	Poisson-Inverse-Gaussian (P-IG)

Light-tailed If the probability (density) function f_X satisfies

$$\lim_{x\to\infty} \frac{f_X^{*2}(x)}{f_X(x)} = \infty,$$

then the distribution with pdf f_X is said to be **light-tailed**. Examples include:

Generalized Inverse Gaussian $(\lambda > 0)$	Poisson
Gamma	Negative Binomial
Exponential	Geometric

For light-tailed distributions, the ratio $f_X^{*2}(x)/f_X(x)$ increases without bound, thus one would get negative values for the joint pdf in eq. (11.2). As a result, eq. (11.2) does not define a proper joint distribution for Poisson, negative binomial or gamma marginals.

For heavy-tailed and moderate-tailed distributions, the ratio $f_X^{*2}(x)/f_X(x)$ is bounded and thus there is a feasible range $(-a, b)$ for ω such that eq. (11.2) does define a proper joint pdf.

In the collective risk model, the right-tail behavior of the aggregate loss distribution is essentially determined by the heavier of the frequency and severity components. If the claim severity distribution is sub-exponential, then the right tail of the compound distribution is governed by the severity distribution and thus is also sub-exponential.

Although eq. (11.2) is not well defined for Poisson frequencies, it can be used to construct multivariate compound Poisson distributions as long as the severity distribution has a heavy right-tail. For example, if each portfolio has a compound Poisson-Pareto distribution, then the aggregate loss distribution for each individual portfolio has a Pareto-type tail.

In practical situations, severity distributions are usually subject to the policy limits. Above the policy limits, the ratio $f_X^{*2}(x)/f_X(x)$ are not defined because $f_X(x)$ is zero but $f_X^{*2}(x)$ may be positive. Thus, for censored severities, eq. (11.2) does not yield a proper joint distribution at the far right tail. In the next section we will show that, in most practical situations, this problem does not seem to appear in the aggregate loss distribution (as opposed to the joint probability distribution).

Remark: For the k-dimensional joint ch.f. in eq. (11.1) we have

$$f_{X_1,\cdots,X_k}(x_1,\cdots,x_k) = f_{X_1}(x_1)\cdots f_{X_k}(x_k)\left[1+\sum_{i<j}\omega_{ij}\left(1-\frac{f_{X_i}^{*2}(x_i)}{f_{X_i}(x_i)}\right)\left(1-\frac{f_{X_j}^{*2}(x_j)}{f_{X_j}(x_j)}\right)\right].$$

Note that the range of ω_{ij}, such that eq. (11.1) defines a proper joint distribution, becomes further restricted as the dimension k increases.

12 Aggregation of Correlated Risks with Given Marginals and Covariance Matrix

Consider the problem of aggregating correlated risk portfolios: We are given the marginal probability distributions for each individual risk portfolio, along with some summary statistics such as the covariance matrix between individual risk portfolios. However, we have no or little knowledge regarding their (multivariate) joint distribution. In order to combine the correlated risk portfolios, we could first construct a multivariate distribution from the given marginals and their covariance matrix. The aggregation of the correlated risks can then be done by using the multivariate distribution. Although *theoretically* there might exist infinitely many multivariate distributions for the given set of marginals and their covariance matrix, finding one specific multivariate distribution can be *practically* difficult.

On the other hand, from the given set of marginals and their covariance matrix, we can easily obtain the first two moments of the aggregate loss distribution. Then we can use some two parameter distributions to approximate the aggregate loss distribution by matching the first two moments. For instance, normal, log-normal, gamma all have two parameters. However, this approximation only utilizes the first two moments, and otherwise completely ignores the information contained in the entire marginal distributions. If we keep the first two moments fixed, we may expect that a set of heavy-tailed marginals would also result in a heavy-tailed aggregate loss distribution. In this section, we propose a method which utilizes the entire marginal distributions.

Suppose we are given k correlated risks $X_1, \cdots, X_k$ with given marginal distributions and their covariance coefficients (ω_{ij}). In the aggregation of correlated risk portfolios, X_j represents the aggregate loss distribution for the j-th portfolio. Without knowing their joint probability distribution we have

$$\begin{aligned} \mathrm{E}[X_1 + \cdots + X_k] &= \mathrm{E}[X_1] + \cdots + \mathrm{E}[X_k] \\ \mathrm{Var}[X_1 + \cdots + X_k] &= \mathrm{Var}[X_1] + \cdots + \mathrm{Var}[X_k] + 2\textstyle\sum_{i<j} \omega_{ij}\mathrm{E}[X_i]\mathrm{E}[X_j]. \end{aligned}$$

Inspired by the multivariate construction in eq. (11.1), now we define a univariate ch.f. as follows

$$\phi_Z(t) = \phi_{X_1}(t) \cdots \phi_{X_k}(t) \left\{ 1 + \sum_{i<j} \omega_{ij} \left[1 - \phi_{X_i}(t)\right] \left[1 - \phi_{X_j}(t)\right] \right\}. \tag{12.1}$$

A random variable Z with a ch.f. in eq. (12.1) has the following characteristics:

- Z has the same mean and variance as that of $X_1 + \cdots + X_k$.
- When eq. (11.1) defines a proper joint distribution for $(X_1, \cdots, X_k)$, then Z has the same probability distribution as that of $X_1 + \cdots + X_k$.
- Even if eq. (11.1) does not define a proper joint distribution for $(X_1, \cdots, X_k)$, Z may still have a proper probability distribution. For example, let X_1 and X_2 both have an exponential distribution with mean 1. Although eq. (11.1) does not define a proper bivariate distribution, the univariate ch.f. by eq. (12.1):

  $$\phi_Z(t) = (1+\omega)(1-it)^{-2} - 2\omega(1-it)^{-3} + \omega(1-it)^{-4}$$

  corresponds to a pdf

  $$f_Z(x) = [(1+\omega) - \omega x + x^2/6]xe^{-x},$$

  which is well-defined for $\omega \in (0, 1/2)$.

For the practical purpose of aggregating correlated risks, we can directly work with the univariate ch.f. in eq. (12.1), without being too concerned about whether the joint ch.f. in eq. (11.1) defines a proper joint distribution. We can numerically calculate the distribution of Z by eq. (12.1). If no negative probabilities are encountered, it would suggest that eq. (12.1) defines a proper probability distribution and thus can be used to approximate the aggregate loss distribution. Nevertheless, when eq. (11.1) does not define a proper *joint* distribution, according to our theory in the previous section, *caution* should be exercised and the following two questions should be analyzed:

- How appropriate is this approximation method ?
- How reasonable are the covariance coefficients ?

Now we give an example of the use of eq. (12.1) in combining correlated risks.

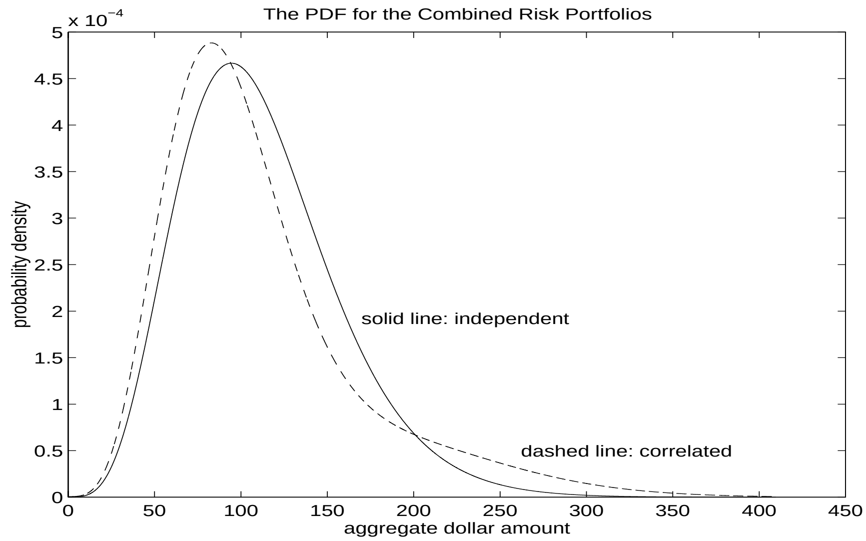
Example 12.1 Consider three correlated risk portfolios with aggregate loss distributions specified as follows.

- Portfolio 1 has a Poisson($\lambda = 10$) frequency and a Pareto($\alpha = 2.4$, $\beta = 5$) severity. The severity distribution is subject to a policy limit of \$20.
- Portfolio 2 has a NB($\alpha = 4$, $\beta = 2$) frequency and a Pareto($\alpha = 1.5$, $\beta = 4$) severity. The severity distribution is subject to a policy limit of \$30.
- Portfolio 3 has a NB($\alpha = 3$, $\beta = 3$) frequency and a log-normal($\mu = 1.2$, $\sigma = 0.7$) severity. The severity distribution is subject to a policy limit of \$15.
- The covariance coefficients are $\omega_{12} = \omega_{13} = 0.2$ and $\omega_{23} = 0.1$.

Eq. (12.1) was used to combine the three correlated risk portfolios. Appendix C provides a pseudo-code for implementing the FFT. Figure 12.1 shows the pdf for the combined risk portfolios in this correlation model, as compared to the pdf in the case of independence between risk portfolios. Here are some descriptive statistics of the aggregate loss distribution for the combined correlated risk portfolios:

$$
\begin{aligned}
\text{Mean} &= 111.72 \\
\text{Standard Deviation} &= 58.36 \\
\text{90th percentile} &= 191.75 \\
\text{95th percentile} &= 234.00 \\
\text{99th percentile} &= 306.50
\end{aligned}
$$

Figure 12.1: Aggregation of three risk portfolios



13 Conclusions

This paper has presented a set of tools for modeling and combining correlated risks. A number of correlation structures were generated using copula, common mixture, component and distortion models. A good understanding of the claim generating process should be helpful in choosing a model as well as in selecting correlation parameters. These correlation models are often specified by (i) the joint cumulative distribution function (i.e. a copula), or (ii) the joint characteristic function. The copula construction leads to efficient simulation techniques which can be readily implemented on a spread sheet. The characteristic function specification leads to simple methods of aggregation by using Fast Fourier Transforms.

In some situations, the correlation structure between risks with given marginal distributions is not fully specified. Instead, only some summary statistics such as the covariance matrix is available. For those cases, we proposed a simple yet general method for combining the correlated risks by FFT. Sometimes this method of aggregation is not based on a proper joint distribution, thus care should be exercised when it is used.

In the high-dimension world of multivariate variables, we may encounter very diverse correlation structures. Regardless of the complexity of the situation, Monte Carlo simulation can always be employed in an analysis of the correlation risk. For instance, in some situations the frequency and severity variables are correlated. With the assistance of Monte Carlo simulation, the common mixture model in section 4 can be adapted to describe the association between the frequency and severity random variables, if both depend on the same external parameter. This external parameter may be chosen to represent the Richter scale of an earthquake, the velocity of wind speed, or several scenarios of legal climate, etc., depending on the underlying claim environment.

Dependency has always been a fascinating research subject as well as part of reality. A good understanding of the impact of correlation on the aggregate loss distribution is essential for the dynamic financial analysis of an insurance company. It is hoped that the set of tools developed in this paper will be useful to actuaries in quantifying the aggregate risks of a financial entity. It is also hoped that this research will stimulate more scientific investigations on this subject in the future.

Bibliography

Burden, R.L. and Faires, J.D. (1989). *Numerical Analysis*, 4th edition, PWS-KENT Publishing Company, Boston.

Cook, R.D. and Johnson, M.E. (1981). "A family of distributions for modelling non-elliptically symmetric multivariate data", *Journal of the Royal Statistical Society* **B 43**, 210-218.

Crow, E.L. and Shimizu, K. (1988). "Lognormal Distributions: Theory and Applications", Marcel Dekker, Inc., New York, NY.

Embrechts, P. and Veraverbeke, N. (1982). "Estimates for the probability of ruin with special emphasis on the possibility of large claims", *Insurance: Mathematics and Economics*, **1**, 55-72.

Fishman, G.S. (1996). *Monte Carlo: Concepts, Algorithms, and Applications*, Springer-Verlag, New York.

Frees, E.W. and Valdez, E.A. (1997). "Understanding relationships using copulas", University of Wisconsin - Madison, working paper, July 1997.

Genest, C. and Mackay, J. (1986). The joy of copulas: bivariate distributions with uniform marginals, *The American Statistician*, **40**, No. 4, 280-283.

Heckman, P.E. and Meyers, G.G. (1983). "The calculation of aggregate loss distributions from claim severity and claim count distributions", *Proceedings of the Casualty Actuarial Society*, LXX, 22-61.

Hesselager, O, Wang, S., and Willmot, G.E. (1998) "Exponential and scale mixtures and equilibrium distribution", *Scandinavian Actuarial Journal*, to appear.

Hogg, R. and Klugman, S. (1984). *Loss Distributions*, John Wiley, New York.

Johnson, N., Kotz, S. (1972). "Distributions in Statistics: Continuous Multivariate Distributions", John Wiley & Sons, Inc., New York.

Johnson, N., Kotz, S. (1975). "On some generalized Farlie-Gumbel-Morgenstern distributions", *Communications in Statistics*, 415-427.

Johnson, N., Kotz, S. (1977). "On some generalized Farlie-Gumbel-Morgenstern distributions -II Regression, Correlation and Further Generalizations", *Communications in Statistics*, 485-496.

Johnson, N., Kotz, S. and Balakrishnan, N. (1997). "Discrete Multivariate Distribution", John Wiley & Sons, Inc., New York.

Johnson, M.E. (1987). *Multivariate Statistical Simulation*, Wiley, New York.

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

Marshall, A.W. and Olkin, I. (1988). "Families of multivariate distributions", *Journal of the American Statistical Association* **83**, 834-841.

Oakes, D. (1989). Bivariate survival models induced by frailties, *Journal of the American Statistical Association* **84**, 487-493.

Panjer, H.H. (1981). Recursive evaluation of a family of compound distributions, *ASTIN Bulletin* **12**, 22-26.

Panjer, H.H. and Willmot, G.E. (1992). *Insurance Risk Models*, Society of Actuaries, Schaumburg, IL.

Robertson, J. (1992). "The computation of aggregate loss distributions", *Proceedings of the Casualty Actuarial Society*, LXXIX, 57-133.

Vaupel, J.W., Manton, K.G. and Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality, *Demography* **16**, No. 3, 439-454.

Venter, G.G. (1983). "Transformed beta and gamma distributions and aggregate losses", *Proceedings of the Casualty Actuarial Society*, LXX, 156-193.

Willmot, G.E. (1987). The Poisson-Inverse Gaussian as an alternative to the negative binomial, *Scandinavian Actuarial Journal*, 113-127.

Appendix A. An Inventory of Univariate Distributions

A.1 Counting distributions

- The Poisson distribution, Poisson(λ), $\lambda > 0$, is defined by a probability function:

$$p_n = \Pr\{N = n\} = e^{-\lambda}\frac{\lambda^n}{n!}, \quad n = 0, 1, 2, \cdots.$$

  It has a pgf

$$P_N(t) = \mathrm{E}[t^N] = e^{\lambda(t-1)}$$

  and $\mathrm{E}[N] = \mathrm{Var}[N] = \lambda$.

- The negative binomial distribution, NB(α, β), $\alpha, \beta > 0$, has a probability function:

$$p_n = \Pr\{N = n\} = \frac{\Gamma(\alpha + n)}{\Gamma(\alpha)n!}\left(\frac{1}{1+\beta}\right)^{\alpha}\left(\frac{\beta}{1+\beta}\right)^{n}, \quad n = 0, 1, 2, \cdots.$$

  It has a pgf

$$P_N(t) = [1 - \beta(t-1)]^{-\alpha}$$

  with $\mathrm{E}[N] = \alpha\beta$ and $\mathrm{Var}[N] = \alpha\beta(1+\beta)$.

  When $\alpha = 1$, the negative binomial distribution NB($1, \beta$) is called the geometric distribution.

- The Poisson-Inverse-Gaussian distribution, P-IG(β, μ), has a pgf

$$P_N(t) = \mathrm{E}[t^N] = \exp\left\{-\frac{\mu}{\beta}[\sqrt{1 + 2\beta(1-t)} - 1]\right\}.$$

  It can be verified that $\mathrm{E}[N] = \mu$ and $\mathrm{Var}[N] = \mu(1+\beta)$. The probabilities can be calculated via a simple recursion (Willmot, 1987):

$$p_n = \frac{2\beta}{1+2\beta}\left(1 - \frac{3}{2n}\right)p_{n-1} + \frac{\mu^2}{n(n-1)(1+2\beta)}p_{n-2}, \quad n = 2, 3, \cdots,$$

  with starting values

$$p_0 = e^{-\frac{\mu}{\beta}[\sqrt{1+2\beta}-1]}, \quad p_1 = \frac{\mu}{\sqrt{1+2\beta}}p_0.$$

A.2 Continuous distributions

- The exponential distribution, Exponential(λ), is defined by

$$S(x) = 1 - F(x) = e^{-\lambda x}, \quad x > 0,$$

  with $\mathrm{E}[X] = 1/\lambda$ and $\mathrm{Var}[X] = 1/\lambda^2$.

- The gamma distribution, Gamma(α, β), $\alpha, \beta > 0$, has a pdf

$$f(x) = \frac{x^{\alpha-1} e^{-x/\beta}}{\beta^\alpha \Gamma(\alpha)}, \quad x > 0.$$

It has a mgf

$$M_X(t) = \mathrm{E}[e^{tX}] = (1 - \beta t)^{-\alpha}$$

and $\mathrm{E}[X] = \alpha\beta$ and $\mathrm{Var}[X] = \alpha\beta^2$.

- The Pareto distribution, Pareto(α, β), $\alpha, \beta > 0$, has a survivor function

$$S(x) = 1 - F(x) = \left(\frac{\beta}{x+\beta}\right)^\alpha = (1 + x/\beta)^{-\alpha}.$$

The mean $\mathrm{E}[X] = \frac{\beta}{\alpha-1}$ exists only if $\alpha > 1$.

- The Weibull distribution, Weibull(β, τ), $\beta, \tau > 0$, has a survivor function

$$S(x) = 1 - F(x) = e^{-(x/\beta)^\tau}$$

with $\mathrm{E}[X] = \beta\Gamma(1 + \tau^{-1})$ and $\mathrm{E}[X^2] = \beta^2 \left[\Gamma(1 + 2\tau^{-1})\right]$.

- The Inverse Gaussian distribution, IG(β, μ), has a pdf

$$f(x) = \mu(2\pi\beta x^3)^{-\frac{1}{2}} \exp\left\{-\frac{(x-\mu)^2}{2\beta x}\right\}, \quad x > 0.$$

It has a mgf

$$M(t) = e^{\frac{\mu}{\beta}[1-\sqrt{1-2\beta t}]},$$

and $\mathrm{E}[X] = \mu$ and $\mathrm{Var}[X] = \mu\beta$.

- The Exponential-Inverse-Gaussian distribution, E-IG(β, μ), has a survivor function:

$$S(x) = 1 - F(x) = e^{\frac{\mu}{\beta}\left\{1-(1+2\beta x)^{\frac{1}{2}}\right\}}, \quad x > 0,$$

with moments (Hesselager/Wang/Willmot, 1998):

$$\mathrm{E}[X] = \frac{\beta + \mu}{\mu^2}, \qquad \mathrm{Var}[X] = \frac{5\beta^2 + 4\beta\mu + \mu^2}{\mu^4}.$$

- The log-normal distribution, LN(μ, σ^2), has a pdf

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{1}{2}\left[\frac{\log(x)-\mu}{\sigma}\right]^2}, \qquad x > 0,$$

with $\mathrm{E}[X] = \exp\left[\mu + \frac{\sigma^2}{2}\right]$ and $\mathrm{Var}[X] = \exp\left[2\mu + \sigma^2\right]\ \left[\exp(\sigma^2) - 1\right]$.

A.3 Parameter uncertainty and mixture models

In modeling insurance losses, actuaries/underwriters are called upon to pick a frequency distribution and a severity distribution based on past claim data and their own judgement. Actuaries/underwriters are fully aware of the presence of parameter uncertainty in the assumed models. As a way of incorporating parameter uncertainty, mixture models are often employed.

- The most popular frequency distributions are the negative binomial family of distributions. In modeling claim frequency, the negative binomial $\mathrm{NB}(\alpha, \beta)$ can be interpreted as a mixed Poisson model, where the Poisson parameter λ has a $\mathrm{Gamma}(\alpha, \beta)$ distribution. This can be seen from the pgf

$$P_N(t) = \mathrm{E}[t^N] = \mathrm{E}_\lambda[\mathrm{E}(t^N|\lambda)] = \mathrm{E}_\lambda[e^{\lambda(t-1)}] = M_\lambda(t-1) = \{1-\beta(t-1)\}^{-\alpha}$$

- A popular claim severity distribution is the Pareto distribution which has a thick right tail representing large claims. The $\mathrm{Pareto}(\alpha, \beta)$ distribution can be interpreted as a mixed exponential distribution, where the exponential parameter λ has a $\mathrm{Gamma}(\alpha, \frac{1}{\beta})$ distribution. This can be seen from the survivor function

$$S(x) = \mathrm{E}_\lambda[e^{-\lambda x}] = M_\lambda(-x) = (1 + x/\beta)^{-\alpha} = \left(\frac{\beta}{\beta + x}\right)^\alpha .$$

- A more flexible family of claim severity distribution is the Burr distribution (including Pareto as a special case). The $\mathrm{Burr}(\alpha, \beta, \tau)$ distribution can be expressed as a Weibull-Gamma mixture. This can be seen from the survivor function

$$S(x) = \mathrm{E}_\lambda[e^{-\lambda x^\tau}] = M_\lambda(-x^\tau) = (1 + x^\tau/\beta)^{-\alpha} = \left(\frac{\beta}{\beta + x^\tau}\right)^\alpha .$$

  The $\mathrm{Burr}(\alpha, \beta, \tau)$ family includes the $\mathrm{Pareto}(\alpha, \beta)$ as a special member when $\tau = 1$.
  For $\tau > 1$ the $\mathrm{Burr}(\alpha, \beta, \tau)$ distribution has a lighter tail than its $\mathrm{Pareto}(\alpha, \beta)$ counterpart.
  For $\tau < 1$ the $\mathrm{Burr}(\alpha, \beta, \tau)$ distribution has a thicker tail than its $\mathrm{Pareto}(\alpha, \beta)$ counterpart.

A.4 Log-normal distributions

A.4.1 Univariate log-normal distributions

The normal distribution, $N(\mu, \sigma^2)$, has a pdf

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2}, \qquad -\infty < x < \infty.$$

It has a moment generating function

$$M_X(t) = \mathrm{E}[e^{tX}] = \exp\left[\mu t + \frac{1}{2}\sigma^2 t^2\right].$$

If $X \sim N(\mu, \sigma^2)$, then $Y = e^X$ has a log-normal distribution with a pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\sigma y} e^{-\frac{1}{2}\left[\frac{\log(y)-\mu}{\sigma}\right]^2}, \qquad y > 0.$$

The moments of Y are

$$\mathrm{E}[Y^n] = \exp\left[n\mu + \frac{n^2\sigma^2}{2}\right], \quad n = 1, 2 \cdots.$$

Specifically,

$$\begin{aligned} \mathrm{E}[Y] &= \exp\left[\mu + \frac{\sigma^2}{2}\right] \\ \mathrm{Var}[Y] &= \exp[2\mu + \sigma^2]\,[\exp(\sigma^2) - 1] \\ \mathrm{E}[Y - \mathrm{E}Y]^3 &= \exp\left[3\mu + \frac{3\sigma^2}{2}\right]\,[\exp(3\sigma^2) - 3\exp(\sigma^2) + 2]\,. \end{aligned}$$

A.4.2 Bivariate log-normal distributions

Let X_1 and X_2 have a bivariate normal distribution with joint pdf

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho)^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right)\right]\right\}.$$

Then X_1 and X_2 have marginal distributions $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, respectively. (X_1, X_2) has a covariance matrix

$$\begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

where ρ is the correlation coefficient between X_1 and X_2. Note that $\rho = 1$ if, and only if, $\Pr\{X_1 = aX_2 + b\} = 1$ with $a > 0$.

Now consider the variables $Y_1 = \exp(X_1)$ and $Y_2 = \exp(X_2)$. Note that $\log(Y_1Y_2)$ has a $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2)$ distribution. We have

$$\begin{aligned} \mathrm{Cov}[Y_1, Y_2] &= \mathrm{E}[Y_1Y_2] - \mathrm{E}[Y_1]\mathrm{E}[Y_2] \\ &= \exp\left\{(\mu_1 + \mu_2) + \frac{1}{2}[\sigma_1^2 + \sigma_2^2 + 2\rho\sigma_1\sigma_2]\right\} - \exp\left[\mu_1 + \sigma_1^2 + \mu_2 + \sigma_2^2\right] \\ &= \exp\left[\mu_1 + \mu_2 + \frac{1}{2}(\sigma_1^2 + \sigma_2^2)\right]\ \{\exp(\rho\sigma_1\sigma_2) - 1\}. \end{aligned}$$

Therefore, the correlation coefficient of Y_1 and Y_2 is

$$\rho_{Y_1,Y_2} = \frac{\exp(\rho\sigma_1\sigma_2) - 1}{\sqrt{\exp(\sigma_1^2) - 1}\sqrt{\exp(\sigma_2^2) - 1}},$$

where $\rho = \rho_{X_1,X_2}$.

A.4.3 Multivariate log-normal distributions

Consider a vector $(X_1, \cdots, X_n)'$ of positive random variables. Assume that $(\log X_1, \cdots, \log X_n)'$ has an n-dimensional normal distribution with mean vector and variance-covariance matrix

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix}, \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1n} \\ \vdots & \vdots & \vdots \\ \sigma_{n1} & \cdots & \sigma_{nn} \end{pmatrix},$$

respectively.

The distribution of $(X_1, \cdots, X_n)'$ is said to be an n-dimensional log-normal distribution with parameters $(\boldsymbol{\mu}, \Sigma)$ and denoted by $\Lambda_n(\boldsymbol{\mu}, \Sigma)$. The probability density function of $\mathrm{X} = (X_1, \cdots, X_n)'$ having $\Lambda_n(\boldsymbol{\mu}, \Sigma)$ is (see Crow and Shimizu, 1988, Chapter 1):

$$f(x_1, \cdots, x_n) = \begin{cases} \dfrac{1}{\sqrt{(2\pi)^n |\Sigma|}\ \prod_{i=1}^n x_i} \exp\left\{-\frac{1}{2}(\log \boldsymbol{x} - \boldsymbol{\mu})'\Sigma^{-1}(\log \boldsymbol{x} - \boldsymbol{\mu})\right\}, & \boldsymbol{x} \in (0, \infty)^n \\ 0, & \text{otherwise.} \end{cases}$$

From the moment generating function for the multivariate normal distribution we have

$$\mathrm{E}[X_1^{s_1} \cdots X_n^{s_n}] = \exp\left(\boldsymbol{s}'\boldsymbol{\mu} + \frac{1}{2}\boldsymbol{s}'\Sigma\boldsymbol{s}\right),$$

where $\boldsymbol{s} = (s_1, \cdots, s_n)'$. Specifically, we have for any $i = 1, 2, \cdots, n$,

$$\mathrm{E}[X_i^r] = \exp\left(r\mu_i + \frac{1}{2}r^2\sigma_{ii}\right)$$

and for any $i, j = 1, 2, \cdots, n$,

$$\mathrm{Cov}[X_i, X_j] = \exp\left\{\mu_i + \mu_j + \frac{1}{2}(\sigma_{ii} + \sigma_{jj})\right\}\{\exp(\sigma_{ij}) - 1\}.$$

A simulation of this multivariate log-normal distribution can be easily achieved by first generating a sample from a multivariate normal distribution and then taking the logarithms.

Remark: This appendix gives the conventional way of defining multivariate lognormal distributions. Alternatively, for the same set of given marginals and their covariance matrix, we can define a multivariate lognormal distribution by using the joint ch.f.

$$\phi_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \phi_{X_1}(t_1) \cdots \phi_{X_k}(t_k) \left\{1 + \sum_{i<j} \omega_{ij} \left[1 - \phi_{X_i}(t_i)\right] \left[1 - \phi_{X_j}(t_j)\right]\right\},$$

which leads to simple aggregation by FFT.

Appendix B. More On Copulas & Simulation Methods

In this appendix, we discuss in greater detail about the construction of copulas and the associated simulation techniques. For simplicity, we confine our discussions to the bivariate case. Our discussions here can be readily extended to higher dimensions ($k > 2$).

B.1 Bivariate copulas

A bivariate **copula** refers to a joint cumulative distribution function $C(u, v) = \Pr\{U \leq u; V \leq v\}$ with uniform marginals $U, V \sim$Uniform[0,1]. It links the marginal distributions to their multivariate joint distribution. Let $S_{X,Y}(x, y)$ be a joint survivor function with marginals S_X and S_Y. Then there is a copula C such that

$$S_{X,Y}(x, y) = C(S_X(x), S_Y(y)), \qquad \text{for all } x, y \in (-\infty, \infty).$$

Conversely, given any marginals S_X, S_Y and a copula C, $S_{X,Y}(x, y) = C(S_X(x), S_Y(y))$ defines a joint survivor function with marginals S_X and S_Y. Furthermore, if S_X and S_Y are continuous, then C is unique.

Note that $S_X(X)$ and $S_Y(Y)$ are uniformly distributed random variables. The association between X and Y can be described by the association between uniform variables $U = S_X(X)$ and $V = S_Y(Y)$. In terms of Monte Carlo simulation, one can first generate a sample pair (u_i, v_i) from the bivariate uniform distribution of (U, V), and then invert them to get a sample pair $x_i = S_X^{-1}(u_i)$ and $y_i = S_Y^{-1}(v_i)$ for (X, Y).

Note that $F_{X,Y}(x, y) = C(F_X(x), F_Y(y))$ and $S_{X,Y}(x, y) = C(S_X(x), S_Y(y))$ may imply different bivariate distributions. Here we assume that a copula is applied to the survivor functions unless otherwise mentioned.

B.2 Distortion of the joint survivor function

Let $g : [0, 1] \to [0, 1]$ be an increasing function with $g(0) = 0$ and $g(1) = 1$. If $S_{X,Y}(x, y)$ is a joint survivor function with marginals S_X and S_Y, then $g[S_{X,Y}(x, y)]$ defines another joint survivor function with marginals $g \circ S_X$ and $g \circ S_Y$. If we assume that, after applying a distortion g, $g[S_{X,Y}(x, y)]$ has uncorrelated marginals:

$$g[S_{X,Y}(x, y)] = g[S_X(x)]\, g[S_Y(y)],$$

then we have

$$S_{X,Y}(x, y) = g^{-1}\left(g[S_X(x)] \cdot g[S_Y(y)]\right), \tag{B.1}$$

which corresponds to the copula

$$C(u, v) = g^{-1}[g(u)g(v)]. \tag{B.2}$$

If we let $h(t) = -\log g(t)$, then equation (B.1) gives the following relation: $S_{X,Y}(x, y) = h^{-1}(h[S_X(x)] + h[S_Y(y)])$, which gives the Archimedean family of copulas (see Genest and Mackay, 1986).

For a bivariate copula C, the Kendall's tau is

$$\tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1.$$

If a copula $C = g^{-1}(g(u)g(v))$ is defined by a distortion g, then

$$\tau = 1 + 4\int_0^1 \frac{g(t)\log g(t)}{g'(t)}dt.$$

Example B.1. The distortion function $g(t) = \exp\{1 - t^{-\alpha}\}$, $\alpha > 0$, corresponds to the Clayton family of copulas with

$$C_\alpha(u,v) = \left\{u^{-\alpha} + v^{-\alpha} - 1\right\}^{-\frac{1}{\alpha}}, \tag{B.3}$$

$$\begin{aligned} C_\infty(u,v) &= \lim_{\alpha\to\infty} C_\alpha(u,v) &= \min[u,v] \\ C_0(u,v) &= \lim_{\alpha\to 0+} C_\alpha(u,v) &= uv. \end{aligned}$$

Thus, C_∞ and C_0 are the copulas for the Fréchet upper bounds and the independent case, respectively.

Example B.2. The distortion function $g(t) = \exp\{-(-\log t)^\alpha\}$, $\alpha \geq 1$, corresponds to Hougaard family of copulas with

$$C_\alpha(u,v) = \exp\left\{-\left[(-\log u)^\alpha + (-\log v)^\alpha\right]^{\frac{1}{\alpha}}\right\}. \tag{B.4}$$

$$\begin{aligned} C_\infty(u,v) &= \lim_{\alpha\to\infty} C_\alpha(u,v) &= \min[u,v] \\ C_1(u,v) &= uv. \end{aligned}$$

Thus, C_∞ and C_1 are the copulas for the Frechet upper bounds and the independent case, respectively.

Example B.3. The distortion function $g(t) = \frac{\alpha^t - 1}{\alpha - 1}$, $\alpha > 0$, corresponds to the Frank family of copulas with

$$C_\alpha(u,v) = [\log\alpha]^{-1}\log\left\{1 + \frac{(\alpha^u - 1)(\alpha^v - 1)}{\alpha - 1}\right\}, \qquad 0 < \alpha < \infty. \tag{B.5}$$

$$\begin{aligned} C_\infty(u,v) &= \lim_{\alpha\to\infty} C_\alpha(u,v) &= \max[u+v-1,0] \\ C_0(u,v) &= \lim_{\alpha\to 0+} C_\alpha(u,v) &= \min[u,v] \\ C_1(u,v) &= \lim_{\alpha\to 1} C_\alpha(u,v) &= uv. \end{aligned}$$

Thus, C_∞, C_0 and C_1 are the copulas for the Frechet lower and upper bounds and the independent case, respectively.

B.3 Common frailty models

Vaupel, Manto and Stallard (1979) introduced the concept of frailty in their discussion of a heterogeneous population. Each individual in the population is associated with a frailty, r. The frailty varies across individuals and thus is modeled as a random variable R with cdf $F_R(r)$. The conditional survival function of lifetime T, given r, is

$$\Pr\{T > t | R = r\} = B(t)^r,$$

in which $B(t)$ is the base line survivor function (for a standard individual with $r = 1$). The unconditional survivor function for a heterogeneous population is

$$\Pr\{T > t\} = \int_0^\infty B(t)^r dF_R(r) = M_R(\log B(t)),$$

where M_R is the mgf of R.

Oakes (1989) uses a bivariate frailty model to describe associations between two random variables as follows. Assume that X and Y both can be modeled by frailty models

$$\begin{aligned} S_X(x) &= \textstyle\int_0^\infty B_1(x)^r dF_R(r) &&= M_R(\log B_1(x)), \quad \text{and} \\ S_Y(y) &= \textstyle\int_0^\infty B_2(y)^r dF_R(r) &&= M_R(\log B_2(y)), \end{aligned}$$

respectively, in which B_1 and B_2 are the baseline survivor functions. Assume that X and Y are conditionally independent, given frailty $R = r$. However, X and Y are associated as they depend on the common frailty variable R. The bivariate frailty model yields the following joint survivor function

$$\begin{aligned} S_{X,Y}(x, y) = \Pr\{X > x, Y > y\} &= \int_0^\infty [B_1(x) \cdot B_2(y)]^r dF_R(r) \\ &= M_R(\log[B_1(x) \cdot B_2(y)]). \end{aligned}$$

For $g(u) = \exp[M_R^{-1}(u)]$, we have

$$g[S_{X,Y}(x, y)] = B_1(x) \cdot B_2(y) = g[S_X(x)] \cdot g[S_Y(y)].$$

In other words, a bivariate frailty model yields a joint distribution which can also be obtained by using the distortion function g.

Example B.4. Assume that the frailty R has a Gamma distribution with $M_R(z) = (\frac{1}{1-z})^{1/\alpha}$, $(\alpha > 0)$. Then $M_R^{-1}(u) = 1 - u^{-\alpha}$, and $g(u) = \exp[1 - u^{-\alpha}]$. Therefore, the common Gamma frailty model yields the Clayton family of copulas given by (B.3):

$$S_{X,Y}(x, y) = \left\{ \frac{1}{S_X(x)^\alpha} + \frac{1}{S_Y(y)^\alpha} - 1 \right\}^{-\frac{1}{\alpha}}, \qquad 0 < \alpha < \infty.$$

This family is particularly useful for constructing bivariate Burr (including Pareto) distributions (see Johnson and Kotz, 1972, page 289).

Example B.5. If the frailty R has a stable distribution with $M_R(z) = \exp\{-(-z)^{1/\alpha}\}$, $\alpha \geq 1$, the corresponding joint survivor function is given by eq. (B.4):

$$S_{X,Y}(x, y) = \exp\left\{ - [(-\log S_X(x))^\alpha + (-\log S_Y(y))^\alpha]^{\frac{1}{\alpha}} \right\}.$$

This family of copulas is particularly useful for constructing bivariate Weibull (including exponential) distributions.

Example B.6. If the frailty R has a logarithmic (discrete) distribution on positive integers with $M_R(z) = [\log \alpha]^{-1} \log[1 + e^z(\alpha - 1)]$, then we get the Frank family of copulas given by eq. (B.5).

Marshall and Olkin (1988) proposed a simulation algorithm for copulas with a frailty construction. This algorithm is applicable to copulas with arbitrary dimensions ($k \geq 2$):

Step 1. Generate a value r from the random variable R having mgf M_R.

Step 2. Generate independent uniform (0,1) numbers $U_1, \cdots, U_k$.

Step 3. For $j = 1, \cdots, k$, set $U_j^* = M_R(r^{-1} \log U_j)$ and calculate $X_j = S_j^{-1}(U_j^*)$.

B.4 The Morgenstern copula

The Morgenstern copula is defined by

$$C(u, v) = uv[1 + \alpha(1-u)(1-v)], \quad (-1 \le \alpha \le 1).$$

This copula cannot be generated by distortion or frailty models.

The following simulation algorithm for the Morgenstern copula can be found in Johnson (1987, p. 185):

Step 1. Generate independent uniform variables V_1, V_2 and set $U_1 = V_1$.

Step 2. Calculate $A = \alpha(2U_1 - 1) - 1$ and $B = [1 - \alpha(2U_1 - 1)]^2 + 4\alpha V_2(2U_1 - 1)$.

Step 3. Set $U_2 = 2V_2/(\sqrt{B} - A)$.

For the Morgenstern copula, the Kendall's tau is

$$\tau(\alpha) = \frac{2}{9}\alpha, \quad (-1 \le \alpha \le 1),$$

which is limited to the range $(-\frac{2}{9}, \frac{2}{9})$. Thus, the Morgenstern copula can be used only in situations with weak dependence.

An extension of the Morgenstern copula to arbitrary dimensions has been given by Johnson and Kotz (1975, 1977).

B.5 Summary and comments

The following table lists the most commonly used bivariate copulas. Except for the reverse monotone copula, they can readily be extended to multivariate cases ($k > 2$).

A Summary of Popular Copulas

Associated Names	Function Form $C(u,v)$	Kendall's tau
Independence	uv	0
Common monotone	$\min[u, v]$	1
Reverse monotone	$\max[u + v - 1, 0]$	-1
Cook-Johnson	$[u^{-1/\alpha} + v^{-1/\alpha} - 1]^{-\alpha}, \quad (\alpha > 0)$	$\frac{1}{1+2\alpha}$
Clayton	$[u^{-\alpha} + v^{-\alpha} - 1]^{-1/\alpha}, \quad (\alpha > 0)$	$\frac{\alpha}{\alpha+2}$
Frank	$\log_\alpha \left\{1 + \frac{(\alpha^u-1)(\alpha^v-1)}{\alpha-1}\right\}, \quad (0 < \alpha < \infty)$	*
Farlie-Gumbel-Morgenstern	$uv[1 + \alpha(1-u)(1-v)], \quad (-1 \le \alpha \le 1)$	$\frac{2}{9}\alpha$
Gumbel-Hougaard	$\exp\left\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{1/\alpha}\right\}, \quad (\alpha \ge 1)$	$1 - \alpha^{-1}$
Normal	$H(\Phi^{-1}(u), \Phi^{-1}(v))$ ** $\quad (-1 \le \rho \le 1)$	$\frac{2}{\pi}\arcsin(\rho)$

* For the Frank copula, $\tau(\alpha) = 1 + \frac{4}{-\log(\alpha)} \left[\frac{1}{-\log(\alpha)} \int_0^{-\log(\alpha)} \frac{t}{e^t - 1} dt - 1 \right]$.
** H is the joint cdf for a bivariate standard normal distribution with a correlation coefficient ρ.
Note that the Cook-Johnson copula with parameter α is the same as the Clayton copula with parameter $1/\alpha$.

Frees and Valdez (1997) have just written a good survey paper on the use of copulas including the associated simulation techniques. In general, the use of copulas permits simple implementation by Monte Carlo simulations, thus can be used in aggregating correlated risks.

B.6 The use of mixed copulas

Assume that joint cdf's $F_{X_1,\cdots,X_k}$ and $G_{X_1,\cdots,X_k}$ have the same marginals $F_{X_1}, \cdots, F_{X_k}$. Then the mixed joint cdf

$$(1-q)\ F_{X_1,\cdots,X_k}(t_1, \cdots, t_k) + q\ G_{X_1,\cdots,X_k}(t_1, \cdots, t_k), \quad (0 < q < 1),$$

also has marginal cdf's $F_{X_1}, \cdots, F_{X_k}$. For this mixed joint pgf, we have

$$\tau[X_i, X_j] = (1-q)\ \tau^F[X_i, X_j] + q\ \tau^G[X_i, X_j],$$

where τ^F and τ^G represent the Kendall's tau implied by the joint cdf's F and G, respectively. In particular, if we let $F_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \prod_{j=1}^k t_j$ represent the independent copula, and $G_{X_1,\cdots,X_k}(t_1, \cdots, t_k) = \min[t_1, \cdots, t_k]$ represent the comonotone copula, then $\tau[X_i, X_j] = q$.

The mixture of joint cdf's can be used to adjust, up or down, the Kendall's tau. For example, if we feel that a common mixture joint cdf F would give too strong of a correlation, then we can mix it with an independent joint cdf G. If if we feel that a common mixture joint cdf F would give too weak of a correlation, then we can mix it with a comonotone joint cdf G'.

Appendix C: Computer Pseudo-Code for Example 12.1

In this appendix, we give a pseudo-code for the *FFT* method of combining the correlated risk portfolios in Example 12.1.

1. First we choose the "number of points" for the *FFT* computation at $4096 = 2^{12}$. Note that this is the maximum "number of points" for the Microsoft Excel *FFT* routine.

2. Then we determine approximately the "entire range" of the aggregate loss amount for the combined risk portfolios. In our example, we may set the entire range at ($0, $800), which is roughly seven times the aggregate mean loss of $110, so that with negligible probability the aggregate loss amount not lying in this range.

3. Next, we choose a span (i.e. interval width) at $0.20, which is roughly equal to $800/4096.

4. Discretize the severity distribution of each individual risk portfolio using a span at $h = \$0.20$. For a severity distribution F_x with a policy limit equal to $M \cdot h$, apply the following discretization method:

$$\begin{aligned}
\Pr\{X = 0 \cdot h\} &= 0, \\
\Pr\{X = 1 \cdot h\} &= F_x\left(\frac{3}{2}h\right), \\
\Pr\{X = j \cdot h\} &= F_x\left(\left(j + \frac{1}{2}\right)h\right) - F_x\left(\left(j - \frac{1}{2}\right)h\right), \quad \text{for } j = 2, 3, \ldots, M-1, \\
\Pr\{X = M \cdot h\} &= 1 - \sum_{j=0}^{M-1} \Pr\{X = j \cdot h\}
\end{aligned}$$

5. Note that the severity distribution for the three risk portfolios are subject to policy limits of \$20, \$30 and \$15, respectively. Given that the span was chosen at \$0.20, the maximum severity points with non-zero probabilities are 100, 150 and 75, respectively. *It is critical to pad (i.e. add) enough zeros to the discretized severity vectors so that each severity vector has the same length, 4096 in this case, as the target aggregate loss distribution.* Let f_1, f_2 and f_3 represent the discretized severity vectors for the three risk portfolios, each of which of length 4096.

6. Perform FFT on each of the severity vectors, f_1, f_2 and f_3. Let $\phi_1 = FFT(f_1)$, $\phi_2 = FFT(f_2)$, and $\phi_3 = FFT(f_3)$ represent the resulting vectors (each of length 4096) from the FFT.

7. Let P_1, P_2 and P_3 be the probability generating functions of the frequency variables for Portfolios 1, 2 and 3, respectively. In our example,

$$\begin{aligned}
P_1(t) &= \exp[10(t-1)] \\
P_2(t) &= [1 - 2(t-1)]^{-4} \\
P_3(t) &= [1 - 3(t-1)]^{-3}.
\end{aligned}$$

Apply P_1, *element by element*, to the FFT transformed vector ϕ_1; and likewise for portfolios 2 and 3. Let $\psi_1 = P_1(\phi_1)$, $\psi_2 = P_2(\phi_2)$ and $\psi_3 = P_3(\phi_3)$.

8. Define new vectors ψ_4 and φ by the following *element by element* operations:

$$\psi_4 = 1 + 0.2(1-\psi_1)(1-\psi_2) + 0.2(1-\psi_1)(1-\psi_3) + 0.1(1-\psi_2)(1-\psi_3)$$

and

$$\varphi = \psi_1\psi_2\psi_3\psi_4.$$

9. Finally, perform an inverse FFT on φ and let $g = IFFT(\varphi)$. Note that g is a probability vector with a span of \$1/8, which approximates the aggregate loss distribution for the combined risk portfolios.

10. As warning, one should exercise caution in the selection of the span, h, for discretization of the severity distributions. A too large span would affect the accuracy of the discretization. A too small span may produce some "wrapping" (non-zero probabilities at the high points near 4096).