Modelling dependencies:
An Overview

Martyn Dorey
Phil Joubert
Coomaren Vencatasawmy

Presented to the Finance & Investment Conference,
19-21 June 2005, Whittlebury Hall, Silverstone

1 Introduction

In 1888, Francis Galton discovered the concept of correlation while doing some quantitative work on heredity. From 1890 onwards a group of persons came forward to fill up the gaps in Galton's work and to extend it in various directions. The most prominent member of this group was Karl Pearson. The product moment formula for correlation coefficient was given by Pearson in 1896 [1]. This is the familiar formula

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma(X)\sigma(Y)} \qquad (1)$$

Since then correlation analysis has been used quite intensively in the social sciences to ascertain the relationship between occurrences of economic or social events. One of the earliest examples of using correlation involved an anthropologist investigating whether some bones belonged to a skeleton by calculating the correlation between the lengths of various bones for each skeleton divided by the length of the skeleton [2]. We will call this type of correlation event correlation.

Qualitatively, the notion of correlation has a much longer history. The first manifestations of spatial data arose in 1686 in the form of data maps which were used in a qualitatively way to infer a physical cause of monsoon rains [3]. The use of spatial models came later in [4] which was concerned with the distribution of particles through a liquid. Although spatial dependence was observed in agricultural fields, most efforts were aimed at removing them prior to analysis [5]. Spatial models have gained popularity in the last decade and have been used in areas such as ecology, geology, climatology and environmental science. In forestry, for example, spatial models are used to model patterns of tree growth.

Temporal correlation, dependence of measurements taken at different points from the same process in time, have grown in use since the 1950's. Although many observations have a time dimension, often temporal correlation is ignored, instead a cross-section of data is analysed. In the past few years, spatio-temporal models have been used to describe dynamic systems such as ecological and climatic phenomena.

Nowadays, the notion of correlation is central to financial theory. The Capital Asset Pricing Model (CAPM) and the Arbitrage Pricing Theory (APT) use correlation as a measure of dependence between different financial instruments [6]. Furthermore, the importance of correlation has often been emphasised in the context of the pricing of derivatives instruments whose pay-offs depend on the joint realisations of several prices or rates. Examples of such derivatives products are basket of options, swaptions and spread options [7]. Although insurance has traditionally been built on the assumption of independence and the law of large numbers has governed the determination of premiums, the increasing complexity of insurance and reinsurance products has led recently to increased interest in the modelling of dependent risks, especially with the emergence of more intricate multi-line products [6]. With the ICA requirements of the FSA, robust and defensible approaches to modelling dependencies are required. In Enterprise Risk Management (ERM) the modelling of dependencies between lines of business is critical.

Dependencies arise when one factor affects more than one variable. The insurance premium cycle can result in the loss ratios of different classes of business moving in the same direction. Concentration of risks in a given sector, for example the energy sector, can result in increased claims such as directors and officers (D&O), errors and omissions (E&O), surety and others. Extreme events, such as hurricanes, can also result in dependencies between classes of business which are unrelated in normal conditions. Dependencies can be both very tricky to model and also not intuitive. Quite often dependencies occurs at different levels- for example if the risk profile of a particular class of business is broken into different sections by size and compared to another class of business, then different dependences can be found depending on the section compared. In ecology complex dependence structures, built up from several factors, are quite common. In the financial world the dependence structures vary with the volatility of the market. The estimation of dependence in non-volatile conditions can be very tricky depending on the amount of data available, the quality of the data available and complexity of the dependence structure. Quite often, it best to impose a dependence structure rather than trying to empirically determine and validate a structure. One should always look out for spurious structures that may be due to biases in the sampling approach used.

There is a need to understand the different alternatives available for modelling dependencies and assess the methods available to parameterise and validate such models. Finding a model that is robust to certain conditions can be very tricky. In this paper we will look at some of the issues in modelling dependencies. We will describe some of the measures used to estimate correlation and some of the approaches for modelling dependencies while explaining some of the pitfalls inherent in them.

2 Spurious correlations

Given, the long history of applying correlation analysis in other fields, it would be a pity to ignore what is already known about the empirical estimation of correlation structures. Correlation is only one particular measure of stochastic dependence among many. It is the canonical measure for spherical and elliptical distributions and being a linear measure it cannot capture the non-linear dependence relationships that exist among most real life factors [6]. It is always important to bear in mind that correlation does not imply causation.

It is possible to obtain a significant value for a coefficient when in reality the two functions are absolutely uncorrelated. Spurious correlation can be due to standard ways of processing data, for example one should be very wary of correlating ratios or indices. For example, two financial ratios may both be influenced by inflation. The two may show a strong correlation, but this is simply an artefact of inflation. Stripping inflation out may result in two uncorrelated indices.

Fallacies can also be caused by mixing different records. Suppose that a drug is effective only on women and the population tested is predominantly men. In this case a spuriously high correlation is obtained only because some women are present in the sample [8, 9, 10]. Correlating time series can also produce spurious correlation especially due to noise or finiteness of the time series [11].

3 Measures of correlation

Suppose we have a pair of datasets (X,Y) and we wish to empirically determine the correlation coefficient between them. There are a number of methods for estimating the correlation coefficient and we will look at some of the most common ones. The most common approach is the Pearson's moment approach. This assumes a bivariate normal distribution and a linear relationship. This coefficient is given by (1)

For non-elliptical and non-linear correlation coefficients, one can use the Kendall's tau or the Spearman's rho. Note that while the former can give values which are very different from Pearson r, the latter can be numerically identical to Pearson r, especially if it is applied blindly. The Kendall's tau is defined as

$$\tau = 1 - \frac{4Q}{n(n^2 - 1)}$$

where Q is the number of inversions between the rankings of x and y. An inversion is any pair of objects (i,j) such that r_i-r_j and r'_i-r'_j have opposite signs

The Spearman's rho is defined as

$$\upsilon = 1 - \frac{6D^2}{n(n^2 - 1)}$$

where

$$D^2 = \sum_1^n \left(r_{y_i} - r_{x_i} \right)^2$$

where r denotes the rank.

Note that the standard error of these coefficients can be estimated using bootstrapping [12]. There are other alternatives for estimating the correlation when non-linearity is suspected. In [13] the observation that close values of X gives rise to close values of Y and thus the statistic is given by

$$K=\Sigma\{I[ABS(X_i-X_j)<\delta]I[ABS(Y_i-Y_j)<\varepsilon]\}$$

A large value of K, which can be compared to a reference distribution, will indicate strong relationship. The Moran's coefficient [14] replaces the Xs and Ys by their ranks and then calculates a moment correlation coefficient using the ranks. The coefficient is then compared to a reference distribution. Another rank correlation that is robust to outliers is introduced in [15] and in [16] another robust correlation coefficient is proposed and is based on the median.

4 Modelling and simulating multivariate distributions

One approach for simulating correlated multivariate distributions is through the correlation matrix. Using historical data a correlation matrix is determined, which is then decomposed to simulate the random numbers.

Note that the correlation matrix should be positive semi-definite and that the multivariate distribution will be the same as a weighted linear combination of the variables constituting the multivariate distribution. For spherical and elliptical distributions such as Poisson, gamma, normal, inverse Gaussian, the distribution of the weighted sum of linear components is the same as the distribution of the components, except for the correlation matrix. For other distributions, finding the multivariate distribution can be very tricky.

Although this approach is intuitive, restrictions on the distributions and positive definiteness of the correlation matrix can be problematic. Furthermore, it is quite difficult to structure complicated correlation structures through this method. A number approaches have been proposed for rendering the correlation positive definite. In [17] the matrix is decomposed into its eigenvalues and eigenvectors, the highest eigenvalues are chosen and the matrix is reconstruct using them. In [18] a review of existing approaches is given and a new approach is proposed. This consists of decomposing the correlation matrix as a block matrix such that the matrices in the main diagonals are positive definite while the matrices in the other diagonals are transpose of each other. The lower right matrix is then further decomposed in the same way as above. The correlated random numbers are then simulated using the Cholesky decomposition of the final matrix. The process is repeated for each stage of the decomposition and the random numbers are stacked horizontally. Note that this approach can result in a large number of parameters and it might help to parameterise the correlation matrix so that its elements can be determined from a pre-determined function with lower number of parameters.

5 Copulas

One method of modelling dependencies which has become very popular recently is the copula. The word copula is a Latin noun which means ‘a link, tie bond’. Mathematically, a copula is a function which allows us to combine univariate distributions to obtain a joint distribution with a particular dependence structure. The word itself was first employed in a mathematical or statistical sense by Abe Sklar [19].

Sklar’s theorem, which is central to the theory of copulas, states that for a given joint bivariate distribution function and the two marginal distributions, there exists a copula function that relates them. If both marginal are continuous, then the copula function exists. Conversely, if given the copula function and the marginal distribution functions then the joint distribution is given by applying the copula function on the marginal distributions. The theorem describes how functions join together one-dimensional distribution functions to form multivariate distribution functions. Sklar named the function knowing that the word copula is a grammatical term that links a subject and predicate. Sklar’s theorem states that A joint distribution can be expressed as inter-dependency C applied to the individual distributions. More precisely:

Sklar’s theorem

Let F_{xy} be a joint distribution with margins F_x and F_y. Then there exists a function $C:[0,1]^2 \to [0,1]$ such that

$$F_{xy}(x,y) = C(F_x(x), F_y(y)) \qquad (4.1.1)$$

If X and Y are continuous, then C is unique; otherwise, C is uniquely determined on the (range of X) x (range of Y).

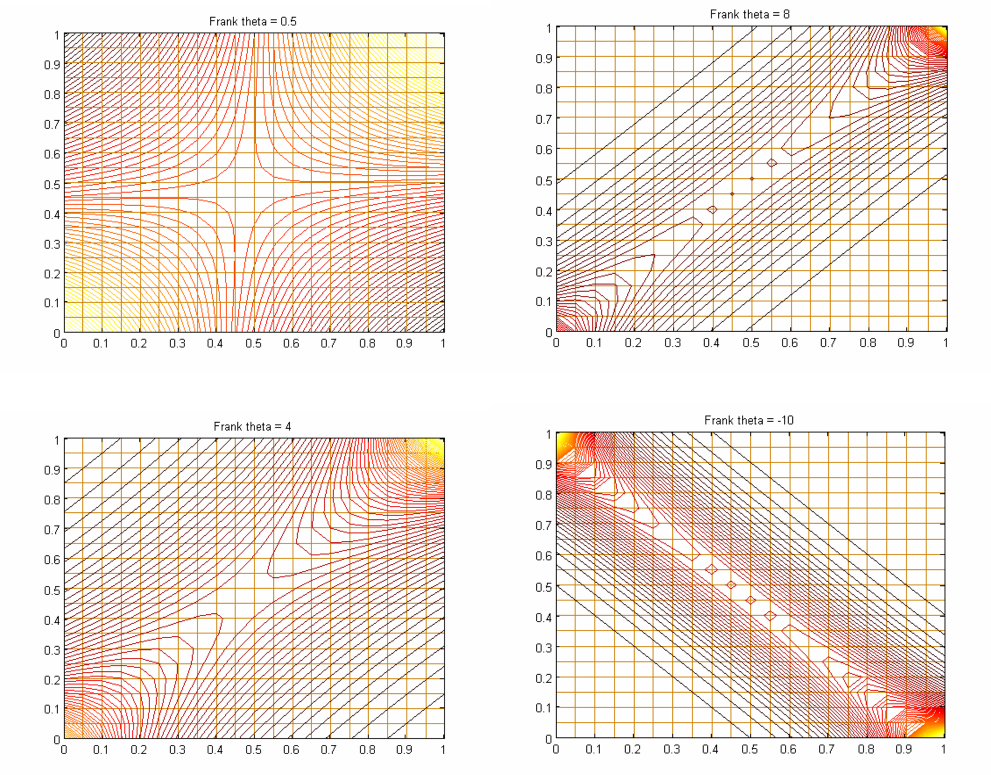
Conversely if C is a copula and X and Y are distribution functions, then the function Fxy defined by 4.1.1 is a joint distribution with margins F_x and F_y.

Using a copula to build a multivariate distribution is flexible because no restrictions are placed on the marginal distributions [20]. For example, if we have two marginal distributions - one with a beta distribution with parameters $\alpha=5$ and $\beta=5$, and the other with a lognormal distribution with parameters $\mu=0$ and $\sigma=1$. Then we can use a copula which is a member of the Frank’s Family and given by

$$C(u,v)=-1*\ln(1+(e^{-\sigma u}+1)(e^{-\delta v}-1)/(e^{-\delta}-1))/\delta$$

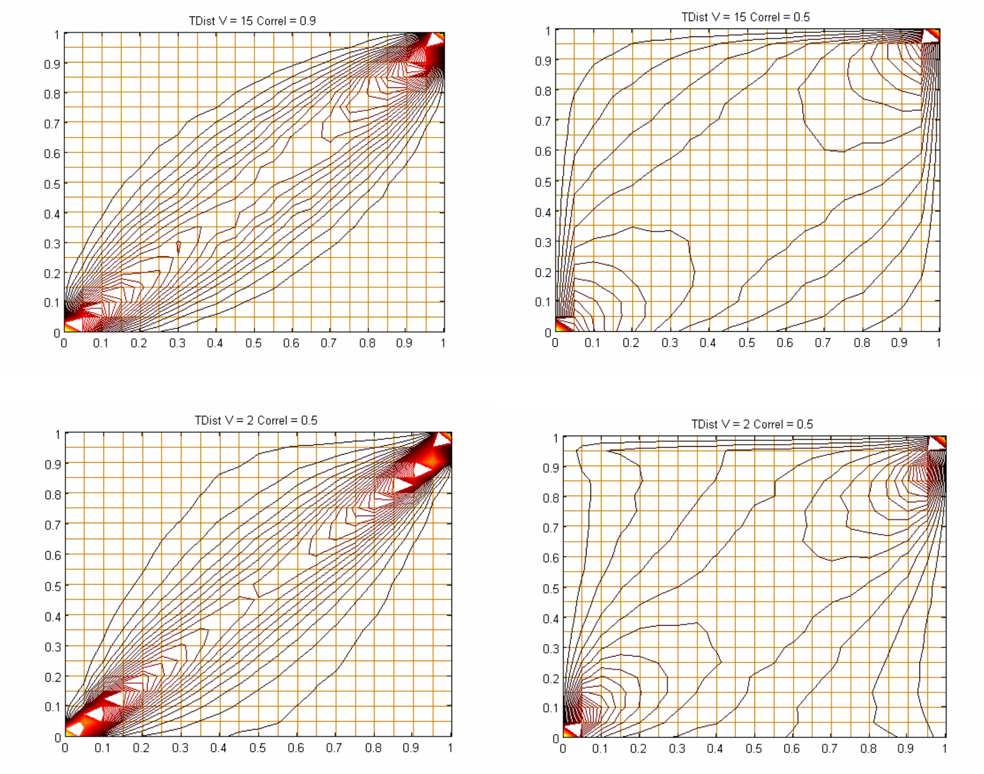
Simply by using equation 4.1.1 we generate a new joint distribution. The parameter δ determines the level of dependence of between the marginals.

Figure 5.1 Contour maps of a Frank Copula function



There exists a range of copulas the most common being the Gumbel copula for extreme distribution, the normal copula for symmetric correlation, the Archimedean copula the t-copula for dependence in the tail [21, 22]. The functional form of the t-copula is somewhat complicated but easy to simulate. However more recent work on the t copula shows that it can be generalised to give asymmetric dependence [23].

Figure 5.2 Sampled contour maps of a T copula function



[24,25] propose reserving the degree-of-freedom parameter as a user-specified simulation input, allowing the user to subjectively induce the extent of tail dependence between assets. A user can approximate a Gaussian by entering a high degree-of-freedom (say 15), or select a lower degree of freedom for higher tail dependence (between 1 or 2). This is useful not only for more traditional VaR analyses, but also for stress tests in which the degree of extreme co-dependence is of critical importance. Certainly the authors agreed approach with this and found a degree of freedom of 2 to offer realistic equity tail dependence.

Estimating parameters from data is more problematic. When the right and left tails are quite different the t-copula would not be indicated, but if only the right tail behaviour is important a fit to that could be sought. The main practical obstacle to the use of the t-copula is that there is only one parameter to control tail association and different pairs of variates might have different tail association [26]. Ways around this have been found by the authors through conditioning, and this will be discussed separately in the future.

Building a multivariate copula

In order to minimise the effort in building a multivariate copula practitioners often use the same copula function for all dependencies. For example, all dependencies are described as a t-distribution where v=2, or all variables are described as Gaussian.

Although a great deal of the literature considers the dependency structure between variables, the practitioner will still have to build the marginal distributions. Different approaches can be taken, such as empirical distributions or parametric best fit. Using empirical distributions results in a cumulative histogram of steps, The discrete nature of the steps is often not desirable. As a result many practitioners start with an empirical distribution and apply a cubic spline or kernel smoothing technique to interpolate between the steps. Consideration also needs to be given to the tails. The tails could be an abrupt minimum or maximum, or they can be fitted using Extreme Value Theory (EVT) related techniques; such as a Gumbel distribution [27]. Fitting an EVT tail would be appropriate for equity returns, but would not be appropriate for unemployment, which has a minimum value of 0% and a maximum of 100%.

In addition the modeller can improve the flexibility of a copula through the smart use of pre-processing, for example [24,25] suggest applying a GARCH filter to give i.i.d. observations. The authors found that a simple normalisation of the form (x-mean)/stdev offers an improvement in the dependency estimation and modelling of tails. Indeed pre-processing would remove the need for complicated copulas designs such as the EV t copula described in [23].

.

Conditioning copulas

From the perspective of the practitioner the ability to condition copulas seems to be a very powerful approach. Fitting the copula to all the data is equivalent to fitting a non-linear regression. Forcing certain variables to take a particular value allows the modeller to generate an expected distribution given a variable value Y. This offers benefits over traditional Markov chains.

The conditional distribution can be defined using copulas by differentiating the copulas with respect to the first argument to get Fy|x(y). In an independent case C(u,v)=uv and the conditional distribution of V given U=u is $C_1(u,v)=v$, where $C_1(u,v)$ denotes d(C(u,v))/du. One can use C_1 to simulate the joint distribution. First simulate a value for U, then simulate a value of V from C_1.

The authors have found that conditioning allows different copula forms to be bolted together in one model, i.e. different tail dependencies. As the scope of this subject is beyond this paper we have decided to address this subject in more detail separately.

6 Case Study: Life company ICA

The available capital of a life company is given by its net assets. The Financial Services Authority (FSA) requires institutions to demonstrate that they are adequately capitalised, i.e. that they have enough capital such that the probability of the company failing over a certain time period is sufficiently small. This is known as an ICA – Individual Capital Assessment. In the following we consider how we might evaluate the situation of a life company, and in particular, how we could use copulas to help us.

The model

The value of capital at any future time can be modelled as a random variable, in fact we could write

$$C_t = C_0 + X_t$$

Where C_0 is the current net assets of the insurer, and X_t is some random variable. (We could equally well take C_0 as the *expected* value of capital at time t, and arrange for X_t to have zero mean).

Clearly the financial health of a life insurer is affected by many different risks. The FSA identifies seven categories of risk:

- Market
- Credit
- Insurance
- Business
- Liquidity
- Operational
- Enterprise

We could further break down market risk for example into the risk factors determining:

- domestic equity returns
- foreign equity returns
- bond returns

- property returns

This would lead us to propose

$$X_t = \mathrm{H}_t \cdot R_t$$

Where H_t and R_t are, respectively, vectors of exposures and random variables representing risks.

For example, for a company with no liabilities, \$100 in capital at time zero invested 75% in equities and 25% in short bonds we would have

$$X = \begin{bmatrix} 0.75C_0 \\ 0.25C_0 \end{bmatrix} \bullet \begin{bmatrix} R^e \\ R^b \end{bmatrix}$$

where R^e would be a random variable of returns with (say) mean 8% and standard deviation 20%, and R^b would be similarly defined.

Now if we assume R is jointly normal, then X_t is a linear combination of normal random variables, and hence is also a normal random variable. In fact, if R ~ $N^d(M,\Sigma)$, then X_t ~N(H.M, H' Σ H) where H' is the transpose of the vector H.

It is then trivial to calculate $P\left[C_t \le 0\right]$, simply by noting that

$$\begin{aligned} P[C_t \le 0] &= P[C_o + H \cdot R_t \le 0] \\ &= P\left[Z \le \frac{-C_0}{H'\Sigma H} \right] \\ &= \Phi\left(\frac{-C_0}{H'\Sigma H} \right) \end{aligned}$$

So all that remains is to estimate H and R.

The catch

Above we assumed that all the components of R_t are (jointly) normal. Then evaluating the individual contributions to capital requirements is trivial. Additionally, we can demonstrate how the lack of perfect correlation between the risks lowers capital requirements. This can be done because the joint normal distribution has normal marginals, and both are well understood.

However, in doing so we would ignore two important facts about risks:

1. In general, individual risk factors will not be well approximated by a normal distribution. A distribution with fat tails would be a more appropriate choice.
2. The joint normal distribution restricts the form of the dependence between the marginal variables. We may wish for example, to model a dependency where good returns are uncorrelated, but poor returns are positively correlated

The solution

We turn to the copula tools discussed above instead. Using a copula we can specify arbitrary marginal distributions for the different risk factors based on empirical estimation and/or on theoretical grounds. Then using a copula we can create a joint distribution for the risk factors. One useful copula for this function would be the t copula which will then be used in our formulae above.

7 Case Study: Pricing CDO’s

Most readers will be familiar with the growth in the market for credit derivatives and structured products based on those instruments. In this section we discuss how copula based models are used to price Collateralised Debt Obligations, or CDO’s. We believe this will be of interest to many actuaries given that many life and pension funds are investing in these products in the search for higher returns.

We emphasise that we are describing what we believe to be an emerging industry standard. However, the methods described below have many theoretical shortcomings, and the standard will no doubt change in this fast developing field [26].

What is a CDO?

We begin by explaining the nature of a CDO. CDO’s are a wrapper around a basket of corporate debt instruments, for example high-yield bonds. The CDO is divided into “tranches”, or levels of seniority, each with a promised coupon. Payments from the underlying instruments are passed on, via a Special Purpose Vehicle (SPV), to the purchasers of the tranches, starting with the most senior level, and proceeding down. If defaults occur in the underlying assets, the senior levels are (at first) unaffected, with the lower levels losing some or all of their investment. More defaults will mean more layers are “burnt through”, but the senior level’s principle and interest payments will continue unless almost all the underlying instruments have defaulted – generally considered to be an unlikely scenario.

This is illustrated in figure 6.1.

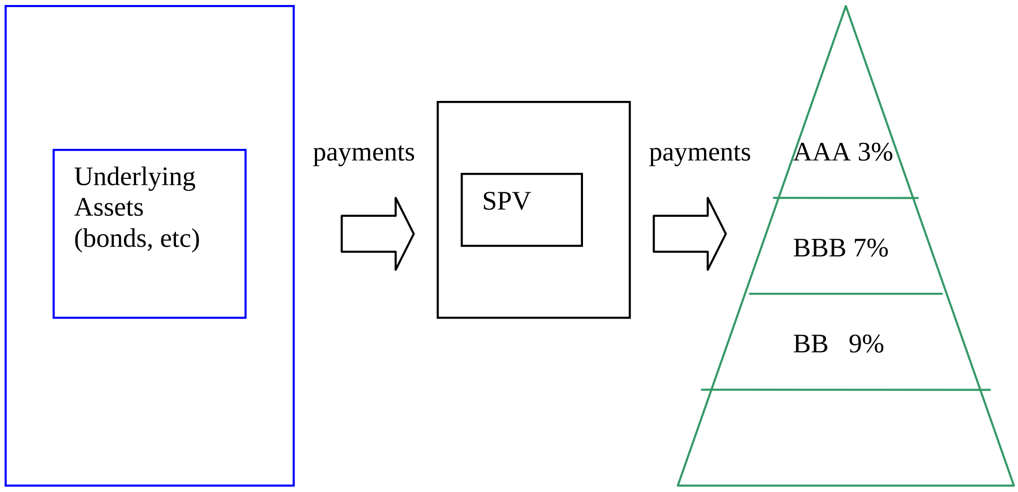


Figure 6.1 A schematic view of a CDO transaction

So a CDO is a method of taking a large number of high risk investments, and creating a several artificial structures (the tranches) with a varying levels of risk.

Synthetic CDO's are designed to be similar from the investor's point of view. They are created, not from the underlying assets themselves, but rather by the use of credit derivatives like credit default swaps.

A credit default swap (CDS) is essentially an insurance contract on a corporate debt: one party pays a regular premium, while the other party agrees to pay an amount if a specific credit event occurs – generally the default of a specific company on a particular debt. The premium is set such that the (risk-neutral or market implied) expected value of payment on default is equal to the value of the premium payments.

In a synthetic CDO the default swap premiums are paid to the SPV, who passes them on to the various tranche holders. The tranche holder's principle will be used to make any necessary payments on default; again the lower level tranches lose their principle before the more senior tranches. From the investor's point of view there is no difference between a synthetic and standard CDO – the cashflows are identical, and the risks are triggered by the same events (default of specified names).

These products are often said to be opaque, the risks poorly identified, understood and priced. We hope to shed some light on the matter, by setting out what we understand has become the market standard method of pricing tranches.

The pricing problem

Pricing a CDO, either synthetic or pure, comes down to calculating the (joint) probabilities of default of the underlying instruments. The problem is that we are attempting to price a basket of credit instruments which may not have independent risks of default. Although the correlation of defaults may be small, it can have a sizeable impact on the risk and price of a tranche.

Additionally, the individual risks of default are traded directly in the corporate bond and credit derivatives market. We need our CDO tranche prices to be consistent with the individual bond and default swap prices if we are to avoid arbitrage opportunities.

Note that, as in all pricing problems, it is not the real-world probabilities that are needed, but the risk-neutral probabilities [28].

A Copula based solution.

If we think of the prices of individual bonds and CDS's as reflecting the *marginal* risks of default, and the price of a CDO tranche as reflecting the *joint* risk of defaults, we see immediately that a solution involving copulas is indicated.

Given that the marginal prices (and hence probabilities of default) are observable, we could assume a copula and then either: by observing the price of the CDO, infer the relevant correlation structure, or, by estimating the correlation structure exogenously, calculate the fair price of the CDO.

The copulas most used by market practitioners are the standard Gaussian, the one-factor Gaussian, and the Clayton copula. We describe in more detail how and why they are used below.

Standard Gaussian copula method

The Gaussian copula is used to generate Monte Carlo simulations of the defaults of the underlying instruments, which are then used to price the CDO [29]. The correlation structure used is the pair-wise asset correlation as used, for example, in

CreditMetrics. Often equity correlation, as derived from historical time series, is used as a proxy for asset correlation.

The observed marginal default distributions, together with the correlations and the Gaussian copula, define the joint default distribution. Pricing can then be done by simulating from the joint distribution, and assessing the payouts in each simulation, as per normal Monte Carlo pricing methods.

We summarise by figure 6.2:

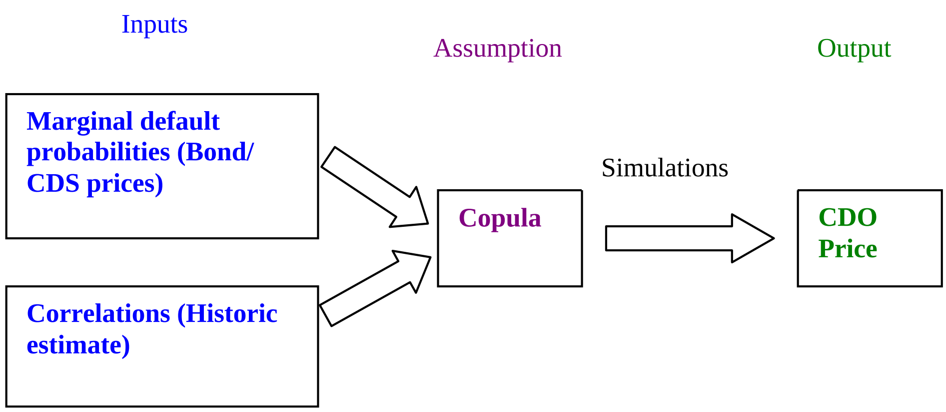


Figure 6.2 Pricing a CDO using a Gaussian Copula

This method is useful since it can be used to price CDO's accurately, given the appropriate marginal distributions and the asset correlation parameters. The disadvantages are that it requires a large number of inputs (the correlation parameters – CDO's can contain hundreds of names) and the computation time required for the Monte Carlo simulation can be onerous. Finally, given a market price for the CDO tranche, we are unable to "invert" this price to find the implied correlation parameters.

One Factor Gaussian and Clayton

Those familiar with the pricing of credit risk will recall the notion of one-factor models [30]. These models postulate that defaults are dependent through a single random factor, often identified with the state of the economy, or some other macro-economic variable. Conditional on the value of that variable, defaults are held to be independent. So during a recession we have more defaults than during a boom, but if we know we are in a recession, company A and company B will default independently (but both with a higher probability than in a boom).

This model further assumes that the underlying portfolio consists of a large number of homogenous assets.

Because the defaults are conditionally assumed to be independent, the conditional joint density function factorises. The unconditional joint density function can then be found by performing a one-dimensional integration over the possible values of the factor. This integration must usually be done numerically, but the computation time and effort is far smaller than for the Monte Carlo method above.

In this model the marginal default probabilities are correlated with the factor variable by a common amount. Since we no have a single correlation factor, we can invert market (observed) prices for CDO tranches, and solve for the "implied correlation".

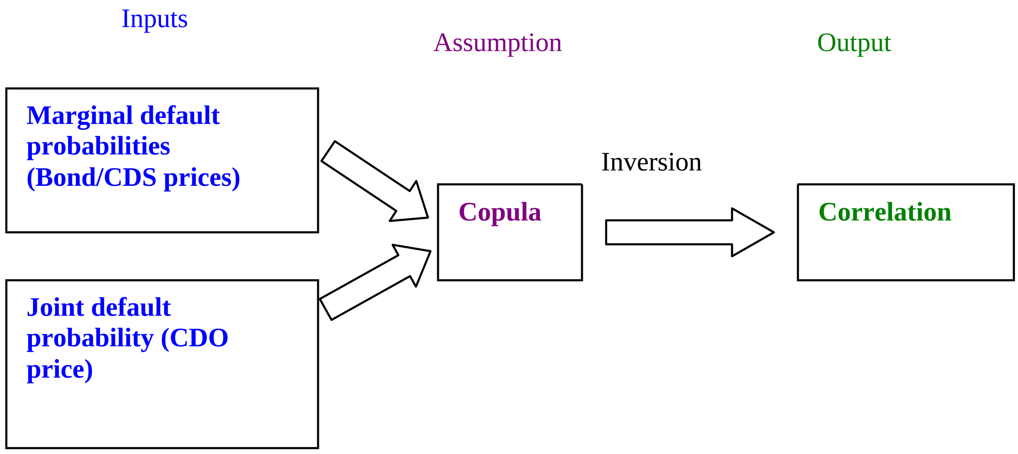


Figure 6.3 Using a one factor model to derive a market implied correlation.

This "inversion" has become market practise, with some talk of implied correlation becoming the credit market equivalent of implied volatility. Indeed, this correlation varies by tranche of a particular CDO, where the model tells us it should be constant, in a way some see as being analogous to the implied volatility smile. Others merely think that this shows that better models are needed.

The Clayton copula is used in a similar way, but is technically more convenient for calculations.

Conclusion

CDO’s are structured assets based on underlying instruments in the credit market. The pricing of a CDO reduces to the estimation of the joint probability of default of the underlying assets. The marginal probabilities of default are observable in the underlying market, making copulas the natural choice for a pricing methodology.

Market practise is to use the Gaussian copula or a factor variant thereof for pricing. This seems to be for modelling convenience rather than for any solid theoretic reason.

Correlations can be estimated, and combined with the marginal default probabilities to produce a joint default probability distribution. This can be sampled from to produce a price for a CDO tranche by Monte Carlo techniques.

Alternatively, the underlying portfolio can be assumed to be homogenous and follow a one factor model. The observed price of a CDO can then be numerically inverted to find the “implied” correlation factor of the underlying risks and the common factor.

We note that this field is still rapidly expanding, and market standards are shifting. Notably, the Gaussian copula which is used in the techniques described above can be shown to significantly underestimate the frequency of multiple extreme defaults, and hence underestimate the risk associated with these products.

Interestingly, in our research for this section, we noticed that much of credit risk valuation involves the analysis of survival probabilities – an area which should be a speciality for actuaries. It is possible that the profession has something to contribute to this field.

About the authors

Martyn Dorey (FIA)

Martyn is a senior consultant at PSolve Asset Solutions. He has10 years experience of research and development within the institutional investment industry. He provides consulting, and licenses techniques on econometric modelling, portfolio construction and asset allocation to investment institutions.

He has a BSc Hons degree in Applied Statistics from the University of Greenwich, and qualified as an actuary in 2004. Prior to joining PSolve in 2003, Martyn worked at a number of leading international investment consultancies, including Head of Quantitative research at Aon Investment Consulting and Quantitative Analyst at Bacon & Woodrow.
martyn.dorey@psolve.com

PSolve Asset Solutions, 126 Jermyn Street, London SW1Y 4UJ..

Phillipe Joubert (FFA)

Phil Joubert is a manager at Deloitte Actuarial and Insurance Solutions. He has worked extensively with life insurers on risk based capital, market consistent valuation, and other projects.

He holds a BSc (Hons) in Pure Mathematics from the University of the Witwatersrand, South Africa, and is currently on sabbatical to pursue an MSc in Mathematics and Finance at Imperial College. He qualified as an FFA in 2003.

pj104@imperial.ac.uk

Deloitte A&IS, Stonecutter court, 1 Stonecutter st EC4A 4TR

Coomaren P. Vencatasawmy

Coomaren is a Risk Financing Analyst at Jardine Lloyd Thompson. He has worked in research and development and as a statistical consultant to insurance companies. He provides structured finance solutions for insurance companies.

He has a PhD in Statistics from the University of Sheffield and an MBA from Edinburgh Business school.

Coomaren_Vencatasawmy@JLTGROUP.COM

Bibliography

[1] Chatterjee, S.K. 2003. *Statistical Thought: A Perspective and History*. Oxford University Press. Oxford.

[2] Kronmal. R.A. 1993. Spurious Correlation and the Fallacy of the Ratio Standard Revisited. *Journal of the Royal Statistical Society A*, **156**, 379-392.

[3] Halley.E. 1686. An Historical Account of the Trade Winds, and Monsoons, Observable in the Seas between and near the Tropicks; with an Attempt to Assign the Physical cause of Said Winds. *Philosophical Transactions*, **183**, 153-168.

[4] Student.1907. On the Error of Counting with a Haemacytometer. *Biometrika*, **5**, 351-360

[5] Fisher. R.A.1935. *The Design of Experiments*. Oliver and Boyd, Edinburgh

[6] Wang S.S. 1997. Aggregation of Correlated Risk Portfolios: Models & Algorithms. Working Paper. CAS

[7] Rebonato, R. 1999. Volatility and Correlation in the Pricing of Equity, FX and Interest Rate Options. John Wiley and Sons. Chichester

[8] Neifeld. M.R. 1927. A Study of Spurious Correlation. *Journal of the American Statistical Association*. **22**, 331-338.

[9] Aldrich, J. 1995. Correlations Genuine and Spurious in Pearson and Yale. *Statistical Science*. **10**, 364-376.

[10] Stigler, S.M. 1989. Francis Galton,s Account of the Invention of Correlation. *Statistical Science*, **4**, 73-79.

[11] Laloux, L., Cizeau, P., Bouchaud, J.P., and Potters M. 1999. Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, **83**, 1467-1470.

[12] Young G.A. 1988. A Note on Bootstrapping the Correlation Coefficient. *Biometrika*, **75**, 370-373

[13] Babour, A.D. and Costi, M. 1993. Correlation Tests for Non-Linear Alternatives. *Journal of the Royal Statistical Society B*, **55**, 541-548..

[14] Moran, P.A.P. 1948. Rank Correlation and Product Moment Correlation. *Biometrika*, **35**, 203-206

[15] Gideon, R.A. and Hollister, R.A. 1987. A Rank Correlation Coefficient Resistant to Outliers. *Journal of the American Statistical Association*. **82**, 656-666..

[16] Abdullah, M.B. 1990. On a Robust Correlation Coefficient. *The Statistician*, **39**, 455-460.

[17] Rebonato, R. and Jackel, P. 1999. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. Preprint

[18] Al-Subaibi, A.A. 2000. Simulating Correlated Multivariate Pseudorandom Numbers. Preprint.

[19] Sklar A 1959 Fonctions de repartition a n dimensions et leurs marges. Pul Inst Statist Univ Paris

[20] Clemen R.T. and Reilly, T. 1999. Correlations and Copulas for Decision and Risk Analysis. *Management Science*, **45**, 208-224.

[21] Venter, G.G. Tails of Copula. *Preprint. CAS*

[22] Nelsen R. R. 1999. **An Introduction to Copulas**. Springer, New-York.

[23] The t Copula and Related Copulas, Stefano Demarta & Alexander J. McNeil, Department of Mathematics May 2004 http://www.math.ethz.ch/~mcneil/ftp/tCopula.pdf

[24] MATLAB copula tutorial material. For details contact Steve Wilcockson, Financial Products Account Manager, The MathWorks Ltd, Matrix House,Cambridge Business Park, Cambridge, CB4 0HH Tel: +44 (0)1223 423200 x246 Fax: +44 (0)1223 423289 www.mathworks.co.uk

[25] Nystrom, K., Skoglund, J. (2002), "Univariate Extreme Value Theory, GARCH and Measures of Risk", Preprint, Swedbank.

[26] Venter, G.G. Quantifying Correlated Reinsurance Exposures with Copulas. *Preprint, CAS*

[27] McNeil, A. and Frey, R. (2000), "Estimation of Tail Related Risk Measure for Heteroscedastic Financial Time Series: An Extreme Value Approach". Journal of empirical finance 7, 271-300.

[28] Burtschell, x., Gregory, J., Laurent, J.-P., *A comparative analysis of CDO pricing models*, 2005

[29] Li, David X., *On Default Correlation: A Copula Function Approach*, RiskMetrics, 2000

[30] Friend, A., Rogge, E., *Correlation at First Sight,* Economic Notes: Review of Banking, Finance and Monetary Economics, 2004