Hardly any risk models properly estimated the negative financial results of the past years. Even the most extreme risk calculations – typically expressed in such terms as ‘Value at Risk’ and required ‘Economic Capital’ - made no allowance for such fierce fluctuations. To some extent this is due to a lack of ‘negative’ data. But there is considerable room for improvement in the modeling of risks, and in particular their dependency. Analyst Bas Tammens explains how this matter should be tackled.

Improve your risk calculations

Multivariate normal probability distributions are frequently used to calculate risk standards of the underlying variables (such as interest rates or returns on equity). A multivariate probability distribution offers a possibility to model dependency between different variables as well as the properties of the individual series. However, using the multivariate normal distribution inappropriately can result in an underestimate of the risk, especially in case of extreme (simultaneous) situations.
By defining the characteristics of the individual variables as accurately as possible (statistically), it is possible to create a more realistic picture of the risks that a company faces. The relationship and dependence between the variables must then be determined correctly in order to obtain a realistic total picture of all the risks.
This article focuses on the structure of multivariate probability distributions and the application of this concept as part of risk management. This concept can for example be applied in the following fields: Summing Value at Risk trade limits, Simulation of economic variables and Aggregation of risk types.

Theory

A multivariate probability distribution consists of two parts:

1. The (univariate) distributions that reflect the characteristics of the individual variables, such as normal, t or exponentially distributed.
2. The copula function that defines the dependence between the different variables (stochasts), such as a Gaussian, t or Archimedean copula. When modeling dependencies of variables, copulas are frequently considered relatively complex. In practice, however, copulas are simpler than people think and implicitly they are already being used regularly.

Whereas the characteristics such as skewness or kurtosis of the (univariate) distributions are frequently known, the characteristics of the copulas are not as a rule. A Gaussian copula defines the linear dependency between stochasts. A t copula can model the dependency between variables in the tail of the distributions. This means that extreme events occur simultaneously more often than expected purely on the basis of the correlation. It is also possible to use the Archimedean copulas to model non-linear connections.

To grasp the structure of a multivariate distribution function, the bivariate normal distribution is described here. This consists of normally distributed univariate random variables and a Gaussian copula (defined by a correlation matrix). Figure 1 is a diagrammatic representation. So, if you use a multivariate normal distribution in the modeling of a certain phenomenon, you already use a (Gaussian) copula implicitly!

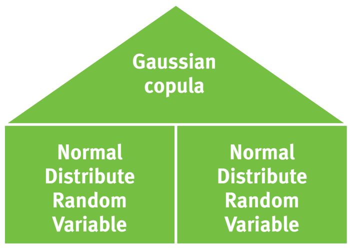

Figure 1. Structure of bivariate normal distribution.

If the assumption of normality does not hold for one of the univariate series, the consequence is that the multivariate normal distribution is not the right model to base the calculations on. It is also possible that the marginal distributions are normally distributed, but that the structure of dependence is not defined by the Gaussian copula. In both situations, a tailor-made multivariate probability distribution can provide a solution. An advantage of a tailor-made multivariate distribution compared with a standard (normal) multivariate probability distribution is that the univariate distributions can be different. These distributions are linked by a copula that describes the dependency structure.

“There is considerable room for improvement in the modeling of risks, and in particular their dependency.”

Example

Say a company wants to determine the total risk that it faces as a result of market and credit risks within various investment portfolios. The first step is to prepare a multivariate probability distribution based on historical data using the steps described above. Next, data is simulated using this distribution. Based on this data, the risk model and the selected reliability (often dependent upon the desired rating of the company), it is then possible to determine the total risk, expressed for example as Value at Risk.

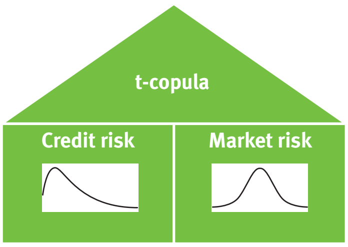

Figure 2. Structure of tailor-made bivariate distribution

Figure 2 shows the marginal distributions based on historical data. For the loss function of the credit risk, the gamma distribution provides the best fit, while the market risk provides the best fit with a t distribution. The data further shows that the extreme events occur relatively often at the same time, so the t copula is more suitable than the Gaussian copula.

To determine the combined risk for the market risk and the credit risk, two simulation steps must be carried out (described in the sidebar). First, the data for the t copula is simulated. Next, this data can be transformed to the area of the individual variables, by means of the inverse cumulative distribution functions of the gamma distribution for the credit risk and the t distribution for the market risk. This simulated data is then used to determine the aggregate loss to obtain for example the value at risk of the combined risk types.

Of course, this bivariate example can be expanded in a simple way to higher dimensions, where for all the individual series (different) distribution function functions can be described and the (multivariate) dependency can be described by a copula function.

Consequences and conclusions

When calculating risk standards, it is crucial to make the right model assumptions. For determination of a multivariate probability distribution, this concerns the choice of the univariate probability distributions and the copula that is used. The impact of wrongly specifying a multivariate distribution can have major consequences. If the bivariate normal distribution is chosen wrongly in the above example, it will have the following consequences:

- the probability of extreme events will be underestimated, because the kurtosis (fatness of the tails) of the selected probability distribution does not correspond with the risk that actually exists;
- the choice of the normal distribution for the credit risk is completely incorrect, as the loss function of the credit risk has only values greater than nil. The normal distribution, however, also gives a probability to events smaller than nil;
- the Gaussian copula underestimates the probability of simultaneous extreme events. As a result, that risk will be underestimated.

Incorrect models by definition provide an incorrect representation of the risks, i.e. the risk standards. Therefore, you should make sure that you have the right specifications of your risk models and the simulations of the underlying (economic) variables. The multivariate normal distribution, which in practice is often wrongly used, can be replaced by a tailor-made multivariate distribution better aligned to the real-life risks that the company faces. This will further improve the risk management of a company and the presented risk standards will increasingly start to represent reality. ‹

IF YOU WISH TO LEARN MORE ABOUT THE MODELING OF RISKS, CONTACT MARTIJN DE GROOT BY PHONE ON +31 (0)35 692 89 89 OR BY E-MAIL AT M.DEGROOT@ZANDERS.EU.

Estimation process

Three steps must be completed in the estimation process in order to prepare a multivariate distribution:

1. Estimate for each variable individually the most applicable (univariate) distribution function.
2. Use the cumulative distribution function of the estimated distributions from step 1 to transform the data to the unity area (the dependence structure described by the copula is defined in this area).
3. Estimate of the copula in the unity area.

Simulation process

When data needs to be simulated from the multivariate tailor-made distribution function, it is necessary to complete the following two steps:

1. Simulate the data from the copula. This data can be found in the unity area.
2. Use the (inverse cumulative) distribution functions that have been estimated for the all the variables separately, to transform the data from the unity area to the area of the original variables.

About the author

Bas Tammens worked for Zanders as an analyst. He studied econometrics at the University of Amsterdam. For his final thesis, Tammens conducted research at a mid-sized Dutch bank/insurer. He studied the use of the concept described in this article for the generation of economic scenarios. Whereas a multivariate t distribution was initially being used, Tammens constructed a tailor-made multivariate distribution that ultimately provided a more realistic picture of the various risks faced by the company. Bas is now associate consultant at Zanders.