Operational risk modelling: aggregating loss distributions using copulas

Capturing the dependence structure between business line/risk event types is an extremely important step for any serious attempt to model operational risk. In this article we show how this can be achieved by using a powerful statistical technique known as copulas

UNDER THE NEW Basel Capital Accord, the advanced measurement approach (AMA) offers financial institutions a measurement methodology for operational risk that allows for modelling with very few constraints. The output of the AMA is a capital charge based on a 99.9% confidence level and one-year holding period covering the scope of the op risk definition. Given the heterogeneous nature of operational risk, breaking it down into manageable categories will facilitate the modelling efforts. These categories can typically represent the loss event types as proposed by the Basel Committee:

- Internal fraud
- External fraud
- Employment practices and workplace safety
- Client, products and business practices
- Damage to physical assets
- Business disruption and system failures
- Execution, delivery and process management

For a better understanding of its working processes, a financial institution needs to categorise its operations by business line (BL), which will also enhance the management and modelling efforts of operational risk.

It is not difficult to understand that dependence effects between business line/loss event-types play a crucial role in operational risk measurement. The dependence issue is crucial, particular when capital needs to be allocated fairly among different business units. Hence, there is a necessity for applying techniques that take this dependence structure into consideration. Put in another way, we need to be able to answer questions such as: "What is the probability that two operational processes are subject to failure?". From a modelling point of view this means we need to focus on the joint distribution of several outcomes, ie, a multivariate distribution.

As loss distributions are generated per loss event type and BL, these need to be aggregated to arrive at a firm-wide capital charge. It is not difficult to understand that Dependence effects between loss event types play an important role, for the aggregation of loss distributions to realistically reflect the firm-wide op risk exposures The dependence issue is crucial, particular when capital needs to be allocated fairly among different business units. Hence, there is a necessity for applying techniques that take this dependence structure into consideration. Put in another way, we need to be able to answer questions such as: "What is the probability that two loss event type are subject to failure?" For example, a loss in the "business disruption and system failures" loss event type will most likely also trigger a loss in the "execution, delivery & process management" loss event type, therefore a dependence structure needs to be constructed. From a modelling point of view, this means we need to focus on the joint distribution of several outcomes, ie, a multivariate distribution. From a modelling point of view, this means we need to focus on the joint distribution of several outcomes, ie, a multivariate distribution.

Previously, modelling was confined to the assumption that all distributions are normal distributions, which could then easily be aggregated. However, this was not a very realistic assumption and the aggregated capital charge lacked any meaningful relation

Figure 1. Constructing bivariate distribution with Poisson marginals and normal copula

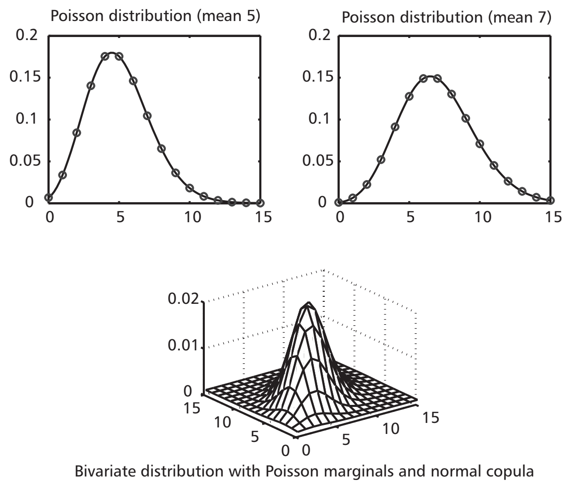
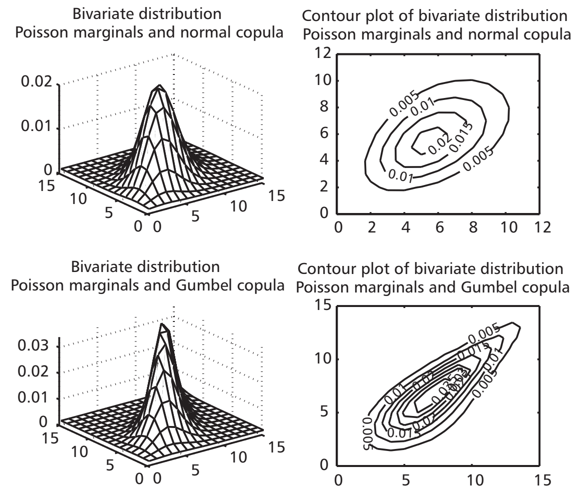


Figure 2. A comparison of tail dependents between normal and Gumbel copula



with the actual levels of the operational risk.

The implementation of a parametric approach, involving gathering risk data and modelling a frequency distribution, to estimate a regulatory risk capital amount, at a certain 99.9% confidence interval similar to credit risk modelling, implies that a dependence structure between different business line/risk event types can be modelled by building a multivariate distribution for the loss distributions. However, except for some specific cases the extension of the multivariate distribution to more than two dimensions is far from straightforward. This is particularly true for the Poisson distribution, which represents one of the most used distributions for fitting operational risk losses.

The difficulties in constructing a multivariate distribution force us to look for alternative techniques for modelling joint events and dependence structures. A possible alternative is to construct a multivariate distribution using a copula function.

Copulas: a possible solution

Copulas have proved to be a useful tool for building and simulating multivariate distributions with given marginals. When dealing with random variables, copulas can be interpreted as cumulative distribution functions with uniform marginals. More precisely, a function C is a copula function if it satisfies the following properties:

1. $C: [0,1]^N \rightarrow [0,1]$;
2. C is grounded and N-increasing;
3. C has cumulative marginal distributions C_i that satisfy the following relationship $C_i = C(1,...,1,u,1,...,1) = u$ for each $u \in [0,1]$

The first property guarantees that the copula function is defined only in the unit hypercube (ie, the N dimensional cube with every dimension defined in [0,1]). The second property asserts that the copula function starts from 0 and that is non-decreasing. The meaning and implication of the third property will become clearer throughout this article.

Given the properties above, copulas can serve as a useful function for building multivariate distributions. If $F_1,...,F_N$ are cumulative univariate distribution functions, then $C(F_1(x_1),...,F_N(x_N))$ is a multivariate cumulative distribution function with marginals $F_1,...,F_N$. This is because the component $F_i(x_i) = u_i$ for each $i = 1,..., N$ is a uniform random variable in [0,1].

Furthermore, as proved in a theorem by Sklar, a multivariate cumulative distribution function with continuous marginals can be uniquely decomposed in the following way:

$$F(x_1,...,x_N) = C(F_1(x_1),...,F_N(x_N))$$

This formula simply states that we can think of any multivariate distribution function as being composed by its univariate marginals linked together by a copula function that gives the dependence structure, ie, the relationship between the marginals.

A direct consequence of Sklar's theorem is the following relationship, which can be used to simulate a specific copula:

$$C(u_1,...,u_N) = F(F_1^{-1}(u_1),...,F_N^{-1}(u_N))$$

So, copulas represent a useful approach to for the understanding and modelling of dependent random variables. They allow us to focus explicitly on the dependence structure by separating the modelling process into two steps:

1. Modelling of the marginals, ie, choice of distribution for each univariate random variable
2. Modelling of the dependence structure, i.e. choice of the copula

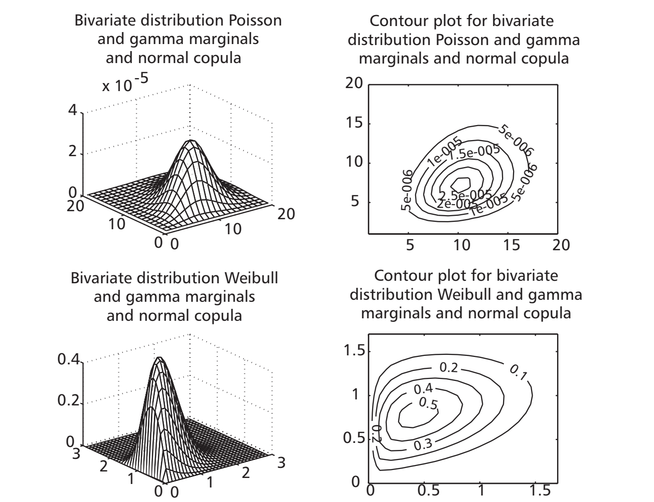
For example, in figure 1, two univariate Poisson distributions with mean 5 and 7 respectively, are linked together by a normal copula to produce a bivariate distribution.

Among the different copula families available, copulas derived from well-known distributions such as the normal and the Student's t are the most useful for practical purposes.

The normal copula is defined as:

$$C_\rho(u_1,...,u_N) = \Phi_{\rho,N}(\Phi_1^{-1}(u_1),...,\Phi_N^{-1}(u_N)),$$

Figure 3. A comparison of bivariate distributions with normal copula and different marginals



where Φ_1^{-1} denotes the inverse of the standard normal cumulative distribution and $\Phi_{\rho,N}$ the N-dimensional normal distribution with correlation matrix ρ. Therefore the normal copula is defined by the knowledge of the correlation matrix. In addition to being derived from the normal distribution, it shares the same tractability of the former but allows for modelling extreme events and different dependency structures than the multivariate normal distribution.

The Student's *t* copula, being obtained from the Student's *t* multivariate distribution satisfies the following relationship:

$$C_{\rho,\nu}(u_1,\ldots,u_N) = T_{\rho,\nu}(t_{1,\nu}^{-1}(u_1),\ldots,t_{N,\nu}^{-1}(u_N))$$

where is the multivariate Student's *t* distribution with correlation matrix and degrees of freedom and is the inverse of the standard Student's *t* cumulative distribution.

The Student's *t* copula shares the same properties of the Student's tdistribution, hence as $\nu \to \infty$ the Student's t copula tends to the normal copula. However, for finite degrees of freedom *v* the Student's tcopula possesses a tail dependency that is independent of the marginal distributions. This tail dependency implies that the Student's tcopula is suitable to model large or extreme events that occur simultaneously.

Archimedean copulas are another important class of copulas. These copulas are easy to construct and possess a great variety of dependence structure. Archimedean copulas are able to define the dependence structure within the data better than the Gaussian copula because they can be fitted to the data. This simply means that the best type of Archimedean copula can be inferred directly from the data at hand.

An example of Archimedean copulas is the Gumbel copula, given by the following function:

$$C_\theta(u,v) = \exp\left\{-[(-ln(u)^\theta + (-ln(v)^\theta]^{1/\theta}\right\}$$

Where the parameter θ is defined in the interval [1, ∞]. The Gumbel copula is non-symmetric and possesses a strong right-tail dependency (see figure 2).

The flexibility and different dependence structures given by the different classes of copulas can be seen in figure 2, where different copulas are compared. In the first row, the bivariate distribution and contour plot are shown for Poisson marginals and normal copula. While in the second row, the dependence structure is given by the Gumbel copula. From figure 2 it is possible to see how the two contour plots vary using a different copula, ie, a different dependence structure.

Hence, the copula approach has several advantages when compared to the approach that uses a specific multivariate distribution to capture the joint probabilities. First of all, it is more flexible. Each risk loss event type can be modelled by a different univariate distribution and the existence of several copula families allows for the capture of a broad range of dependence structures. Secondly, the dependence structure is modelled separately and explicitly. This gives the possibility to describe the dependency between random variables with greater accuracy.

In other words, the risk we are trying to model can be separated in two parts by using copulas. In the first part, the individual risks can be modelled by specific marginal distributions, while in the second part the dependence structure is captured by the copula.

Among the copulas we have discussed, it is certainly the normal copula that is the simplest and easiest to implement. Despite the lack of tail dependency, by modelling the marginal distributions with different marginals, the normal copula is able to capture extreme events and non-normal dependencies. For example, figure 3 shows that by using different marginals and maintaining the normal copula assumption, different bivariate distributions with different tail structures can be built.

Summary

To conclude, the copula framework provides the modeller with a powerful tool that in a risk-sensitive and realistic way can capture the dependence structure between the loss event types and BLs. Failure to capture this dependence will result in having an AMA methodology in place that doesn't properly reflect the operational risks faced by a financial institution and subsequently will produce a misleading capital charge. **OpRisk**

Niklas Hageback (niklas.hageback@kpmg.co.uk) and Leonardo Petz (leonardo.petz@kpmg.co.uk) are managers in KPMG's Financial Service Advisory Practice in London

References

The Basel Committee on Banking Supervision, 2003
The New Basel Capital Accord – Third Consultative Paper.
Buoyé E et al, 2000
Copulas for Finance. A Reading Guide and Some Applications.
Working Paper Groupe de Recherche Operationelle,
Credit Lyonnais, Paris.
Cruz M, 2002
Modeling, Measuring and Hedging Operational Risk.
John Wiley & Son, New York.