Correlations and Dependencies in Economic Capital Models

Authors:
- Richard Shaw MSc(LBS) FIA, HCR Ltd, [richard.shaw@hcr-uk.com]
- Grigory Spivak MSc FIA, Towers Perrin, [grigory.spivak@towersperrin.com]

ABSTRACT

This paper covers a number of different topics associated with dependency modelling within economic capital models. The scope is relatively wide in that we not only address such technical topics as spurious relationships and the different methods of risk aggregation including copulas, but also more general subjects such as how does one communicate the results of this modelling to the Board of directors of an insurance company. This is a difficult subject, not just because of the underlying mathematical methods employed, but more so from the perspective of setting robust and defensible model parameters.

We have endeavoured throughout this paper to include as many numerical examples as possible to help in the understanding of the key points, including our discussion of model parameterisation and the communication to an insurance executive the impact of dependency on economic capital modelling results.

The economic capital model can be seen as a combination of the two key components, the marginal risk distribution of each risk and the aggregation methodology which combines these into a single distribution or capital number. This paper is concerned with the aggregation part, the methods and assumptions employed and the issues arising, and not the topic of the marginal risk distributions which is equally important and complex.

KEYWORDS

Economic capital, dependency, correlation, linear correlation, Spearman correlation, Kendall Tau correlation, copula, Gaussian copula, T copula, Archimedean copula, variance-covariance matrix, structural model, spurious relationship, time series, tail dependency, positive semi-definite matrix, right tail concentration function, coefficient of tail dependence

Date: 21st September 2009

Correlations and Dependencies in Economic Capital Models

CONTENT

Introduction

This paper was sponsored for the UK Actuarial Profession's Financial Insurance Risk Management ("FIRM") conference of June 2009. However, it may be of interest to other actuarial practice areas such as Non-Life insurance and Life assurance given their day to day activities involving economic capital modelling, pricing, capital allocation and similar.

This paper is relatively wide in its scope. We originally started off with a brief that had a large technical bias nudging towards the more complex areas of dependency modelling such as copulas, but as we progressed in our writing we found ourselves addressing more fundamental questions such as:

- What do we mean by correlation
- Are correlations stable over time and how do they vary
- Are diversification benefits realistic
- How often do we confuse spurious relationships for dependency
- What do people mean when they talk about 'tail correlation'
- How does one communicate to the Board the impact of dependency modelling on economic capital results

Diversification modelling is an important topic. Diversification benefits can amount to around 40% of an insurance company's undiversified total economic capital and for a company with a $1 billion undiversified capital, $400m is a lot of credit. It is therefore of great importance that this number is realistic and that any modelling underpinning the result is analytically robust and well documented.

The words dependency and correlation have recently suffered a rather negative press in the wake of the current financial crisis within the banking industry. Typical comments in the press were along the lines of "it was the fault of the Gaussian copula – it doesn't capture tail dependency" or "the correlations were underestimated" or even "Anything that relies on correlation is charlatanism" in respect of structured credit securities and similarly complex financial products. Often a result of mathematical models undone by their weakest link: their assumptions or their statistical properties. In the case of the Gaussian copula such a weak link was the lack of tail dependency in Monte Carlo simulation studies.

There is a clear need for a greater understanding of dependency and correlation and their limitations if we are to avoid a repeat of such an experience for the banking industry and the wider financial community.

Even at the more basic level one's perspective on dependency can dramatically change according to how information is presented. A scatter plot between two risks will be interpreted differently to two risks represented as historical time series data that provide information on any path dependency, which is effectively lost in the scatter plot.

This paper is a practical one in that it not only illustrates the many concepts and ideas with numerical examples but also highlights many practical considerations in measuring, implementing and communicating the impact of correlations and dependency structures on economic capital results.

Executive Summary

Before going into the detail of each of the six main sections it is useful to provide an overview of the topics discussed herein.

1. Why Diversification is Important

This paper is about modelling dependencies in economic capital models. We start with defining economic capital along with the more recent regulatory developments occurring under Basel II, Solvency II and the UK's current ICA regime. Each of these has greatly influenced the scope of the work currently undertaken by many insurance companies and groups worldwide in this area. Lastly, a brief introduction is given to the idea of dependency.

2. Different Measures of Dependency

Section 2 discusses various models of dependency. Correlation and dependency are often used interchangeably and yet they mean quite distinctly different things. The concept of linear dependency is discussed together with the various correlation measures of linear correlation, Spearman correlation coefficient and Kendall Tau correlation. Correlation as a sole measure of dependency has deficiencies and therefore this section goes on to describe the idea of tail dependency and the evolution of mathematical modelling techniques such as copulas that address many of these weaknesses.

Finally, copulas are discussed in detail starting with their mathematical properties and derivation before going on to describe some of the more popular copulas such as the Gaussian copula, t copula and the Archimedean copula family of the Gumbel copula and Clayton copula.

3. Risk Aggregation

Risk aggregation is at the core of insurance company's economic capital modelling efforts. Each of the main aggregation methods is discussed together with their advantages and disadvantages from the perspective of such considerations as model accuracy, methodology consistency and ease of communication. A more sophisticated modelling approach such as Monte Carlo simulation involving copulas is more flexible than the use of a variance-covariance matrix calculation, but comes at the expense of complexities caused by copula selection, parameterisation and an increase in communication issues.

Finally the most sophisticated, intuitively appealing and potentially the most accurate of the options, namely structural modelling (dependency modelling with common risk drivers) raises other issues such as transparency, parameterisation, and the possible inducement of a false sense of accuracy. Our approach has been to be neutral on each method and present an objective description of the strengths and weaknesses.

4. Model Parameterisation

This is probably the hardest topic in economic capital modelling. We begin with the idea of spurious relationships. The key message is that it is very important that general reasoning and economic rationale are used in conjunction with any numerical assessment of data. Thereafter parameterisation is looked at from the perspective of the different risk aggregation methodologies. We have constructed some simple examples to estimate correlation coefficients from financial time series data that illustrate the sensitivity of the results to the time periods used to conduct such studies and the inherent difficulties of determining these correlations.

Some other topics in this section include discussion of using higher than average correlation coefficients in a variance-covariance matrix calculation as a substitute for tail dependence, how to estimate the missing terms in a variance-covariance matrix, positive semi-definite matrices and methods to fit copulas to data.

5. Impact of Dependency Modelling on Economic Capital

We have constructed a hypothetical insurance company, ABC Insurance Company, to illustrate a number of points related to the impact of dependency modelling on economic capital results. We show how the economic capital varies with the use of different copulas and parameters and how the results compare with the variance-covariance approach.

Finally, we provide some numerical examples that complement our earlier discussion of the issues arising from the common use of higher than average correlations in a variance-covariance matrix approach.

6. Communication of Economic Capital Modelling Dependency Impacts

This section covers the issues related to communication of the results, including questions such as how one explains to senior management what a copula is, what it does and how it impacts the company's overall economic capital.

We present a wide range of possible methods that could be potentially used, discussing their advantages and disadvantages. The methods outlined may be of use in the determination of appropriate copulas and their parameters if similar calculations are made from empirical data.

1. Why Diversification is Important

1.1 Economic Capital

A financial institution, be it an insurance company or a bank, faces a multitude of risks that could cause a financial loss. Economic capital is the amount of risk capital that is needed, assessed on a realistic basis, to cover the risks being run such as insurance risk, market risk, credit risk and operational risk.

There are three main components of an economic capital calculation, (i) risk measure, (ii) probability threshold and (iii) time horizon. A company may do economic capital calculations according to an external criteria laid down by the regulators for regulatory capital purposes or other criteria e.g. to satisfy specific standards for maintaining an external rating level such as 'AA' as prescribed by a rating agency.

Currently the most popular risk measure that is used in banking and insurance is Value at Risk ("VaR"). For example, under the UK's Individual Capital Assessment ("ICA") regime an insurance company needs to hold enough capital such that there is a probability of 99.5% of survival over a one-year time horizon, or in other words the probability of insolvency over 12-months is no more than 0.5%.

However, not all risks the company is facing will suffer adverse losses at the same time. Some areas of business may experience adverse financial losses whilst others average losses, or profits. Another way of viewing this is that if one was looking at the overall economic capital for an insurance company at the 99.5% level then this would be less than the sum of the 99.5% individual capital amounts for each risk.

The extent to which the aggregate 99.5% capital differs from a straight sum of the 99.5% individual capital amounts is a measure of the level of diversification between risks. The lower the degree of dependency between risks the greater the diversification benefit.

In other words, the effect of diversification can be expressed as $1 - EC_T / \Sigma\ EC_i$, where EC_T is an aggregate economic capital total for an insurance company and EC_i is an economic capital on a stand alone basis for each risk i.

The extent of the level of diversification between risks varies from company to company but levels in the range of 25% - 50% are common. The recent CRO Forum QIS 4 benchmarking study of 2008 suggested that diversification reduces economic capital by around 40% on average.

1.2 Regulatory Developments

The use of economic capital models is far greater today compared with a few years ago. One of the key drivers to their more regular use within financial institutions has been the evolution of legislation for insurance companies and banks. Two of the more influential pieces of legislation have been the developments arising under Basel II (Banking) and Solvency II (Insurance). [1]

---

[1] It should be noted that the UK's ICA regime, which came into effect 1/1/2005, could be viewed as a forerunner of the impending Solvency II legislation.

Basel II

Basel II is the second of the Basel Accords issued by the Basel Committee on Banking Supervision. Its purpose being to create an international standard that banking regulators can use when creating regulations about how much capital banks need to hold against financial and operational risks.

Basel II uses the three pillars approach, where the first pillar specifies a minimum capital amount, the second pillar is a supervisory review and the third pillar is a market discipline. From an economic capital modelling perspective the three major components of risk that are covered are credit risk, operational risk and market risk. Furthermore within each major risk there are different permissible approaches to the quantification of risk. [2]

Solvency II

Solvency II is a risk-based approach to insurance company supervision based on a three pillar approach similar to the banking industry.

The first pillar contains the quantitative requirements. Within this there are two separate capital requirements, the Solvency Capital Requirement ("SCR") and the Minimum Capital Requirement ("MCR"). The SCR is a risk-based requirement and the key solvency control level. Solvency II sets out two possible methods for the calculation of the SCR, (i) European Standard Formula or (ii) Firms' own Internal Model, i.e. their Economic Capital models. The SCR will cover all the quantifiable risks an insurer or reinsurer faces and will take into account any risk mitigation techniques such as reinsurance.

The second pillar contains qualitative requirements on undertakings such as risk management as well as supervisory activities. The third pillar covers supervisory reporting and disclosure. [3]

1.3 Internal Models within Solvency II

The modelling of dependencies and the calculation of overall diversification benefits goes beyond a pure bottom line insurance company impact. Dependency modelling will be an integral and very important part of a Firm's overall internal model.[4]

An internal model needs to be seen, for model approval, to be instrumental in the decision making process of a company. Such decision making including (i) the allocation of capital by business unit, line of business or risk (ii) risk-based profitability, (iii) risk-based performance targets, (iv) planning etc all aside from the regulatory capital calculations. Each of these activities being sensitive to the underlying dependency structure embedded in such internal models which might make all the difference to a company in deciding which is a good or bad decision in the long run.

[2] For example, under Credit Risk the options are (i) Standardised approach, (ii) Foundation IRB and (iii) Advanced IRB, where IRB stands for "Internal Rating-Based Approach".

[3] Solvency II will also streamline the way that insurance groups are supervised and recognise the economic reality of how groups operate. Groups will be able to use group-wide models and take advantage of group diversification benefits.

[4] The internal model under Solvency II is a lot more than just an economic capital model that measures and quantifies risk.

1.4 Dependency Structures

In simplistic terms risk modelling consist of two main components (i) marginal risk distribution for each risk and (ii) dependency structure that links these risk distributions. There are enough issues alone in estimating the form of the distribution and parameters for the marginal risk distributions before we even consider how they might be linked. The realistic measurement and modelling of dependencies is one of the most difficult aspects of economic capital modelling facing the insurance and banking industries today.

An unrealistic model of a dependency structure could result in an unrealistic optimistic view of an enterprise, despite the fact that the individual capital components themselves may be quite reasonable.

1.4.1 Dependencies in the real world

Dependency may arise due to the impact of macroeconomic conditions on many risks. For example, inflation rates, interest rates, exchange rates and equity values are not only interrelated but they also influence both sides of the balance sheet.

The impact of these risk factors on asset values is obvious e.g. interest rates on bond values or inflation on equity values, but there is also a direct link to the liabilities. The level of inflation rates will influence the loss payments for underwriting losses and reserve development whilst interest rates will directly impact discounted cash value calculations or act as a risk factor for variation in the underwriting cycle.

Other dependencies are not related to macroeconomic causes but may be a function of the common exposures of an insurance company by line of business. For example, when a major event like a hurricane occurs, seemingly unrelated insurance lines of business covering property, casualty and life could all be affected.

1.4.2 Dependency as a mathematical representation

In a perfect world every single risk would be connected via a complex array of equations such that Risk A → Risk B_1, Risk B_2… → Risk C_1, Risk C_2 → … etc. Such a structure is not attainable, and even if it were an accurate parameterisation would not be feasible.

A statistical dependency between two risks is most often described by a single number, the correlation coefficient. But for many situations this one statistic is not sufficient to capture the range of possible relationships between risks and one needs information on the nature of the dependency structure. The term dependency structure, rather than a simple correlation, is typically used where correlations between risks are not uniform across the range of distribution outcomes. For example, the average outcomes for risks A and B may be 50% correlation, while their extreme outcomes are almost perfectly dependent.

In the real world then we have to make do with the tools available which involves measuring the observed risk correlations and then determining the parameters of the model structure that is being used to reflect such observations.

Many of these themes will be explored in more detail in the following sections.

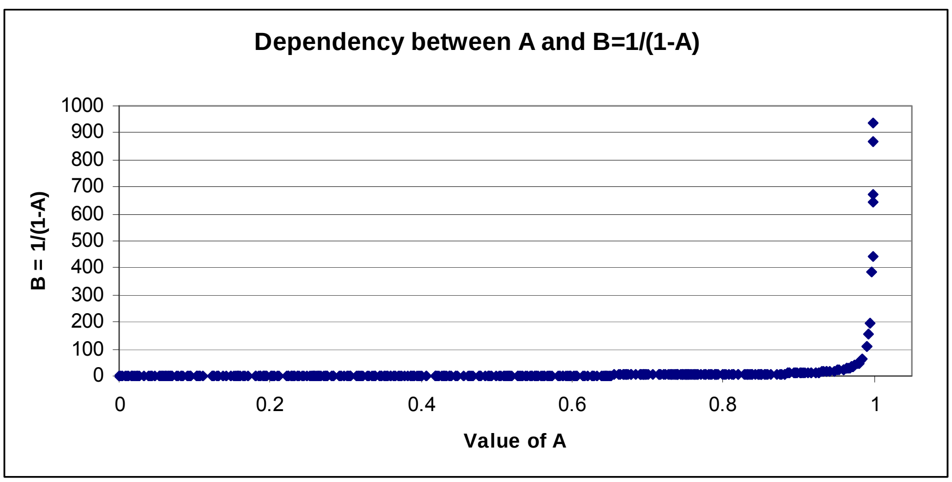
2. Different Measures of Dependency

2.1 What do we mean by dependency

In everyday language a lot of us use the words "correlation" and "dependency" interchangeably. Quite often a correlation coefficient is used in every circumstance when we need to measure the strength of dependency between two random variables. In fact, it is very important to remember that correlation is just a special case of dependency. It quantifies a *linear* relationship between two random variables whilst dependency deals with any kind of relationship.

Dependency between two random variables (e.g. risk factors) means that there is some link between them, i.e. information about one random variable tells you something about the value of the other random variable. One extreme is perfect dependence; if you know the value of one random variable, you know exactly what the value of the other random variable is. The other extreme is independence; the value of one random variable does not enable you to make any predictions about the value of the other random variable.

Dependence between two random variables can be very strong, but such a relationship does not need to follow a linear pattern. Consider the simple example of two random variables A and B. Lets us assume that A is uniformly distributed on [0,1] and also that $B = 1 / (1- A)$. Obviously, there is a strong dependency between them: if we know the value of A, then the value of B is also known. However, the linear correlation coefficient between these two random variables is very low. In fact based on a random sample of 10,000 values, the linear correlation coefficient is only 0.07. The graph below illustrates the reason for this. When A takes a value close to 1, B becomes very large. Therefore, the pattern of the relationship between A and B is very non-linear when A is close to 1.



One of the reasons for the popularity of correlation in finance is that it is used in variance-covariance matrices as part of Modern Portfolio Theory, which is based on the normal (or more correctly, elliptical) distribution. However, in reality, a lot of financial risks that are dealt with in economic capital modelling and other actuarial work are not adequately described by the normal distribution [1], or indeed by an elliptical distribution. Many of these risks exhibit asymmetry and 'fatter' tails than described by the normal distribution, especially in non-life insurance, and so relying solely on correlation as a measure of dependency between risks can be very misleading.

Moreover, by definition, correlation is a constant scalar coefficient. As the market experience over the last year has shown, the dependency structure between random variables can change dramatically with the value of the underlying variables themselves. In stressed market conditions, the implied correlations between various assets classes turned out to be significantly higher than have been observed historically.

Embrechts and others [2] provide the following good summary of the deficiencies of using correlation solely as a measure of dependency:

D1. Correlation is simply a scalar measure of dependency. It cannot tell us everything we would like to know about the dependency structure of risks.

D2. Possible values of correlation depend on the marginal distribution of the risks. All values between -1 and 1 are not necessarily attainable. This means, a model might be impossible to calibrate to certain correlation values.

D3. Perfectly positively dependent risks do not necessarily have a correlation of 1. Perfectly negatively dependent risks do not necessarily have a correlation of -1.

D4. A correlation of zero does not imply independence between risks.

D5. Correlation is not invariant under monotonic transformations. For example, $log(X)$ and $log(Y)$ generally do not have the same correlation as X and Y.

D6. Correlation is only defined when the variances of the risks are finite. It is not an appropriate dependency measure for very heavy-tailed risks where variances appear infinite.

The following sections 2.2 to 2.4 describe the different types of correlation.

2.2 Linear Correlation Coefficient

We begin with by considering a pair of random variables X, Y with finite variances.

2.2.1 Definition 1

The linear correlation coefficient between X and Y is:

$$\rho_L(X,Y) = \frac{Cov[X,Y]}{\sqrt{Var(X)Var(Y)}}, \text{ where } Cov[X,Y] = E[XY] - E[X]E[Y]$$

If we have a sample of *n* observations x_i and y_i where $i = 1, 2, ..., n$, then the sample correlation coefficient, also know as Pearson product-moment correlation coefficient can be used to estimate the correlation between X and Y:

$$r_{xy} = \frac{\sum x_i y_i - n\overline{xy}}{(n-1)s_x s_y} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - \left(\sum x_i\right)^2}\sqrt{n\sum y_i^2 - \left(\sum y_i\right)^2}}$$

Linear correlation has the following important properties:

- The correlation is 1 in the case of an increasing linear relationship, −1 in the case of a decreasing linear relationship, and some value in between in all other cases. The closer the coefficient is to either −1 or 1, the stronger the correlation between the variables. $\|\rho_L[X,Y]\| = 1$ if and only if there exist a, b≠0 such that Y = a + bX
- If the variables are independent then the correlation is 0, but the converse is not true because the correlation coefficient detects only linear dependencies between two variables. Suppose the random variable X is uniformly distributed on the interval from −1 to 1, and $Y = X^2$. Then Y is completely determined by X, so that X and Y are dependent, but their correlation is zero; they are uncorrelated.
- However, in the special case when X and Y are jointly normally distributed, zero correlation is equivalent to independence.
- Linear correlation is invariant under a linear transformation :
$\rho_L[a_1 + b_1 X, a_2 + b_2 Y] = sign(b_1 b_2) \times \rho_L[X,Y]$ for all real a_1, a_2 and $b_1, b_2 \neq 0$
- Linear correlation is not invariant under an arbitrary non-linear monotonic transformation T: $\rho_L[T(X), T(Y)] \neq \rho_L[X,Y]$

The last observation is a serious weakness of the linear correlation. By applying a non-linear monotonic transformation T to two variables *X* and *Y* we are only rescaling the marginal risk distributions, not changing the dependency structure between the underlying random variables, and yet the correlation between them does change.

The generalisation of correlation to the n-dimensional case is straightforward.

2.2.2 Definition 2

Consider vectors of random variables *X* = $(X_1,...,X_n)^t$ and *Y* = $(Y_1,...,Y_n)^t$.
Then given pairwise covariances *Cov*[*X,Y*] and correlations $\rho(X,Y)$ for an *n*×*n* correlation matrix we define:

$Cov[X,Y]_{ij} = Cov[X_i, Y_j],$

$\rho[X,Y]_{ij} = \rho[X_i, Y_j],$ for $1 \leq i, j \leq n$

Such *nxn* matrices have to be symmetric and Positive Semi-Definite (“PSD”) (See Appendix 1 for a definition of PSD).

Rank correlation is an alternative to the use of linear correlation as a measure of dependency. The two most common types of rank correlation are (i) Spearman coefficient and (ii) Kendall Tau correlation. Both of them are commonly used.

2.3 Spearman coefficient

Definition 3

Spearman Coefficient is: $\rho_S[X,Y] = \rho[F_X(X), F_Y(Y)]$
where: $F_X(X)$ and $F_Y(Y)$ are cumulative density functions of X and Y, i.e. their ranks.

In practice, a simple procedure is normally used to calculate ρ_S. If we are given two vectors $X=(X_1, \ldots, X_n)$ and $Y=(Y_1, \ldots, Y_n)$ that represent observations of the random variables X and Y, then ρ_S between X and Y is simply a linear correlation between the vectors of ranks of X_i and Y_i.

Rank correlation, and this refers to both Spearman Coefficient and Kendall Tau (See the next section), does not have the limitations of conditions D1, D3, D5 and D6.

The following property holds for rank correlation: $\rho_{rank}[T(X), T(Y)] = \rho_{rank}[X,Y]$ for any non-linear monotonic transformation T. Rank correlation assesses how well an arbitrary monotonic function could describe the relationship between two variables without making any assumptions about the underlying distribution frequencies of these variables. So we only need to know the ordering of the sample for each variable, not the actual values themselves. Therefore, rank correlation does not depend on marginal distributions of both variables. For this reason it can be used to calibrate copulas from empirical data.

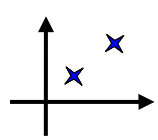
Having said this, the limitations identified in D1 and D4 still hold. It is possible to construct examples of random variables which are highly dependent on each other but have either a low or zero rank correlation coefficient.

2.4 Kendall Tau correlation

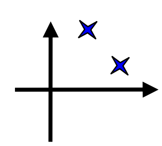
The Kendall Tau correlation measures dependency as the tendency of two variables, X and Y, to move in the same (opposite) direction. Let (X_i, Y_i) and (X_j, Y_j) be a pair of observations of X and Y.

If $(X_j - X_i)$ and $(Y_j - Y_i)$ have the same sign, then we say that the pair is concordant, if they have opposite signs, then we say that the pair is discordant. The following graphs illustrate concordant and discordant pairs in the (x,y)-plane:

Concordant pair:



Discordant pair:



Definition 4

Suppose, we have a sample of n pairs of observations. Let C stand for the number of concordant pairs and D stand for the number of discordant pairs. A simple intuitive way to measure the strength of a relationship is to compute $S=C-D$, a quantity known as Kendall S.

The normalised value of S, namely $\tau = \dfrac{S}{\frac{1}{2}n(n-1)}$

is known as the Kendall Tau correlation coefficient, or Kendall tau.

This measure has a simple intuitive meaning. For example, in a sample with $\tau = 1/3$ then 2 sets of observations $(X_{i,}, Y_i)$ and (X_j, Y_j) are twice as more likely to be concordant than discordant.

The same comments that we made about the correlation properties at the end of section 2.3 also hold here.

In section 2.5 we demonstrate the differences in the values arising from use of these different measures of correlation in the case of a simple example involving 10 joint data observations for two risks A and B.

2.5 Numerical Example

Let us consider the table with 10 joint observations from the two risk factors A and B:

Observations	**Risk A**	**Risk B**	**Risk A Rank**	**Risk BRank**
Observation 1	0.5	0.2	5	1
Observation 2	0.6	0.9	6	8
Observation 3	0.4	0.6	4	5
Observation 4	0.8	0.3	8	2
Observation 5	0.3	0.4	3	3
Observation 6	0.2	0.7	2	6
Observation 7	0.9	0.5	9	4
Observation 8	0.7	0.9	7	8
Observation 9	0.1	1	1	10
Observation 10	100	0.8	10	7
Linear Correlation	0.21			
Spearman Correlation	-0.19			
Kendall's Tau	-0.16			
Concordant pairs	19			
Discordant pairs	26			

The Linear correlation coefficient is equal to 0.21 whereas the Spearman correlation is equal to -0.19. This latter calculation involving the correlation of the ranks between the two risks (which are listed in the last two columns of the table). We note that the Spearman correlation is very different from the linear correlation. This is because the linear correlation is affected by one outlier (the last observation).

Kendall's Tau is equal to -0.16 and is calculated as (19-26)/45. We note that it is close to the Spearman correlation, as they are both rank correlations and therefore are not affected by the large outlier value of Observation 10.

2.6 Tail Dependency

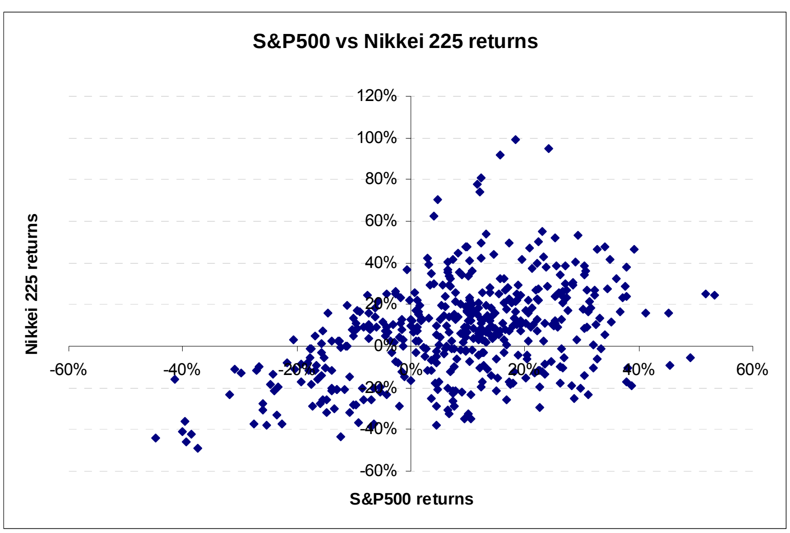
The overriding deficiency of all the correlation measures that we have described so far, namely linear and rank correlation is that they are only scalar measures of dependency and do not allow us to model how dependency changes with economic circumstances (limitation D1 in section 2.1).

For example, they do not allow us to model dependency between risks conditional on the underlying values of the risks themselves, which is often a feature of actual events be it the recent financial crisis or natural disasters such as earthquakes or hurricanes where the observed dependency between risks tends to increase in the event of such scenarios.

A feature of tail dependency is that one extreme event or series of events will trigger risks that are normally assumed to be independent or otherwise have low correlation. Turning to the insurance industry it is recognised that very large events can trigger multiple lines of business. The 9/11 World Trade Centre attack is a prime example of this where large insurance loss amounts were seen in property, business interruption, marine, workers compensation and life insurance lines of business. However, this was not the not the end of the losses as the consequences included falling asset values on insurer's balance sheets.

The recent financial crisis, or 'credit crunch' as it is colloquially known has had noticeable impacts within the financial markets. An otherwise pattern of observed relatively low levels of dependency between various financial asset classes, such as equities, fixed income, credit risk and foreign exchange rates being replaced by severe losses occurring at the same time separately within each. Dependency empirically observed in the market at a time of extremely bad economic conditions (tail dependence) tends to differ structurally from the dependency levels observed in 'normal' market circumstances.

For example, see the following graph showing annual returns of two stock market indices: S&P500 and Nikkei 225 in each month for the last 40 years.



We can see from the previous graph that in benign markets, when the values of the two indices are highly positive, the correlation between them is relatively weak (the points on the RHS are quite widely spread). But in adverse conditions, when the returns are highly negative, the correlation is relatively larger (the points on the LHS form much more of a concentrated pattern). In fact, the linear correlation coefficient calculated for all months when the S&P500 return was negative was 0.9, whereas the linear correlation coefficient calculated for all months when the S&P500 return was positive was 0.4.

The challenge of a good economic capital model is to capture the main features of complex and unpredictable relationships between multiple real world risks with relatively simple mathematical structures. Any differences between the model and reality being accentuated in times of stress.

Given the complex realities of the real world there are in effect three modelling choices:

- Use of correlation matrices together with the variance-covariance approach to capital aggregation.
- Use of copulas (see 2.7)
- Structural models where common risk drivers, e.g. inflation, are simulated with the joint distribution structure of risks being determined by the underlying mathematical relationships that exist between the risks.

These different methods are discussed in more detail in section 3.

2.7 What is a Copula

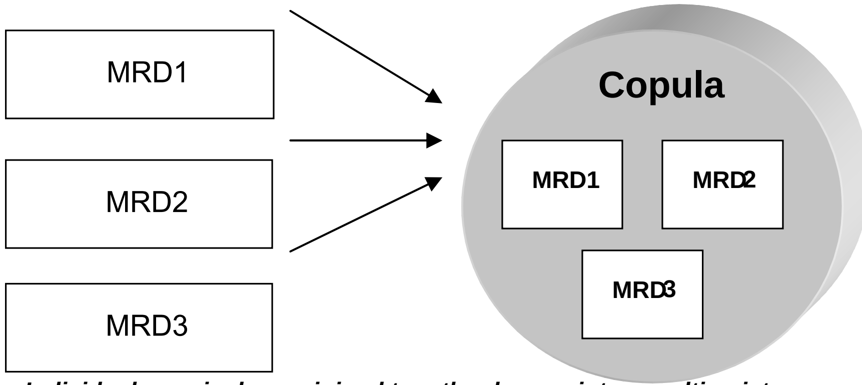
Let us start by considering two random variables X and Y. A range of outcomes for each on a stand alone basis can be represented by a marginal risk distribution ("MRD"), given by its two-dimensional Probability Density Function ("PDF") or Cumulative Density Function ("CDF").

In addition we might happen to know the distribution law which describes the joint distribution of any pair of values (X,Y), i.e. the three-dimensional surface. Visually one can think of loss amounts on the x-axis and y-axis for X and Y respectively with the z-axis representing the value of either the joint PDF or CDF.

If the joint distribution is known, it gives us the best possible information about the behaviour of both variables in aggregate. However, in practice, if we have a large number of risk variables, such as equity, fixed interest, property, non-life underwriting risk, non-life reserving risk etc. it is very difficult to specify a multivariate joint distribution between all risks.

One way of dealing with this difficulty is to split the problem into two parts:

- The first part which describes the individual behaviour of each risk in isolation, i.e. the stand alone marginal risk distribution and
- The second part (which in itself is a distribution function) the dependency structure between the risk variables. This second part is where the copula comes in.



Individual marginal risk distributions ***…joined together by a copula….*** ***…into a multivariate joint distribution.***

Advantages of Copulas:

- Copulas use is consistent with a typical actuarial and financial risk modelling process whereby marginal risk distributions for each risk are first determined and then one considers separately the aggregation process.
- There are a range of different copulas that can be used, each varying in their mathematical properties.
- Copulas are very flexible in that one can combine a varied number of marginal risk distributions together with a varying number of copula distributions. Various types of copulas can be selected depending on one's views on such characteristics of a dependency structure as skewness, kurtosis and tail dependence.
- Even for a selected copula type, there is a wide range of dependency structures that are possible from the use of different copula parameters.
- If one chooses to have a simple model for dependency (e.g. correlation matrix) combined with asymmetric heavy-tailed distributions, this can be done using a Gaussian copula with non-Gaussian MRDs.
- Copulas can more accurately reflect the dependency structure between risks than correlation coefficients can. They avoid the deficiencies of correlations described in section 2.1, in particular, using a suitable copula allows the modelling of a non-zero tail dependency.
- Copulas allow us to express dependencies in terms of quantities of loss distributions. A multivariate loss function constructed using a copula allows the estimation of losses at any given percentile level.
- Most types of copulas are easily simulated using Monte-Carlo methods.
- Copulas are gaining greater recognition as best practice by the various international actuarial and supervisory organisations, which should help in the internal model approval process.

Disadvantages of Copulas:

- There is usually not enough data to perform a credible calibration of a copula, especially in the tail. By definition the extreme joint loss events from various risks that one is trying to reflect in the modelling process are sparse in historical data.
- Any economic capital model becomes more of a 'Black Box'. There is often a lack of transparency in the modelling process. The model is harder to understand and check by a non-mathematician.

- Communication both internally and externally becomes more of an issue when dealing with non-technical people. This should not be underestimated given the advent of Solvency II and the Pillar III disclosure requirements.
- Copulas are essentially static models and a more realistic way of modelling dependency through time would be through use of stochastic process or time series models.

2.8 Copula Mathematics

2.8.1 Definition 5

An n-dimensional copula is a multivariate joint distribution on $[0, 1]^n$ such that each marginal distribution is uniform on [0, 1], i.e. copula C is a distribution function P($U_1 \leq u_1$, …, $U_d \leq u_d$) of a random vector (U_1, …, U_d) such that for every k P($U_k \leq u$)=u for each $u \in [0,1]$

More specifically,
$C : [0,1]^n \rightarrow [0,1]$ is a copula if:

- $C(\overline{u}) = 0$ when $\overline{u}$ has at least one 0 component
- $C(\overline{u}) = u_i$ when $\overline{u} = (1,...,1,u_i,1,...1)$
- $C(\overline{u})$ is n-increasing, for example for n=2:
  For any (a_1, a_2) and (b_1, b_2) such that $a_k \leq b_k$:
  $F(b_1,b_2) - F(a_1,b_2) - F(b_1,a_2) + F(a_1,a_2) \geq 0$

Sklar's Theorem is fundamental to the use and application of copulas.

2.8.2 Sklar's Theorem

If $F(x_1,...,x_n)$ is a joint distribution function with marginal risk distributions $F_1(x_1),...,F_n(x_n)$ then there exists a copula C such that for every $\overline{x} \in R^n$ $F(x_1,...,x_n) = C(F_1(x_1),...,F_n(x_n))$. Moreover, if $F_1(x_1),...,F_n(x_n)$ are continuous, then C is unique.

This theorem means provides the mathematical framework that allows the copula to aggregate the individual marginal risk distributions to derive a joint distribution.

Furthermore:

If C is a copula for $(X_1,...,X_n)$ then for every set of strictly increasing transformations $T_1,...,T_n$ C is also a copula for $(T(X_1),...,(X_n))$.

This means that a copula of a random vector with continuous marginal distribution functions is invariant under strictly increasing transformations of the components of the random vector. We note that this property is similar to the rank correlation properties.

2.9 Different types of Copula and Copula selection

An important distinguishing characteristic of different copulas is their behaviour in the tail. Therefore, we first define more rigorously what we mean by tail dependence.

2.9.1 Definition 6

Let (X,Y) be a 2-dimensional random vector with marginal distribution functions F_X and F_Y.

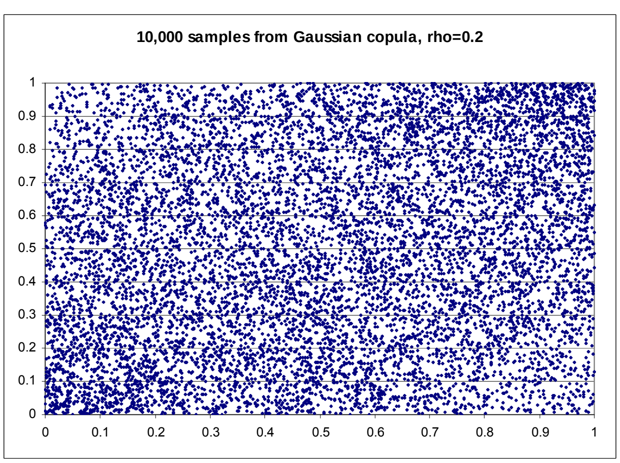
The Coefficient of Upper Tail Dependence of (X,Y) is defined as follows:
$\lambda_U(X,Y) = \lim_{u \uparrow 1} P(Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u))$, provided that the limit $\lambda_U \in [0,1]$ exists.
The Coefficient of Lower Tail Dependence of (X,Y) is defined as follows:
$\lambda_L(X,Y) = \lim_{u \downarrow 0} P(Y \le F_Y^{-1}(u) \mid X \le F_X^{-1}(u))$, provided that the limit $\lambda_L \in [0,1]$ exists.
For example, if (X,Y) is a 2-dimensional random vector with the copula C then it can be shown that:
$\lambda_U(X,Y) = \lim_{u \uparrow 1}\left(\frac{1-2u+C(u,u)}{1-u}\right)$, provided that the limit exists, and
$\lambda_L(X,Y) = \lim_{u \downarrow 0}\left(\frac{C(u,u)}{u}\right)$, provided that the limit exists.

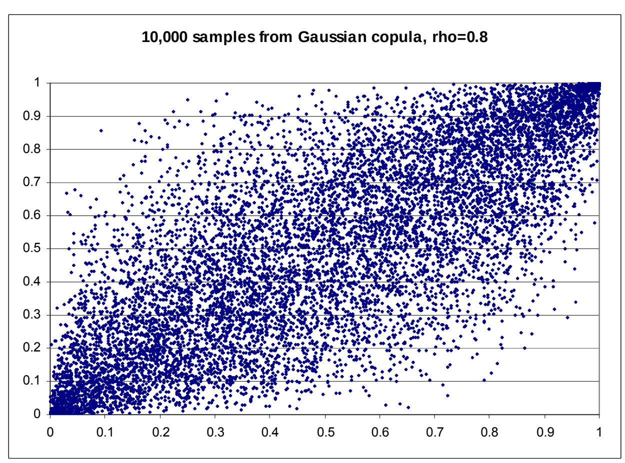
2.9.2 Gaussian copula

The Gaussian copula is the copula of the d-dimensional normal distribution with linear correlation matrix R. It is given by the following formula:
$C_R(u) = \Phi_R^d\left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_d)\right)$, where Φ_R^d denotes the d-dimensional standard normal distribution function with linear correlation matrix R, and Φ^{-1} denotes the inverse of the standard normal distribution function.

The Gaussian copula's tail dependencies are zero, i.e. $\lambda_U = \lambda_L = 0$. This limitation means that the Gaussian copula is not suitable for modelling dependency with heavy tails.

The graphs below show simulations from a 2-dimensional Gaussian copula, (i) the first graph shows results based on a Gaussian copula with a correlation rho = 0.2, and (ii) the second graph one with correlation rho = 0.8. Clearly, the points on the second graph are closer to the straight line.





The correlation matrix R which is a key input to the Gaussian copula needs:

- To be symmetric with unity diagonal elements
- All of its pairwise values to be between -1 and 1
- To be Positive Semi-Definite ("PSD") (see Appendix 1).

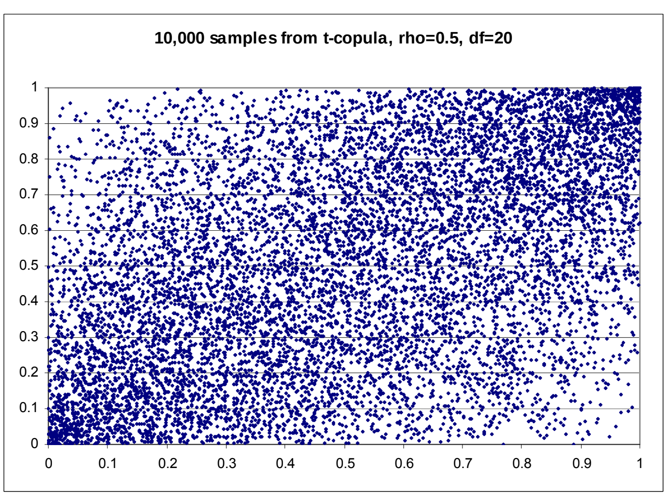
2.9.3 t copula

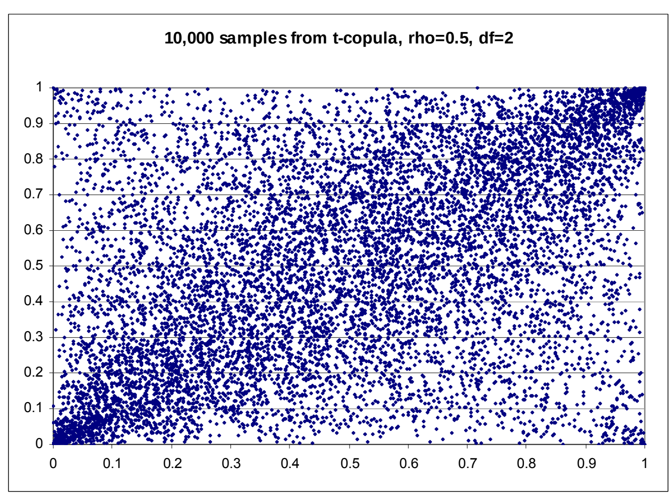
The t copula is constructed from the multivariate t distribution in much the same way as the Gaussian copula is derived from the multivariate normal distribution.

Some of its properties are similar to the Gaussian copula, but there are differences:

- Similarly to the Gaussian copula the t-copula is an elliptical ("bell shaped") function. This makes it mathematically tractable. The t copula is easily extended to the multidimensional case, unlike some other copulas which are limited to two risks only.
- The t copula can be easily simulated just like the Gaussian copula.
- The t copula has non-zero tail dependency coefficients. This is very important, because it allows us to model positive tail dependency between risks.
- Like the Gaussian copula, it requires a correlation matrix R as a parameter input.
- However, in addition it also requires a degrees-of-freedom ("df") parameter. The df parameter determines the strength of the tail dependency, the lower the value of df the greater the tail dependency.
- The t copula is symmetrical and its left and right tail dependencies are equal. This is not the perfect solution given that economic capital modelling is predominantly concerned with only one side of the distribution.
- A limitation of the t copula when modelling more than two risks is that aside from the pairwise correlation coefficient themselves there is only one variable, the df that controls the tail dependency structure. This means that all pairs of risk have the same tail dependency, which is clearly not realistic. This limitation can be overcome by the generalisation of a t-copula commonly known as the IT copula. For details see Ventor et al [3].
- The bivariate t copula with n degrees of freedom and correlation ρ has the following tail dependence coefficients: $\lambda_L(X,Y) = \lambda_U(X,Y) = S_{n+1}\left(\sqrt{(n+1)(1-\rho)/\ (1+\rho)}\right)$ where S_{n+1} is the t-distribution survival function Pr(X>*x*) with n+1 degrees of freedom.

The graphs below show simulation results for a bi-variate t copula with (i) correlation rho=0.5 and df=20 and (ii) correlation rho = 0.5 and df =2. It can be seen from the second graph that the copula with df =2 has a higher tail dependence.





2.9.4 Archimedean copulas

Another family of copulas frequently used in actuarial modelling, in particular in non-life insurance, is the Archimedean family. The most common types of copulas from this family are the Gumbel and Clayton bi-variate copulas.

The specific feature of the Archimedean copulas is the ability to model particularly heavy tail dependence. Unlike the t copulas, the Archimedean copulas are asymmetric. They allow the modelling of dependency structures where tail dependency only exists on one side of the distribution, i.e. either upper or lower tail dependence.

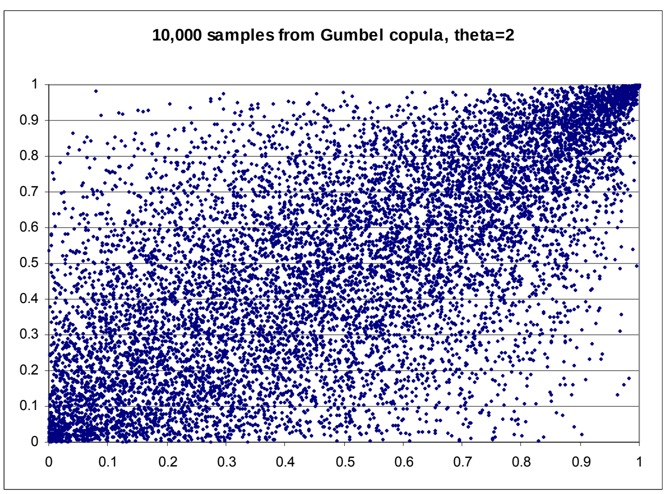
Unlike the Gaussian and t copulas, they are not derived from multivariate distribution functions using Sklar's theorem.

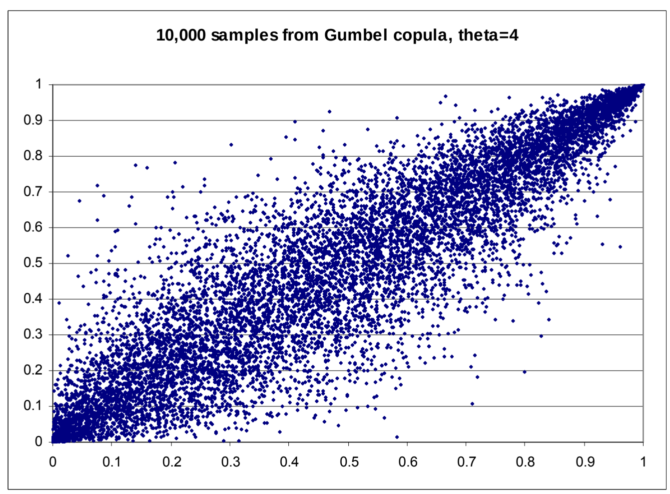
Another distinguishing feature of the Archimedean copulas compared to the Gaussian copula and the t copula is that they do not require a correlation matrix R as an input. Instead, they include a parameter which controls the tail dependency between two risks.

2.9.5 Gumbel copula

The Gumbel family of copulas is given by the formula:
$C_\theta(u_1,u_2) = \exp\left(-\left[(-\ln u_1)^\theta + (-\ln u_2)^\theta\right]^{1/\theta}\right)$, for $\theta \geq 1$. By substituting this expression into the formula for the coefficient of tail dependence we get:
$\lambda_U = 2 - 2^{1/\theta}$ and $\lambda_L = 0$.

The graphs below show simulations from a 2-dimensional Gumbel copula with the parameters (i) theta=2 and (ii) theta=4. It is clear from the second graph that the copula with theta=4 has a higher tail dependence.





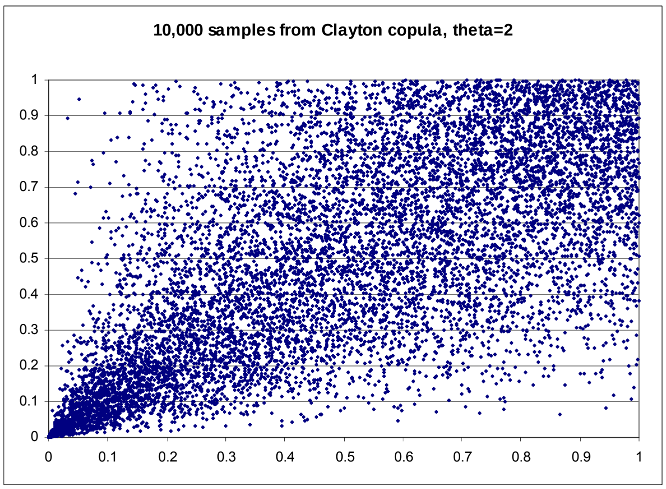
2.9.6 Clayton copula

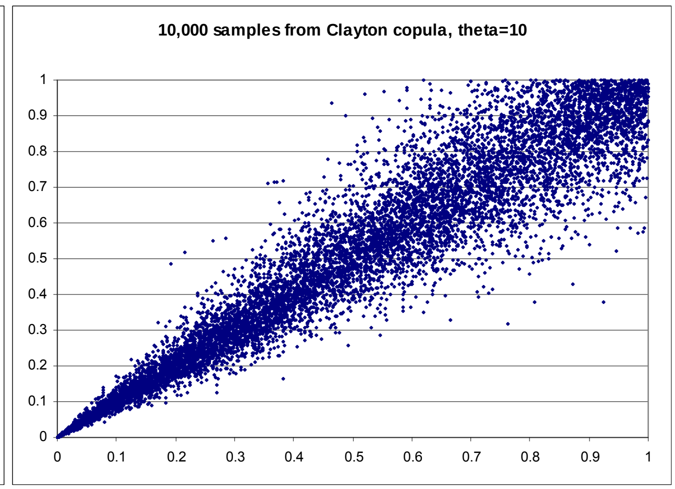
The Clayton family of copulas is given by the formula:

$C_\theta(u_1,u_2) = \left(u_1^{-\theta} + u_1^{-\theta} - 1\right)^{-1/\theta}$, for $\theta > 0$. Then

$\lambda_L = 2^{-1/\theta}$ and $\lambda_U = 0$.

The graphs below show simulations from a 2-dimensional Clayton copula with the parameters (i) theta=2 and (ii) theta=10. It is clear from the second graph that the copula with theta=10 has a higher tail dependence.





2.9.7 Copula Selection

At the moment, the most commonly used copula in economic capital modelling is the Gaussian copula. It is relatively easy to understand, mathematically tractable and can be programmed easily to generate simulated output within an economic capital model. A limitation is that it does not induce tail dependency for extreme losses.

A natural progression on from the Gaussian copula is to consider the use of other copulas such as the t copula, or those from the Archimedean copula family such as the Gumbel, Clayton and Frank copulas.

In economic capital modelling we are of the opinion that the most obvious copula to investigate beyond the use of the Gaussian copula is the t copula (and later its extension the IT Copula). Even with the use of the t copula there still remains the issue of determining both the (i) correlation matrix and (ii) tail dependency parameter so as to produce results that give the requisite tail dependency at high loss percentiles.

The question of the ‘correct’ level of tail dependency is beyond the scope of this paper.

3. Risk Aggregation

3.1 Risk Aggregation Framework

A prima facie reason for the consideration of different dependency modelling structures is risk aggregation in computing overall economic capital levels for insurance companies and banks. Typically economic capital is calculated by first of all assessing the individual risk components and then considering possible techniques to aggregate these components to derive an overall capital number. This approach is a feature of the first four methods that we discuss in sections 3.3 to 3.7.

3.2 Risk Aggregation Methodologies

Insurance companies and banks differ in their approaches to risk aggregation, some techniques being more sophisticated than others. The following is a list of different methodologies in increasing order of complexity:

- Simple Summation (no allowance for diversification benefits)
- Fixed Diversification percentage
- Variance-covariance matrix
- Copulas
- Structural modelling, i.e using common risk drivers [5]. This method is often used in combination with the above methods, such as variance-covariance matrix or copulas.

There are various trade-offs to consider with each method:

- Model accuracy such as the ability to model heavy tailed risks
- Methodology consistency
- Numeric accuracy and availability of data to perform a realistic calibration
- Intuitiveness and ease of communication
- Flexibility
- Resources

Each of these will now be discussed in turn, starting with a description and discussion of each and then listing their advantages and disadvantages.

3.3 Simple Summation

This involves adding together the stand alone marginal risk capital amounts. It ignores potential diversification benefits and produces an upper bound for the economic capital number. Mathematically this is equivalent to assuming a perfect dependency between risks, e.g. 100% correlation.

Advantages:

- No data is required to calibrate the model correlations
- Computational simplicity
- Ease of communication of method and results
- It is deemed to be conservative

[5] The word Structural modelling is being used to describe the process where underlying common risk drivers are identified and their interactions modelled. Their impacts on risks dependent on the underlying mathematical relationships that exist between risks and their common risk drivers.

Disadvantages:

- This method overestimates the amount of required capital, and therefore incurs a cost of holding extra capital
- Does not allow for meaningful interactions between risks

3.4 Fixed Diversification percentage

This method is very similar to the straight summation as described in 3.3 above, however it assumes a fixed percentage deduction from the overall capital figure.

Advantages:

- Data simplicity
- Computational simplicity
- Ease of communication of method and results
- Recognition of diversification effects

Disadvantages:

- A crude method, but allows for some diversification benefit to reduce the capital
- Does not allow for meaningful interactions between risks
- Fixed diversification is not sensitive to changes in underlying risk exposures
- Does not capture non-linearities

3.5 Variance-covariance matrix

This method allows for a fuller pattern of risk interactions with the assumption of differing pairwise correlations between risks. The overall level of diversification between risks is dependent on the levels of these correlations.

Within this approach there are various considerations.

3.5.1 Risk Dimensions – Economic Nature vs Organisational

Many financial institutions, in particular large insurance groups, consist of various subsidiaries, business units (“BU”) or similar organisations. When faced with this situation there are two important dimensions of risk classification:

- Economic nature of the risk – insurance, market, credit, operational risk etc.
- Organisational structure – business lines or legal entities

The economic nature of the risk considers aggregating risks into silos by risk-type across the whole group e.g. equity risk by consideration of the aggregate risk at a group balance sheet level. By contrast the organisational risk grouping would consider organisation silos before the aggregation to a group capital total. This approach deals with inter-risk relationships earlier on in the process and takes advantage of known corporate structures. An organisational classification presents far less difficulty than a classification by risk where definitions of risk may be imprecise across different organisations.

A third approach features aspects of both and operates at a lower level of risk granularity. In this situation the unit of risk that is worked with is of the form "Organisation / Risk" e.g. UK / Equity, France / Fixed Interest etc, the aggregation process thereafter working from this base level.

However, whereas at face value this would seem to be conceptually a more accurate approach there are other issues to consider like the smaller volumes of data at this finer level of granularity and the difficulties of estimating cross-terms in the enlarged correlation matrix. For example what would be the correlation coefficient between "UK / Equity" and "France / Fixed Interest" given the most likely scenario that correlation assumptions would only have been determined between the Equity and Fixed Interest risks within each business units such as the UK or France.

Section 4.7 discusses the issues arising in trying to determine these cross-terms.

3.5.2 Risk Granularity

The finer the level of risk classification (i.e. a more granular subdivision of risk) within a variance-covariance matrix, the lower the intra-risk diversification (i.e. diversification within a risk category) and the greater the inter-risk diversification (i.e. the diversification between risk categories).

A simple example of this would be in considering non-life insurance risk where one could envisage either (i) just one risk category for non-life insurance risk say in a variance-covariance matrix with other non-insurance related risks or (ii) non-life insurance risk split further into broader groupings by line of business of Property, Casualty and Other insurance risks together with the same non-insurance related risks. All other things being equal scenario (i) would result in greater intra-risk diversification but lower inter-risk diversification than scenario (ii).

Differences in approaches will generally lead to differences in the economic capital number given the complexity of re-working all of the various risk dependency relationships.

3.5.3 Subdivisions of the variance-covariance matrix

Sometimes the economic capital calculation will feature a series of "nested" variance-covariance matrices. A topical example of this is the current method used within the standard formula approach to the Solvency II Solvency Capital requirement ("SCR") as detailed in the Solvency II QIS 4 Technical Specification [4].

Table 1 details the correlation matrix that is used to aggregate the individual capital amounts for each of the five main risk categories to derive the Basic SCR. Market Risk capital ("SCR Market") is one of the major risk capital categories within this process.

Table 2 shows the "nested" Market Risk correlation matrix, i.e. the matrix that is used to aggregate the individual capital amounts in respect of different types of market risk e.g. interest rates, equities etc to derive an overall Market Risk capital number for use in Table 1.

Table 1

	SCR Market	SCR Default	SCR Life	SCR Health	SCR Non-Life
SCR Market	100%	25%	25%	25%	25%
SCR Default	25%	100%	25%	25%	50%
SCR Life	25%	25%	100%	25%	0%
SCR Health	25%	25%	25%	100%	25%
SCR Non-Life	25%	50%	0%	25%	100%

Table 2

	Mkt Interest	Mkt Equity	Mkt Property	Mkt Spread	Mkt Concentration	Mkt Currency
Mkt Interest	100%	0%	50%	25%	0%	25%
Mkt Equity	0%	100%	75%	25%	0%	25%
Mkt Property	50%	75%	100%	25%	0%	25%
Mkt Spread	25%	25%	25%	100%	0%	25%
Mkt Concentration	0%	0%	0%	0%	100%	0%
Mkt Currency	25%	25%	25%	25%	0%	100%

It is a common practice for many life insurance companies to aggregate individual stress tests results by using the variance-covariance matrix approach. Some non-life insurance companies [6] use this approach as well. However, more and more companies (at the moment mostly in the non-life area, although with life offices also gradually moving in the same direction) use more sophisticated mathematical models involving copulas and structural modelling.

<u>Advantages:</u>

- More accurate calculation than the previous two methods
- Relatively simple, intuitive and transparent
- Facilitates a consensus of typical correlations for use by companies
- The use of a cascade of correlation matrices permits the easy addition of further risks, from a new business unit, subsidiary or risk category
- Correlation is the only form of dependency that a lot of non-specialists are familiar with. This makes communication easier than some of the more sophisticated methods described in sections 3.6 and 3.7.

<u>Disadvantages:</u>

- Risks where we have empirical evidence of correlations (mainly reliable market data) are very few and so there is a heavy reliance on a subjective 'expert opinion' to determine correlations.
- The variance-covariance matrix approach implies that the underlying risks are normally (or elliptically) distributed
- Underestimates the effects of skewed distributions and does not allow for potential heavier dependency in the tail
- The value of correlations is sensitive to the underlying marginal risk distributions
- A correlation matrix has to satisfy certain conditions (e.g. be PSD). These are often ignored in practice
- All cause-effect structures cannot be properly modelled
- Does not capture non-linearities

[6] The word insurance company has been used in a generic sense to also mean Lloyd's syndicates, reinsurers and similar insurance enterprises.

[8] For the variable X at time t, X_t an AR(n) model is of the form $X_t = a + \Sigma^n b_i X_{t-i} + \sigma.\varepsilon_t$ where the ε_t are i.i.d. unit normal random variables. n = 2 or 3 are common models of the underwriting cycle.

3.6 Copulas

As described in section 2.7 the copula is a function that combines the marginal risk distributions to form a joint risk distribution. The advantages and disadvantages of copulas have been described in some detail in section 2.7. The list below briefly summarises the main points given the overall objectives we had set out in section 3.1.

Advantages:
- This method is more flexible than the use of a variance-covariance matrix
- There are a range of different copulas that can be used, each varying in their mathematical properties e.g. the symmetry, strength of tail dependence, etc.
- Copulas enable the user to build models that reflect reality more e.g. heavy tails
- It is possible to allow for non-linearities and other higher order dependencies
- Copulas are easily simulated using Monte-Carlo methods.

Disadvantages:
- Copula selection is non-trivial. There are many considerations.
- Parameterisation of copulas is very difficult, by definition the joint loss data that is needed is sparse in the tail
- The full marginal risk distribution is needed for each risk, rather than in the case of the variance-covariance matrix approach only the relevant capital number.
- The bottom-up Monte Carlo simulations approach is more demanding computationally, especially if the number of risks is high
- The communication of the method and results can be much harder

3.7 Structural Modelling

This method is intuitively very appealing as it can reflect directly, more than any other methods, possible relationships that might exist between different risks. The method can be describe as the simulation of common risks, for every risk, instead of modelling each risk separately and then aggregating them into a joint loss distribution. The economic capital model implicitly captures any diversification benefits between risks.

Within the structural modelling approach the risk dependencies are likely to be at a more granular level than one would see within the variance-covariance matrix approach. For example, some typical structural models capture dependencies such as:
- Rate movements between lines of business
- Large loss frequency between lines of business
- Inflationary link between loss reserves and loss volatility
- Inflationary link between loss reserves and asset values

3.7.1 Non-Life Underwriting Cycle

For non-life insurers the underwriting cycle is one of the more obvious candidates for some form of structural modelling. The non-life underwriting cycle can be thought of as a recurring pattern of increases and decreases in insurance prices and profits.

The cycle exhibits characteristics of a dynamical system with feedback loops and common economic and social shocks. Each line of business typically has its own cycle and cycles are often linked across lines of business within any one company.

Models often focus on some form of profitability measure based on the loss ratio or combined ratio (loss ratio + expense ratio). For the dependent variable of interest then a number of predictor variables are possible:

- Previous value of the variable over prior time periods
- Other company financial variables such as reserves, investment income and capital
- Regulatory and/or rating variables
- Financial market variables such as interest rates and equity returns
- Econometric variables such as inflation, GDP growth etc.

3.7.2 Static vs. Non-static models

In a simple autoregressive ("AR") [8] model of the underwriting cycle, where the combined ratio is being modelled, the marginal risk distribution of the combined ratio in future years t is conditional on the prior history of combined ratios in years t-1, t-2 etc. As such we can think of the marginal risk distribution as being non-static in that it evolves over time. This is in contrast to the typical copula simulation approach where the marginal risk distribution is fixed.

3.7.3 Simple Example

A simple example of this more purist way of thinking of dependencies is where the value of a risk in any given scenario is made up of the impact from a set of simulated common risk factor drivers plus a simulated residual component for that risk. Furthermore, the residual components may themselves be subject to correlation.

In simplified forms of Collateralised Debt Obligation ("CDO") modelling it was common to assume the latent variable was the asset return of a counterparty. Default was then deemed to occur when the value of the counterparty's asset return in any particular scenario fell below some 'asset' threshold, itself related to the value of its liabilities.

One could represent the asset return for each counterparty as a multi-factor model. In the simplest case we will consider a 2 factor model which consists of a (i) systematic component and (ii) non-systematic part.

Let the systematic component X be the "state of the economy", R^2_i the "counterparty asset return correlation with the market" and ε_i the counterparty-specific (or residual risk) part. It is further assumed that the variables X and ε_i are normally distributed. In this example X is the underlying common risk driver, with the values of R_i and ε_i also determining the degree of dependency between risks

For each counterparty i = 1 and 2 the asset return can be defined as:

$AR_i = [R^2_i]^{0.5} X + [1 - R^2_i]^{0.5} \varepsilon_i$

Furthermore, given AR_1 and AR_2 the asset correlation ρ_A is then:

$\rho_A = \text{Corr}(AR_1, AR_2) = [R^2_1]^{0.5} \times [R^2_2]^{0.5}$

Example: $R^2_1 = 50\%$ and $R^2_2 = 20\%$ then $\rho_A = 15.81\%$

The R^2 represents the proportion of the asset return that can be explained by variation in the state of the economy i.e. it's systematic risk. The non-systematic part consists of both counterparty specific pieces and non-counterparty specific pieces that are common to groups of credit assets but are not deemed to be systematic in nature, e.g. exchange rates. The R^2 can be determined for a company by computing the correlation of the asset value of the company with an index of asset values that represents the universe of companies.

Advantages:
- Theoretically it is a very appealing and intuitive method
- Potentially the most accurate in imitating the way the 'real world' works with a series of external and internal shocks to a company
- Could be used in combination with other methods, e.g. an inflation variable may be used as a common risk driver for expense and claims risks with some further correlation between the expense and claims risks due to other factors
- Possible to capture non-linearities through structural risk relationships

Disadvantages:
- The most demanding in terms of inputs
- It is not feasible to model all common risk factors at the lowest level
- If lots of common risk factors are simulated using a Monte Carlo approach, this puts a very high demand on computing power
- Transparency and results communication becomes an issue – 'Black Box' approach
- Parameterisation issues relating to the structural relationships
- Could lead to an overly complicated model providing a false sense of accuracy

3.8 Aggregation Approaches

Companies often use a combination of the different methods that we have described so far. For example, each insurance company within an insurance group may have models that operate in sufficient detail for its own purposes, but all companies within the group use common economic model output and disaster scenarios that imply dependency when risk is viewed at a group level.

Reinsurance Credit Risk – Modelling granularity

The modelling of reinsurance credit risk, i.e. the loss associated with the failure of reinsurance counterparties is a good example of the different levels of modelling granularity.

One of the key dependencies for non-life insurance companies is between Catastrophe underwriting risk ("Cat UW risk") and Reinsurance credit risk ("RI Credit risk"). If a low frequency, high severity natural catastrophe loss event were to occur, then a non-life company writing property related lines of business may see a large increase in reinsurance credit risk, due to both an increased likelihood of default by its reinsurers and an increased exposure i.e. larger reinsurance recoveries.

Different Methods are possible to reflect this dependency structure between Catastrophe UW Risk and Reinsurance credit risk.

- If Cat UW risk and reinsurance credit risk are separate entries within a variance-covariance matrix then the dependency structure between them could be proxied by a higher than average correlation. The marginal risk capital calculation for reinsurance credit risk being based on the expected level of reinsurance recoveries and other risk factors.
- The next level of complexity could involve Cat UW risk including the RI Credit risk credit risk associated with UW risk. Given that it is common for Cat UW risk modelling to involve the simulation of gross losses and reinsurance recoveries such a method would facilitate more accurately the varying reinsurance loss exposure for calculation of the associated reinsurance credit risk marginal capital. In this scenario a variance-covariance RI Credit risk marginal capital calculation would only be in respect of reinsurance recoveries associated with prior year's business.
- Another layer of complexity would involve allowing stochastic interest rates in the discounting of the reinsurance loss payments in the RI Credit risk calculation given that the reinsurance loss exposure is on a present value basis.
- More complex methods could involve loss given default and reinsurance default rates being a function of insurance losses and / or asset values etc.

The last two methods, whilst being more intuitive, than the earlier methods, do at the same time introduce more uncertainly in terms of both model risk and parameter risk.

3.9 CRO Forum Internal models benchmarking survey

The CRO Forum "Internal models benchmarking study, Summary results" survey as carried out by Oliver Wyman (30/1/09) provides a useful overview of the current economic capital risk aggregation calculation methods adopted by 16 members [9] of the Chief Risk Officers ("CRO") Forum, and two Associate members.

Risk Aggregation (page 23)
In terms of an overall risk aggregation the most popular approaches were (i) variance-covariance matrix [10] (60%) and (ii) simulation (30%) with only 5% of respondents citing the use of copulas. [11]

Non-Life Underwriting Cycle (page 35)
Within non-life business lines currently 44% of the survey respondents capture the underwriting cycle within their economic capital models and within 2-3 years the majority expect to be allowing for the underwriting cycle. Currently only about 25% of the companies are using autoregressive models to do this.

As was discussed in section 3.7, the underwriting cycle is an obvious candidate for some form of structural dependency model.

---

[9] The members of the CRO Forum include some of the world's largest insurance groups such as AIG, Aviva, AXA, ING, Munich Re, Prudential, Swiss Re and ZFS.
[10] This includes widespread use of stochastic models for significant risk types such as non-life underwriting risk and reserve run-off risk, and also stochastic models for market risk.
[11] It does appear unusual that Copulas are mentioned in a different category to simulation given that copula outputs rely on monte carlo simulation techniques.

4. Model Parameterisation

Parameterisation of the variables used to model dependency structures in economic capital models is very often difficult . Many issues arise, not only in the estimation of parameters themselves e.g. correlations for use in a variance-covariance matrix calculation, but how these parameters evolve over time as a result of changes in economic indicators, business cycles or underwriting cycles.

Some of the typical questions arising are:
- Estimation of correlation coefficients for use in a variance-covariance matrix
- Estimation of parameters and / or correlation coefficients for use with a copula
- What sources of data and information are needed for the parameterisation exercise
- What sources of data and information is currently available
- How accurate, reliable and credible are the sources of data and information available

Before addressing each of these points in turn it is useful to consider the topic of spurious relationships and a discussion of whether relationships between variables do exist or are incorrectly perceived to be by chance as a result of the data under study.

4.1 Spurious Relationships

In statistics a spurious relationship [12] is a mathematical relationship in which two occurrences have no causal connection, yet it may be inferred that there is one. "Correlation does not imply causation" is often used to point out that correlation does not imply that one variable causes the other. However, the presence of a non-zero correlation may hint that a relationship does exist.

Edward Tufte [5] puts it succinctly:
"Empirically observed covariation is a necessary but not sufficient condition for causality" or in other words "Correlation is not equal to causation; it is only a requirement for it".

4.1.1 Correlation does not imply causation

1. A occurs in correlation with B
2. Therefore, A causes B

In this type of logical fallacy a conclusion about causality is made after observing only a correlation between two or more factors. When A is observed to be correlated with B it is sometimes taken for granted that A is causing B even when no evidence supports this. This is a logical fallacy as four other possibilities exist:
(a) B may be the cause of A
(b) An unknown factor C may be causing both A and B
(c) The 'relationship' is coincidence or so complex or indirect that it is more effectively called coincidence
(d) B may be the cause of A at the same time as A is the cause of B.

Determining a cause and effect relationship requires further study even when the result is statistically significant.

[12] Sometimes called spurious correlation or spurious regression

Examples of each, drawn from everyday life as analogies, are:

(a) "The more firemen fighting a fire (A), the bigger the fire is going to be (B). Therefore firemen cause fire ". In reality it is (B) the fire severity influencing how many firemen are sent (A).

(b) "Sleeping with one's shoes on (A) is strongly correlated with waking up with a headache (B)". This ignores the fact that there is a more plausible lurking variable excessive alcohol (C) giving rise to the observed correlation.

(c) " With a decrease in the number of pirates, there has been an increase in global warming, therefore global warming is caused by a lack of pirates"

(d) " According to the ideal gas law PV = nRT, given a fixed mass, increased temperature (A) results in increased pressure (B), however an increase in pressure (B) will result in an increase in temperature (A). The two variables are directly proportional to each other and independent.

With regards economic capital modelling simple examples of (a) and (b) are:

(a) Increasing domestic demand and inflation (A) often leads to the Government having to increase short-term interest rates (B) to counter potential over-heating in the economy, evidence of positive correlation. Conversely falling short-term interest rates (B) is likely to lead increased demand, which once spare capacity is utilised, to increasing inflation (A), however in this case, evidence of negative correlation.

(b) The large negative correlation between equity returns (A) and credit spreads (B) during 2008 could be viewed as a consequence of the financial crisis (C)

Some observed correlation relationships are one way A → B. For example, a very severe natural catastrophe could lead to a large decrease in equity markets, but a large fall in equity markets is not likely to result in a natural catastrophe.

4.1.2 Economic Logic

When determining correlations between risks one should consider the questions:

- Is the relationship logical (rather than spurious)
- Is there statistical evidence for the hypothecated relationship

An example of a relationship that satisfies both of these questions is in the simple example of a yield curve. The three year bond yield is closely related to the two year and four year bonds yields, which is intuitive given that yield curve movements are often thought of as a combination of (i) parallel shift and (ii) slope changes. Furthermore. empirical studies support positive correlations between adjacent points of the yield curve

4.1.3 Correlation and Linearity

The Pearson correlation coefficient indicates the strength of a linear relationship between two risks, but it alone is often not sufficient to evaluate the strength of this relationship. Appendix 2 shows scatter plots of Anscombe's quartet, a set of four different pairs of variables created by Francis Anscombe [11]. The *y* variables have the same mean, standard deviation, linear correlation and regression line and yet in all four cases the distribution of the variables is markedly different.

These numerical examples demonstrate that the correlation coefficient, as a statistic summary, cannot replace a more detailed examination of the data patterns that may exist.

4.1.4 Spurious Regression Example

Consider 2 random walk time series X_t and Y_t as follows:
$X_t = X_{t-1} + \varepsilon_t$
$Y_t = Y_{t-1} + \delta_t$

where:
ε_t and δ_t are N(0,1) distributed
$X_0 = Y_0 = 5$.

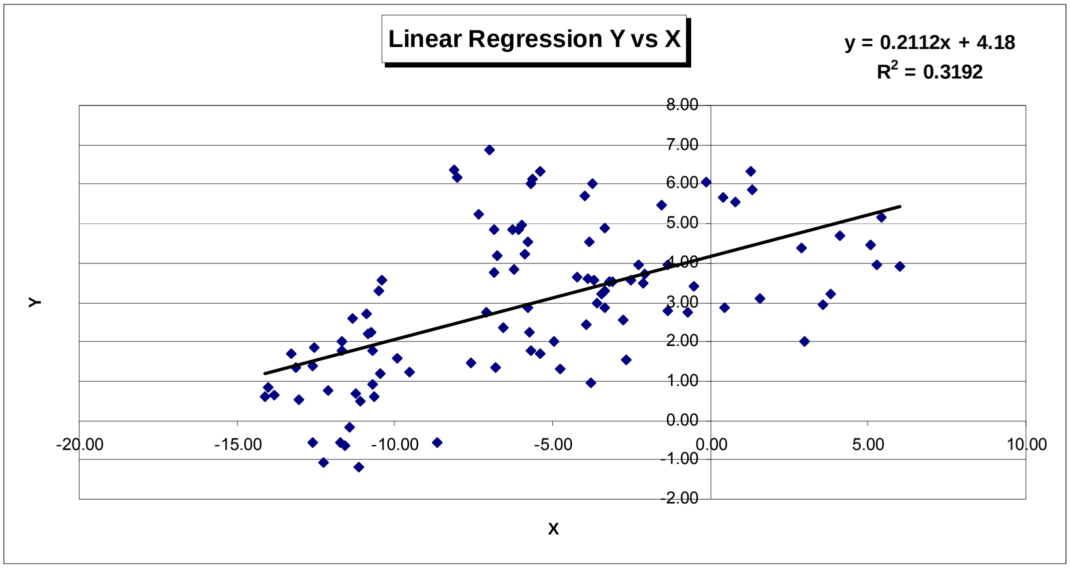
Given this, a random sample of 100 values for each of X and Y for t = 0 to 99 has been generated. Using this output the linear correlation and R^2 have been calculated.

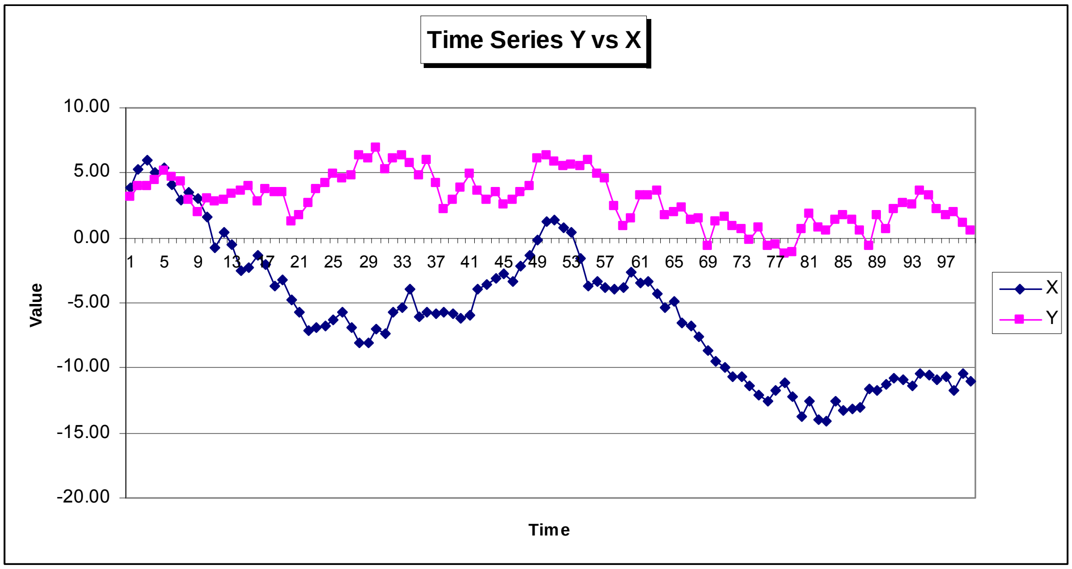
X and Y are not related and yet it is common, in repeated runs, to observe very high correlations far in excess of those expected from sampling error in the N(0,1) values.

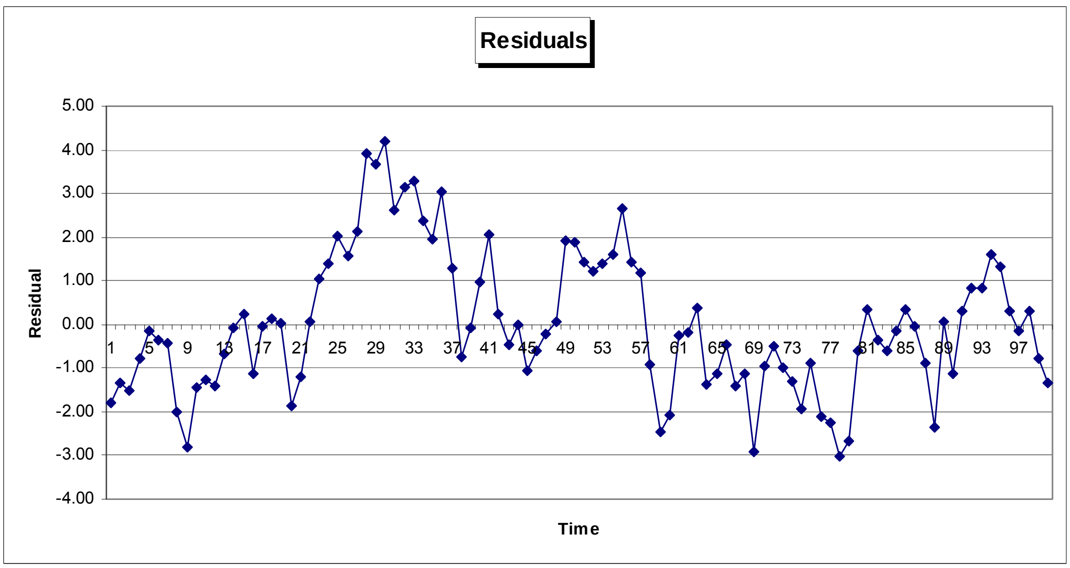
The following graphs show (i) Linear regression between X and Y, (ii) Time Series of X and Y (ii) Residuals from the linear regression.

In this random example, the observed correlation between X and Y is 56.5% and 5.9% between the residuals (any difference from zero due to sampling error). The time series plots for X and Y look plausible for atypical financial variables and yet we conclude they have a large positive correlation when in fact there no relationship between them at all. Moreover, there is significant autocorrelation in the residuals leading one to reject the linear regression model as a measure of the relationship.

In fact this is illustrative of the fact that trending variables, which are often a feature of economic and financial time series data, are likely to lead to a regression with high values of R^2, regardless of whether they are related or not. Differencing variables (changes) eliminates trends and thus avoids spurious regression. So it is important to consider the nature of the variables being used to determine the correlation.







4.2 Variance-covariance matrix correlation parameterisation

The following is a list of possible approaches:

- Empirical estimation using historical time series data
- Use of expert judgement or industry benchmarks
- Ranking method, e.g. using Low, Medium and High rankings

Dealing with each of these in turn:

4.2.1 Use of historical time series data

A starting point in determining appropriate correlation estimates would be an estimate based on the historical time series data of underlying risks. It could be argued that any estimation based on internal data is the most appropriate given it will reflect an insurance company's actual experience and any differences in its business and risk profile.

Considerations arising include:
- Choice of index or data on which the time series is based
- Length of time series data
- Data frequency e.g. weekly / monthly / annual
- Dealing with data gaps, data credibility
- Data weightings, e.g. perhaps giving more weight to recent time periods
- Prospective views

Very often a published index is preferable to actual company data, not only because of the likelihood of a longer data history but to minimise data errors and secular risk. For example it may be more pragmatic for a UK insurance company when estimating Equity risk correlations to use the FTSE 100 Equity Total Return Index rather than its own data which could would contain a varying mix of equities from year to year.

Often data is not complete therefore companies need to apply different techniques or a combination of techniques to overcome such shortcomings. Such techniques being:
- Secondary data – sometimes companies supplement their internal data by secondary data from either public sources or from external data suppliers. For example, a useful source of data for underwriting risk is reinsurers' data.
- Simulating data – data can be enhanced by simulating historical data that is not currently observable. This synthetic data is itself the output from a model and appropriate parameters.

Data quality can be an issue and can generally be grouped under 3 main headings:
- Consistency – is the data consistent and collected in a standard format ?
- Completeness – is the data thorough, e.g. taking into account missing dates ?
- Accuracy – is the data correct? Common issues with data accuracy are processing errors, miscoding, bulk coding and bias.

Each in itself can have an undue influence on the credibility or otherwise of the parameterisation process.

One of the more important aspects of data quality is the treatment of outliers. It is important to check whether any series contains outliers and if so try to understand the reasons for their occurrence. If an insurer thinks that the outlier reflects an anomaly that may repeat itself in the future then it will often retain the observation. If the outlier is for an event that is unlikely to occur again (perhaps because of a change in exposure), or is perceived not to be material for the future, then sometimes companies disregard it from the data or assigning a lower weight (<100%) than other data points.

Any omission of a data outlier, unless an incorrect data entry is not a good idea in that it allows for future *unexpected* events that may have been unforeseen when performing the analysis.

Correlations do vary over time and in a lot of cases quite markedly, so the analysis of historical time series data should not just be a once in a while exercise but an analysis that is performed quite regularly over time.

4.2.2 Expert judgement / Industry benchmarks

Very often company-specific data is not available or is of poor quality. Perhaps an index proxy for a risk is not suitable or the correlations estimated from company own data or index proxies vary too much over time.

In such situations entries in a variance-covariance matrix may be filled on the basis of expert judgement. In such cases the parameters are based on the consensus of risk officers, underwriters, business managers, actuaries and other specialists in an organisation who understand the nature of risks being modelled. This is frequently complemented with an input from external consultants and industry benchmarks.

Furthermore, expert judgement or opinion is a good starting point before any time series analysis has taken place serving very much as a reference point.

This approach introduces an element of subjectivity but may be necessary if the prospective view of risk is different to that captured in the historical data. Expert opinion and judgement becomes more important when looking at extremes of risks where by definition they are unlikely to be very common in data series.

The reliance on expert judgement is likely to vary by risk category. For example, this is likely to be more the case for operational risk than would be for equity risk. Furthermore, the reliance on this approach for risks may be more common for those organisations that are smaller in size and lack the capacity, scope and economies of scale to estimate correlations based on their own experience.

4.2.3 Low, Medium and High rankings

In the absence of any data analysis organisations may fall back on the use of Zero, Low, Medium or High correlation rankings or similar grading based on a subjective assessment of the main risk pairings. The correlation coefficients allocated to these groupings are determined in advance based on prior studies. But what do we actually mean by Medium correlation, which could be regarded as rather arbitrary. It could mean 30%, 40%, 50% or somewhere in between. In addition, the actual values within each category are often dependent on the nature of their use.

The following illustrates the range of realistic correlations that may appear in each of the correlation groupings, but others could equally be valid.

Correlation Rankings

Correlation	Negative	Positive
Zero	0	0
Small	-0.3 to -0.1	0.1 to 0.3
Medium	-0.5 to -0.3	0.3 to 0.5
High	-0.5 to -1.0	0.5 to 1.0

Categorising dependencies as either Low, medium or High is sometimes underplayed by commentators as an approach that should be limited to situations when there is no more accurate approach. This worryingly paints a misleading picture that correlation assessment with voluminous time series data sets is in anywhere near reliable or accurate.

4.3 Estimation of correlation coefficients from historical time series data

Estimating “credible” correlation coefficients from historical data series is far from easy. The difficulty lies not in the calculation of different correlation coefficients, using the relevant mathematical formulae and data used, but the sensitivity of the results to the time periods used and the secular risk of the underlying variables.

Asset related data such as equity returns and fixed interest rates are usually more frequently available, more homogenous and less subjective than insurance line of business related data. Given this, we have used financial time series data in the following analysis.

We decided to investigate a number of questions:

- How do correlation estimates vary with differences in the length of time series data?
- How do annual correlation estimates vary from year to year?

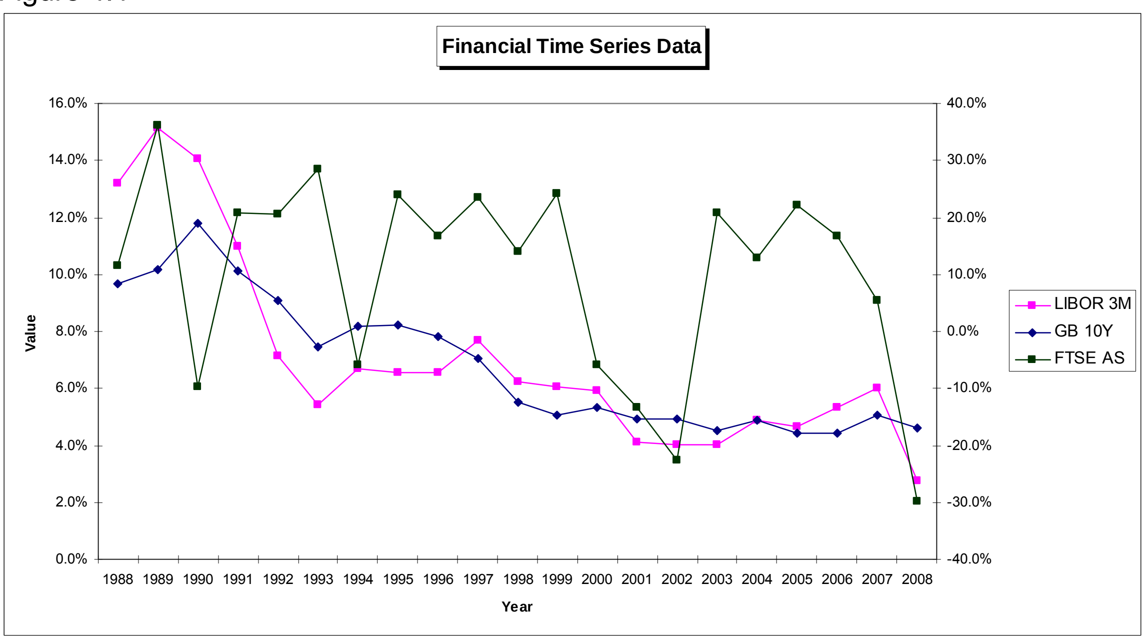
Figures 4.1 and 4.2 show annual time series data for 1988 to 2008 for the following:

- LIBOR 3M rate [13]
- UK 10 year Government Bond yield (“GB 10Y”)
- FTSE All Share Total Return 12-months ending 31/12/YY (“FTSE AS”)
- Credit Spread for AAA rated 10 year corporate bonds as at 31/12/YY (“CS AAA”)
- Credit Spread for BBB rated 10 year corporate bonds as at 31/12/YY (“CS BBB”)

4.3.1 Estimating correlation coefficients with variation in the length of time series data

Figure 4.1 shows a line graph of the values for the first 3 indices. Figure 4.2 shows a graph for GB 10Y vs CS AAA vs CS BBB.

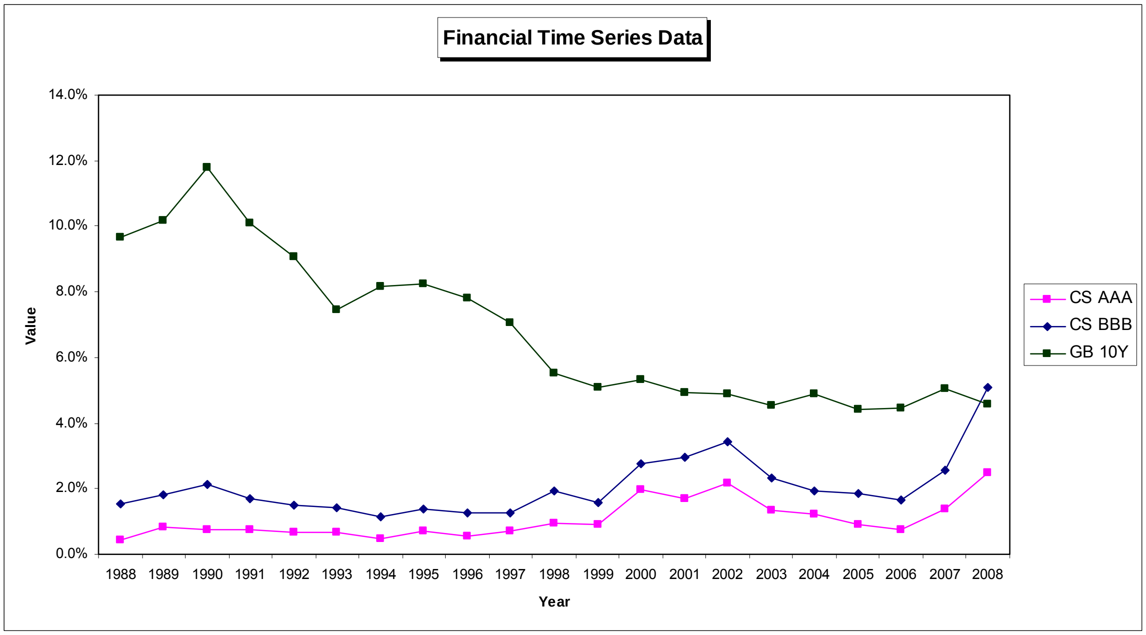
Figure 4.1



[13] The time series data was obtained from Bloomberg with the exception of CS AAA and CS BBB data which was obtained from Iboxx.

On inspection of Figure 4.1 it is clear that the FTSE AS annual return shows a cycle with peaks and troughs over the period 1988 to 2008. Both LIBOR 3M and GB 10Y feature downward trends over the same period and furthermore the LIBOR 3M and GB 10Y graphs crossover at a number of different points, indicating the changing shapes in the yield curve, either upward sloping, downward sloping or inverted yield curves.

Figure 4.2



For Figure 4.2 however, the general decrease in GB 10Y contrasts with the relatively flat credit spreads until 2006 onwards when there has been a sharp rise in the values of both CS AAA and CS BBB reflecting the changing market dynamics following the onset of the current financial crisis.

These relationships between key risk variables changing over time have consequences when one is faced with the estimation of the pairwise correlation coefficients. These comments are not restricted to asset risks in isolation as an insurance liability or combined asset / liability time series graphs often show similar patterns.

Pairwise correlation coefficients have been estimated between the five different risk types assuming four different time periods, namely (i) 20 years to 2007, (ii) 15 years to 2007, (iii) 10 years to 2007 and (iv) 5 years to 2007. The year 2007 has been chosen rather than 2008 so as to minimise as far as possible the influence of the recent financial crisis on the results comparison. Ideally, monthly data would have been better but for some of the risks the data was only available on an annual basis for time periods longer than 10 years ago. However, further analysis in the next section on a limited subset of these risks has been performed using monthly data.

With only 5 data points for the 5 year period 2003 to 2007 the results will be very sensitive to sampling error, as will the results for the 10 year period, although to a more limited extent. The results are presented in Table 4.1

The four sets of tables clearly show not only the sensitivity of the pairwise correlation coefficients to the time periods chosen for the analysis, but how on some occasions there is a reversal of the correlation signs. These results in conjunction with figures 4.1 and 4.2 illustrate the sensitivity of correlation estimates to the choice of data period and perhaps the limitations of a correlation coefficient as measures of dependency.

Table 4.1

20 Years	Start	1988			
	End	2007			
	LIBOR 3M	GB 10Y	CS AAA	CS BBB	FTSE AS
LIBOR 3M	100.0%	86.4%	-46.9%	-28.3%	18.7%
GB 10Y	86.4%	100.0%	-59.1%	-43.7%	14.7%
CS AAA	-46.9%	-59.1%	100.0%	93.8%	-59.6%
CS BBB	-28.3%	-43.7%	93.8%	100.0%	-67.5%
FTSE AS	18.7%	14.7%	-59.6%	-67.5%	100.0%

15 Years	Start	1993			
	End	2007			
	LIBOR 3M	GB 10Y	CS AAA	CS BBB	FTSE AS
LIBOR 3M	100.0%	68.7%	-61.1%	-68.5%	34.9%
GB 10Y	68.7%	100.0%	-60.2%	-64.8%	20.1%
CS AAA	-61.1%	-60.2%	100.0%	97.0%	-71.5%
CS BBB	-68.5%	-64.8%	97.0%	100.0%	-72.5%
FTSE AS	34.9%	20.1%	-71.5%	-72.5%	100.0%

10 Years	Start	1998			
	End	2007			
	LIBOR 3M	GB 10Y	CS AAA	CS BBB	FTSE AS
LIBOR 3M	100.0%	62.3%	-38.1%	-45.7%	33.9%
GB 10Y	62.3%	100.0%	27.4%	21.0%	-31.1%
CS AAA	-38.1%	27.4%	100.0%	94.9%	-89.3%
CS BBB	-45.7%	21.0%	94.9%	100.0%	-92.9%
FTSE AS	33.9%	-31.1%	-89.3%	-92.9%	100.0%

5 Years	Start	2003			
	End	2007			
	LIBOR 3M	GB 10Y	CS AAA	CS BBB	FTSE AS
LIBOR 3M	100.0%	62.6%	-0.2%	17.6%	-85.1%
GB 10Y	62.6%	100.0%	71.5%	62.6%	-93.2%
CS AAA	-0.2%	71.5%	100.0%	89.0%	-48.8%
CS BBB	17.6%	62.6%	89.0%	100.0%	-50.8%
FTSE AS	-85.1%	-93.2%	-48.8%	-50.8%	100.0%

4.3.2 Estimating correlation coefficients using monthly time series data

Further analysis has been performed on the following three sets of monthly time series data for the years 1999 to 2008:

- LIBOR 3M rate as at the end of each month (same notation as before)
- FTSE 10 Year Gilt Yield as at the end of each month (“FTSE G 10Y”)
- FTSE All Share Total Return on a rolling 12-months basis as at the end of each month (same notation as before)

Table 4.2 investigates the correlation between LIBOR 3M and FTSE G 10Y under a number of different calculation bases.

- Column "12 Mths" shows the correlation estimated for each year in isolation using 12-months worth of data. For example for year 2005, 81.9% is based on the monthly data from Jan 2005 to through to Dec 2005.
- Column "Cum YE 08" shows the correlation estimated using monthly data assuming a start year (column "Year") through to Dec 2008. For example for year 2005, 74.5% is based on monthly data from Jan 2005 to Dec 2008. This calculation imitating the often used process of estimating correlations from the "last x years of data".
- Column "Cum YE 07" is a similar calculation to Cum YE 08 but assuming monthly data through to Dec 2007.

The latter 2 calculations are also shown in Figure 4.3. Observations of note are:

- Annual correlations are very volatile from year to year.
- There is more stability in using cumulative monthly data as per Cum YE 08 or Cum YE 07 but even here there is a fair degree of volatility. Both of these calculations appear to indicate a long-term average of around 50%.

However, these calculations are looking at correlations between interest rates of different terms. Figure 4.4 shows similar calculations to Figure 4.3 for LIBOR 3M vs FTSE AS but this time using 20 years of data. From inspection of this figure it is clear how volatile even these calculations are based on the longer time series data set.

Table 4.2

LIBOR 3M vs FTSE G 10Y					
Year	**12 Mths**	**Years Data**	**Cum YE 08**	**Years Data**	**Cum YE 07**
1999	21.7%	10	50.5%	9	52.8%
2000	64.0%	9	50.6%	8	53.3%
2001	32.5%	8	36.0%	7	35.3%
2002	89.2%	7	35.5%	6	32.9%
2003	56.2%	6	50.0%	5	48.5%
2004	1.7%	5	54.2%	4	49.2%
2005	81.9%	4	74.5%	3	75.5%
2006	50.1%	3	71.2%	2	67.3%
2007	-13.4%	2	70.5%	1	-13.4%
2008	81.8%	1	81.8%		

Figure 4.3

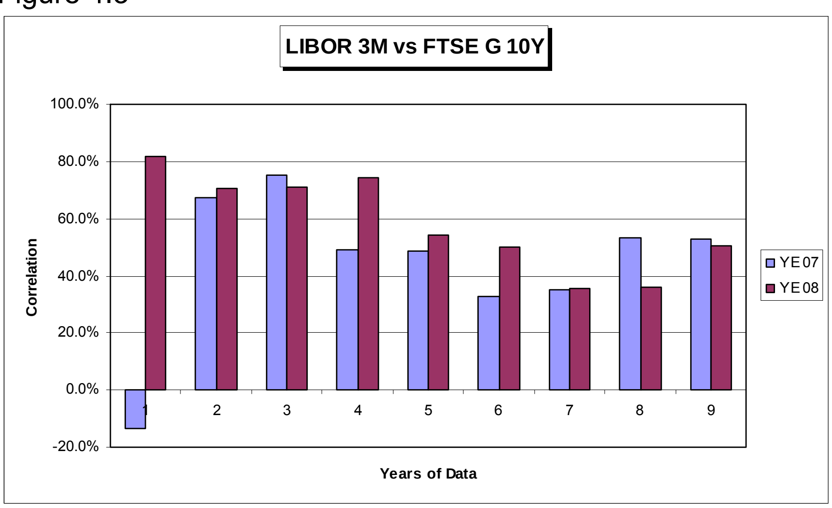
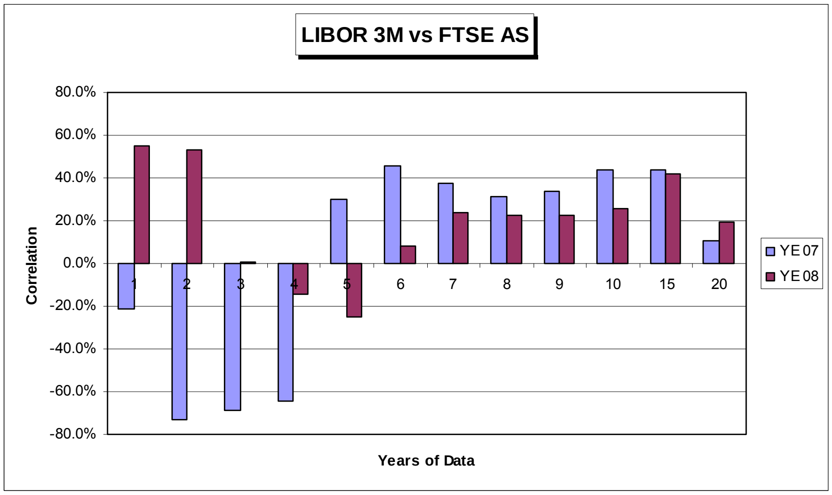


Figure 4.4



4.4 Copula Parameterisation

The practical issues arising when trying to estimate a correlation coefficient from real data were highlighted in section 4.3. With the parameterisation of copulas there is a further issue of how to parameterise the factors that determine the degree of tail dependency between risks in extreme loss scenarios.

For the t copula (or its IT copula extension) one has to estimate degrees-of-freedom parameters, whether it is just one parameter as for the more straightforward t-copula, or n parameters for the more complex IT-copula. One way of doing this is to estimate the tail dependencies for each pair of risk factors from historical data or expert opinion / judgement, and then use an iterative algorithm which finds the set of degrees-of-freedom parameters implying the closest possible tail dependencies using the approaches described later in this section.

If we aim to model tail dependency more accurately, whichever type of copula we are using, we will need to estimate extra parameters which influence the tail dependency.

Estimating copula parameters needs good quality data. Problems arising include:

- Data may not be available, and if available, either of poor quality or incomplete. Asset related data such as returns for equity and fixed interest rate risk are generally more readily available and homogenous than some of the data related to insurance line of business related risks.
- Data may be sparse due to the low frequency of the risk e.g. credit default risk.
- The frequency of the data will also be dependent on the nature of the risks. Asset related time series data is available at least daily whereas insurance loss data may typically only be available on either a quarterly or annual basis.
- If we try to estimate parameters from real data, then we are going to have to use data from economic periods where extreme events occurred and by definition these very low frequency events are scarce data sets.
- Managing the difficult trade-off of having a long enough historical time series to be representative of such events versus the potential secular risk arising.

4.5 Tail Dependency

In simple terms, the relationships between risks may change under the extreme scenarios that could give rise to the financial failure of an insurance company. Historical data based on observations in normal market conditions may indicate that correlations between certain risks are low, but in times of stress such as catastrophes or liquidity crisis many markets and risks could be adversely affected at the same time leading to much higher observed dependency between risks.

A common approach used by some insurance companies to reflect such tail dependency when performing a variance-covariance matrix economic capital calculation is to use correlations that are larger than an average correlation that has been estimated from empirical data in normal market conditions.

The SCR capital aggregation formula within the Solvency II Pillar I framework illustrates this line of reasoning. For example, paragraph 1.84 (p20) of QIS3 Calibration of the underwriting risk, market risk and MCR (2007) [6] says:

" In view of the insufficiency of currently available data, the setting of these correlation coefficients will necessarily include a certain degree of judgement. This is also true because, when selecting correlation coefficients, allowance should be made for non-linear tail correlation, which is not captured under a "pure" linear correlation approach. To allow for this, the correlations used should be higher than simple analysis of relevant data would indicate. "

Furthermore, paragraph TS.VIIIA.7, (p124) says:

" For the aggregation of the individual risk modules to an overall SCR, linear correlation techniques are applied. The setting of the correlation coefficients is intended to reflect potential dependencies in the tail of the distributions, as well as the stability of any correlation assumptions under stress conditions. "

The usage of higher correlations to reflect high tail dependency is a practical, transparent and intuitive method. In addition, the economic capital will be larger than would be the case from using so-called 'average' correlations as assessed within normal market conditions.

However there are certain theoretical shortcomings when higher correlations are used in conjunction with the variance-covariance matrix approach to capital aggregation:

- Higher correlations do not model tail dependency. Tail dependency is a measure of dependence between the risk factors defined using a limit. If you use a Gaussian copula, its tail dependence will be zero irrespective of whether you are using higher correlation parameters or not.
- There are usually no theoretical foundations for their selected values.
- Their implied 'tail correlation' [14] are sensitive to the underlying marginal risk distributions, especially when risks are not normally distributed. In fact in certain scenarios you can get counter-intuitive results.

[14] The implied 'tail correlations' referred to here are explained in detail in section 5.4

4.6 Variance-covariance matrix cross-terms

As stated in section 3.5 there are many financial institutions, in particular large insurance groups that consist of various subsidiaries, business units ("BU") or similar organisational subdivisions. Not only is there a need to calculate economic capital at an individual BU level but also at an overall aggregate level as well.

One of the more common approaches to economic capital modelling, that enables an insurance group to report economic capital at both an (i) aggregate capital level and (ii) BU capital level, is through the use of an enlarged correlation matrix. Data in such a matrix is of the form "Organisation / Risk" e.g. UK /Equity or France/Fixed Interest etc.

In addition, if an insurance company is modelling dependency within an internal model through the use of either a Gaussian or a t copula then it will need the aggregate correlation matrix as described above.

The first stage in the derivation of the overall aggregate correlation matrix is the defining of the correlation matrix at an individual BU level. Once this has been done there is then the issue of the correlation matrix data entries for the cross-terms such as UK / Equity, France / Fixed Interest etc.

.

4.6.1 Example

Let us consider a simple example of a company with 2 business units A and B.

Each BU has 2 risk factors: a_1 and a_2 in BU A and b_1 and b_2 in BU B.

Le the correlation matrix for BU A be $\begin{bmatrix} 1 & a_{12} \\ a_{12} & 1 \end{bmatrix}$ and

the correlation matrix for BU B be $\begin{bmatrix} 1 & b_{12} \\ b_{12} & 1 \end{bmatrix}$.

Then the enlarged correlation matrix for the whole company would look like:

$$\begin{bmatrix} 1 & a_{12} & ? & ? \\ a_{12} & 1 & ? & ? \\ ? & ? & 1 & b_{12} \\ ? & ? & b_{12} & 1 \end{bmatrix}$$

For example an insurance company might have 15 BUs each of which has 10 separate risk categories. In this case the enlarged correlation matrix will be of size 150 x 150.

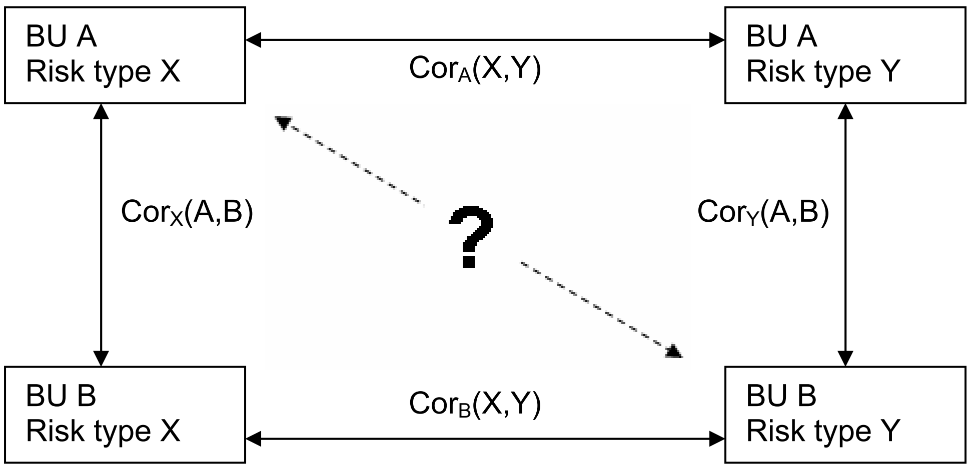
There are several questions immediately raised by this approach:

- How does one estimate cross terms such as the correlation between a_1 and b_1 etc.
- Once the cross-terms are filled in, we still need to make sure that the resulting correlation matrix is PSD (see section 4.8).

One possible approach to this common problem was proposed by Group Consultaf in 2005 [7]

4.6.2 <u>Group Consultatif example of calculating cross-terms</u>

Consider 2 risk types X and Y, and 2 business units BU A and BU B.



The question arises of how to estimate the cross-term identified by **?**.

The proposed approximation of the correlation between risk type X in BU A ("X_A") and risk type Y in BU B ("Y_B") is given by:

Correlation $\rho(X_A, Y_B) = \dfrac{Cor_X(A,B) + Cor_Y(A,B)}{2} \times \dfrac{Cor_A(X,Y) + Cor_B(X,Y)}{2}$.

One of the conditions on the correlation factor we are trying to estimate is expressed by the following double inequality (See Appendix 3 for its derivation):

$$\rho_{X,Y}\rho_{Y,Z} - \sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)} \le \rho_{X,Z} \le \rho_{X,Y}\rho_{Y,Z} + \sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)} \quad (*)$$

Consider the following simple example. There are two permissible routes to get from X_A to Y_B, these are described in the "Path taken" in the following table. Let us assume that $Cor_X(A,B)$ = -0.7, $Cor_Y(A,B)$ = 0.9, $Cor_A(X,Y)$ = 0.7 and $Cor_B(X,Y)$ = 0.8.

Path taken	**Lower Bound**	**Upper Bound**
$Cor_X(A,B)$ & $Cor_B(X,Y)$	-0.988	-0.132
$Cor_A(X,Y)$ & $Cor_Y(A,B)$	0.319	0.941

Then using the proposed Group Consultatif formula ρ (X_A,Y_B) = 0.075.

Using the formula in (*) we have estimated the lower and upper bounds of the correlation according to which of the two possible paths have been taken. The correct correlation between X_A and Y_B should lie in the permitted ranges of each path. Quite clearly this is not the case.

It should be noted that this formula will not always produce sensible answers and illustrates the general issues in estimating cross-terms using similar formulae.

4.7 Positive Semi-Definite Matrices

Once we have filled in the elements of the correlation matrix there is a risk that the resulting matrix may not be Positive Semi-Definite, i.e. it will not be a consistent correlation matrix.

If a matrix is not PSD, an insurance company may typically still use it in the variance-covariance approach to calculate its economic capital. However, the calculation approach used with an inconsistent correlation matrix might lead to counter-intuitive results like a total diversified economic capital higher than the total undiversified capital.

Furthermore, if an insurance company is using a copula approach that requires a correlation matrix e.g. Gaussian copula or T copula then the model will not work as a matrix has to be PSD for it to be inverted as part of the Monte Carlo simulation process.

It is relatively easy to find a solution in mathematical literature which alters a given inconsistent correlation matrix (a non-PSD symmetric matrix with unity diagonal) until it is PSD. See for example the eigenvalue method described in [8] or in Embrechts et al [9].

The problem faced is that if a solution is relatively simple, it can produce quite large changes in the specified correlations and there is no way of knowing in advance if these changes are sensible or what is the impact on the resultant economic capital.

Often insurance companies will want to impose certain constraints on the PSD algorithm e.g. certain key correlation are left unchanged, or can only deviate with a small tolerance. In such cases the algorithms become a lot more complicated and the resulting calculations an iterative process.

4.8 Fitting Copulas to Data

There are two general approaches to estimating copula parameters from a data set. These methods are described comprehensively in Embrechats et al [9]

The two main methods are:

- Maximum Likelihood Estimation
- Method of Moments

4.8.1 Maximum Likelihood Estimation

The maximum likelihood estimation method consists of the following general procedure applied to any data set. For the sake of illustration we will work in 2 dimensions with the two risks X and Y. By extension the same procedure is valid for dimensions > 2:

- Marginal risk distributions are fitted separately for each of X and Y
- The so determined marginal risk distributions are then used to invert the joint data observations of (x,y) into a matrix U of joint values (u,v), where.u and v are defined by $F_X(x) = u$ and $F_Y(y) = v$; x and y are values from X and Y respectively and u and v are values on the unit interval [0,1].
- A copula is then fitted to the matrix U using the method of maximum likelihood

Even when the copula is the main object of interest, one still has to estimate the marginal risk distributions, as the copula data are never directly observed in practice.

The success or otherwise therefore of the statistical quality of the estimates of the copula parameters will depend on the quality of the marginal risk distribution estimates.

The marginal risk distributions in the first step can be chosen using either a:

- Parametric estimation method known as the Inference Functions for Margins ("IFM")
- Non-Parametric estimation with variant of empirical distribution known as Canonical Maximum Likelihood ("CML").

The first of these methods involves the fitting of an appropriate parametric model to the marginal risk data in question using Maximum Likelihood or some other method. However, in the case of very sparse data a variant on this is to make use of a priori marginal risk distribution for the risk of interest.

The second method involves the estimation of an empirical cumulative distribution function from the data, one method involving the divisor being n+1 rather than n such that the maximum or minimum points of the data set do not correspond to either u = 0 or 1. In addition Kernel Smoothing may be adopted to produce a smooth rather than irregular shaped curve.

To implement the Maximum Likelihood method the copula density needs to be derived. The Maximum Likelihood Estimate is generally found by the numerical maximisation of the resulting log-likelihood function.

Appendix 8 provides an illustration of the CML method to 10 years worth of joint monthly data for 3M LIBOR vs FTSE All Share Total Return. The first graph shows the scatter plot of the joint(x,y) values and the second graph the scatter plot of the joint (u,v) after the CML method has been used to fit an empirical CDF to the data. In this example, Kernel smoothing was applied to the empirical CDF.

Both the Gaussian and t copulas were fitted, the following parameters being derived:

- Gaussian copula: Correlation = 29.52%
- t copula: Correlation = 27.61%; t copula d.f. = 2.64

4.8.2 <u>Method of Moments</u>

The method-of-moments consists of using an empirical estimate of Kendall's tau rank correlation (or alternatively Spearman's rank correlation) to derive an estimate of a copula parameter.

This simpler procedure uses sample rank correlation estimates. This method has the advantage that marginal risk distributions do not need to be estimated and consequently inference about the copula does not depend on margin assumptions.

We illustrate the method by estimating the correlation matrix parameter used for Gaussian and t copulas.

Both the Gaussian and t copulas require a linear correlation matrix as an input parameter. This matrix is the linear correlation matrix related to the corresponding multivariate Normal distribution. This means that, if the user is estimating the correlation parameters from sample data, then most likely the estimated correlation coefficient will be different from that required as input to the Gaussian and t copulas to perform Monte Carlo simulation, unless that is we assume that the marginal risk distributions being used are normally distributed.

The theoretically correct approach for Monte Carlo simulation is to calculate the Kendall tau correlation coefficients between risks from sample data and then to convert these into linear correlation parameters using the following formula:

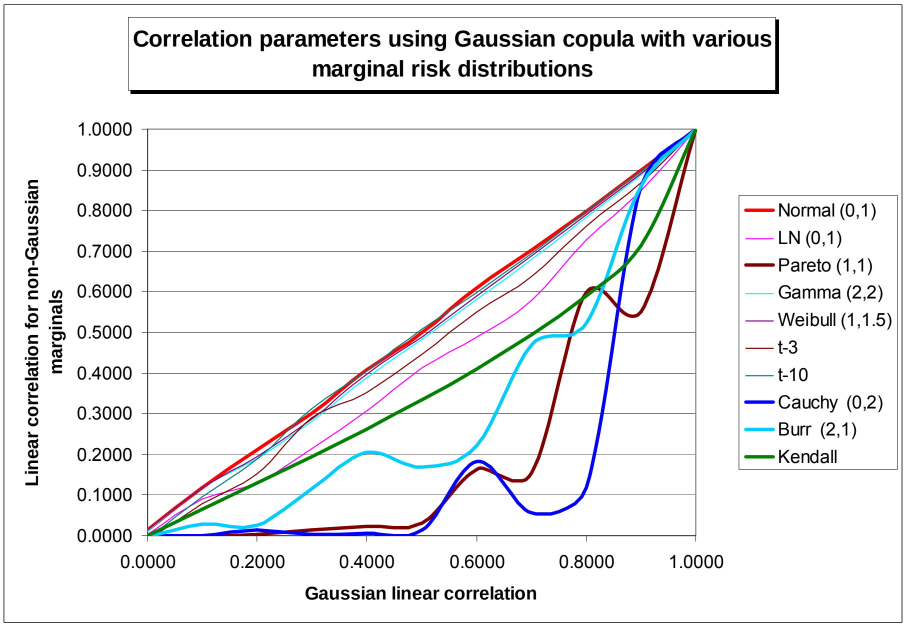
$$\rho_{Gaussian} = \sin\left(\frac{\pi \rho_{Kendall}}{2}\right).$$

This formula works for any elliptical copula, of which the Gaussian and t-copulas are the most notable.

In practice, if sample data arises from marginal risk distributions related to the Normal distribution family such as the Student t, Chi Squared, Gamma and Lognormal distributions then the estimated correlation parameters $\rho_{Gaussian}$ are generally quite close to the estimated linear correlation coefficients from the sample data.

However, the Pareto, Burr and Cauchy distributions are examples of marginal risk distributions where the estimated correlation parameters $\rho_{Gaussian}$ are very different to the estimated linear correlation coefficients from the sample data.

The following graph shows the correlation parameters sampled from two-dimensional distributions using a Gaussian copula with various marginal distributions.



5. Impact of Dependency Modelling on Economic Capital

This section illustrates, through the use of a case study, the impact of different correlation and dependency structures on the economic capital modelling results of a hypothetical insurance organisation ABC Insurance Company.

5.1 ABC Insurance Company

ABC Insurance company writes non-life insurance business in the UK. It calculates economic capital using (i) copula simulation and (ii) a variance-covariance matrix approach. This will involve separate risk distributions by risk category, a correlation matrix and an appropriate copula.

For the sake of convenience and to illustrate the concepts, the marginal risk distributions are all initially assumed to be Lognormal with a Coefficient of Variation ("CV") equal to 25%.[16] Furthermore the pairwise correlation coefficients between all risks are assumed to be identical.

5.1.1 Risk Distributions

The following is a list of the main risk categories and parameters of the risk distributions.

Risk Type	Distribution	Mu	Sigma	E(X)	SD(X)	CV(X)
Equity	Lognormal	7.5706	0.2462	2,000	500	25%
Property	Lognormal	7.5706	0.2462	2,000	500	25%
Interest Rate	Lognormal	7.5706	0.2462	2,000	500	25%
Credit Spread	Lognormal	7.5706	0.2462	2,000	500	25%
Credit Default	Lognormal	7.5706	0.2462	2,000	500	25%
UW - Cat	Lognormal	7.5706	0.2462	2,000	500	25%
UW Non-Cat	Lognormal	7.5706	0.2462	2,000	500	25%
Reserve	Lognormal	7.5706	0.2462	2,000	500	25%
Expenses	Lognormal	7.5706	0.2462	2,000	500	25%
Operational	Lognormal	7.5706	0.2462	2,000	500	25%

5.1.2 Correlation Matrix

The correlation matrix is shown below for a constant correlation of 25%[17] between risks.

CORRELATION MATRIX

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Property	2	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Interest Rate	3	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Credit Spread	4	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25
Credit Default	5	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25
NL UW - Catastrophe	6	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25
NL UW Non-Catastrophe	7	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25
NL Reserving	8	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25
Expenses	9	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25
Operational	10	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00

[16] Results are also shown for the (i) Lognormal distribution with a CV of 50% and the (ii) Normal distribution with a CV of 25%. These other marginal risk distribution parameters are shown in appendix 6.
[17] Results are also shown for other correlations i.e. 10% and 50%.

5.2 Modelling Assumptions

The analysis performed considers:

- The impact of different copulas and parameters on the economic capital modelling results. The modelling shows results for four different copulas, namely:
  - Gaussian Copula
  - t Copula with 10, 5 and 2 degrees of freedom ("d.f.")
- Furthermore, economic capital numbers are also shown using the variance-covariance matrix approach ("V CV") to risk aggregation
- Results are shown at varying percentiles ranging from 75% to 99.95%
- The economic capital is based on a Value at Risk ("VaR") risk measure over 12-months and Capital = Loss (%) – E(Loss) [18] [19]
- The copula simulation results are based on 25,000 simulations for each copula.

5.3 Modelling Results

The results of this modelling exercise are discussed in sections 5.3.1 to 5.3.3.

5.3.1 Lognormal risks (CV = 25%, Correlation = 25%)

Economic Capital - 25% Correlation

Percentile	Return	Gaussian	t - 10 df	t - 5 df	t - 2 df	V CV
75.0%	4	1,760	1,685	1,578	1,421	1,658
90%	10	3,688	3,610	3,582	3,418	3,763
95%	20	4,928	4,906	5,004	4,889	5,182
99%	100	7,423	7,916	8,177	9,049	8,212
99.5%	200	8,391	9,087	10,031	11,052	9,455
99.95%	2,000	11,082	13,926	14,929	18,544	13,468

% change cf Gaussian Copula

Percentile	Return	Gaussian	t - 10 df	t - 5 df	t - 2 df	V CV
75.0%	4	0.0%	-4.2%	-10.3%	-19.3%	-5.8%
90%	10	0.0%	-2.1%	-2.9%	-7.3%	2.0%
95%	20	0.0%	-0.4%	1.6%	-0.8%	5.2%
99%	100	0.0%	6.6%	10.2%	21.9%	10.6%
99.5%	200	0.0%	8.3%	19.5%	31.7%	12.7%
99.95%	2,000	0.0%	25.7%	34.7%	67.3%	21.5%

The economic capital results for the t copula and the variance-covariance matrix approaches have been expressed as percentage changes +/- % of the economic capital arising from use of the Gaussian copula, e.g. at 99% the economic capital using the t copula with 5 d.f. (8,177) is 10.2% higher than that arising from the Gaussian copula at the same percentile (7,423).

If we consider the Gaussian copula as our reference point then quite clearly there is a very wide range of outcomes depending on the approach used to aggregate risks.

[18] We have used the notation 90%, 99.5% to mean the worst 1 in 10, 1 in 200 year result etc.

[19] Loss(%) = Loss amount at % percentile of interest e.g. 99%; E(Loss) = Expected Loss

Comments:

- When the marginal risk distributions are non-normal and positively skewed, then at higher percentiles the V CV capital approach to capital aggregation gives larger economic capital than from use of the Gaussian copula.
- When the marginal risk distributions are normally distributed then the V CV and Gaussian economic capital results should be identical, ignoring sampling error.
- With lognormal marginal risk distributions, the variance-covariance matrix approach to economic capital at the higher percentiles produces similar capital to a t copula. For example at 99% the V CV economic capital is 8,212 which is slightly larger than the economic capital of 8,177 arising from the use of the t copula with 5 d.f.
- As the percentile increases, the larger the implied d.f (lower tail dependency) for the t copula so as to give the same economic capital as the V CV approach.

Appendix 7 provides the same exhibit as above but shows the sensitivity of the results to different correlation coefficients. Results are shown for correlations of 10%, 25% and 50%. When the correlation increases, there is a decrease in the economic capital margin, for the t copula and V CV approaches, over the Gaussian copula method

5.3.2 Lognormal risks (CV = 25% and 50%, Correlation = 25%)

A natural question to ask is what is the sensitivity of the economic capital to variation in the CV of the underlying marginal risk distributions. In this simple example, the CV of the Lognormal distribution has been doubled form 25% to 50%.

Economic Capital - 25% Correlation						**CV 25%**
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-4.2%	-10.3%	-19.3%	-5.8%
90%	10	0.0%	-2.1%	-2.9%	-7.3%	2.0%
95%	20	0.0%	-0.4%	1.6%	-0.8%	5.2%
99%	100	0.0%	6.6%	10.2%	21.9%	10.6%
99.5%	200	0.0%	8.3%	19.5%	31.7%	12.7%
99.95%	2,000	0.0%	25.7%	34.7%	67.3%	21.5%

Economic Capital - 25% Correlation						**CV 50%**
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-4.7%	-11.2%	-24.4%	-17.0%
90%	10	0.0%	-0.6%	-3.3%	-5.5%	0.8%
95%	20	0.0%	3.4%	0.4%	3.7%	7.3%
99%	100	0.0%	6.0%	10.9%	23.5%	18.2%
99.5%	200	0.0%	11.5%	14.1%	32.9%	24.0%
99.95%	2,000	0.0%	13.8%	29.1%	57.1%	38.9%

The economic capital in the tables is expressed as percentage changes +/- % of the economic capital arising from use of the Gaussian copula.

Comments:

- At the higher percentiles, 95% and above, a larger CV results in a larger economic capital margin for the V CV approach compared to the Gaussian copula capital.
- Furthermore, with a higher CV the V CV approach results in a lower implied d.f for the equivalent t copula giving the same economic capital as the V CV approach. For example, at 99% with a CV of 25% the V CV approach is similar to a t copula with 5 d.f. whereas with a CV of 50% the V CV approach looks to be similar to a t Copula with d.f. about half-way between 2 and 5 d.f.

5.3.3 Normal vs Lognormal risks (CV = 25%, Correlation = 25%)

The analysis here considers the sensitivity of the economic capital results to variation in the assumed marginal risk distributions. A comparison has been made of the results arising from use of the Normal distribution compared to those arising from assuming the Lognormal distribution as used in earlier sections.

Economic Capital - 25% Correlation					**Normal**	**CV 25%**
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-2.7%	-7.4%	-15.0%	0.9%
90%	10	0.0%	-1.7%	-3.3%	-7.4%	0.8%
95%	20	0.0%	-1.1%	-0.2%	-3.3%	0.6%
99%	100	0.0%	3.7%	5.4%	13.1%	0.2%
99.5%	200	0.0%	5.5%	12.1%	19.5%	1.0%
99.95%	2,000	0.0%	19.6%	19.5%	35.8%	2.2%

Economic Capital - 25% Correlation					**LogNorm**	**CV 25%**
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-4.2%	-10.3%	-19.3%	-5.8%
90%	10	0.0%	-2.1%	-2.9%	-7.3%	2.0%
95%	20	0.0%	-0.4%	1.6%	-0.8%	5.2%
99%	100	0.0%	6.6%	10.2%	21.9%	10.6%
99.5%	200	0.0%	8.3%	19.5%	31.7%	12.7%
99.95%	2,000	0.0%	25.7%	34.7%	67.3%	21.5%

The economic capital in the tables is expressed as percentage changes +/- % over and above the respective economic capital arising from use of the Gaussian copula.

Comments:

- When the marginal risks distributions are Normal then the V CV approach to capital aggregation should give economic capital results identical to those arising from use of the Gaussian copula.
- The last column in the top table above shows percentage value differences that are nearly, but not exactly equal to 0%. Any differences from 0% are due to sampling error, even with 25,000 simulations.

5.4 Use of 'tail correlations' instead of tail dependence

It was mentioned in section 4.5 that the use of 'tail correlations' within a variance-covariance matrix approach to replicate the effect of tail dependence has some serious methodological limitations.

An alternative approach which would allow companies to get around these difficulties would be to use a copula with positive tail dependency, such as a t-copula. Nevertheless, in reality many companies are currently using higher than average correlations, referred to as 'tail correlations', to reflect views about tail dependence.

A natural question to ask is how these 'tail correlations' compare to the correlation matrices used as an input parameter for a Gaussian and t-copulas. For the purpose of the numeric examples in this section we assume that that the pairwise correlation coefficients between risks are all equal to 25%.

We then calculate the effective correlation coefficient x% such that:
Economic Capital V CV (x%) = Economic Capital t copula (25%) at any given percentile.

We will now compute what these x% correlations should be in the case of two scenarios, (i) Normal marginal risk distributions and (ii) Lognormal marginal risk distributions.

Implied Correlation = V CV Sum				**Normal**	**CV 25%**
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**
75.0%	4	24.3%	22.4%	19.3%	14.5%
90%	10	24.4%	23.2%	22.1%	19.4%
95%	20	24.5%	23.8%	24.4%	22.2%
99%	100	24.9%	27.6%	28.9%	34.9%
99.5%	200	24.3%	28.3%	33.3%	39.4%
99.95%	2,000	23.5%	38.4%	38.2%	52.7%

25% Correlation

Implied Correlation = V CV Sum				**LogNorm**	**CV 25%**
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**
75.0%	4	29.6%	26.2%	21.6%	15.4%
90%	10	23.6%	22.1%	21.6%	18.7%
95%	20	21.5%	21.2%	22.6%	21.0%
99%	100	18.4%	22.4%	24.7%	32.7%
99.5%	200	17.3%	22.2%	29.5%	38.2%
99.95%	2,000	13.3%	27.5%	33.3%	57.4%

25% Correlation

<u>Comments</u>:

- For the Normal distribution the correlations are very close to 25% which is consistent with the V CV approach being equivalent to use of the Gaussian copula.
- For the Lognormal distribution the implied 'tail correlation' is generally lower than that of the Normal distribution at the higher loss percentiles.
- For any given percentile and t copula the implied 'tail correlation' is sensitive to the choice of distribution, which in these examples have quite modest CVs.
- With the Lognormal distribution, for any given percentile e.g. 99.5% the range of implied 'tail correlation' varies from 22.2% to 38.2% for a t copula with 10 d.f. to one with 2 d.f. There is also a marked variation by percentile. In the case of a t copula with 10 d.f. the 'tail correlation' is lower than the average correlation at 99.5%.

The conclusion we draw from these results is that it can be very difficult to come up with an adequate set of 'tail correlations' without any regard for the theoretical underpinnings. The results can vary quite significantly depending on what level of tail dependency the company is aiming to model.

6. Communication of Economic Capital Modelling Dependency Impacts

One of the key challenges facing organisations is how to communicate the effect of different dependency structures on the economic capital results.

This section describes some very simple measures that could be adopted by firms in their communication either internally to the board of directors and senior management or externally to various stakeholders.

The methods outlined may be of use in the determination of appropriate copulas and their parameters if similar calculations are made from empirical data.

For the sake of simplicity we shall continue with the use of ABC Insurance Company to illustrate the calculations for each of the measures of interest.

6.1 Communication Measures

The following is a list of possible measures. It is not exhaustive but illustrative of different approaches, some more complex than others:

- Economic Capital Aggregation
- Joint Probability Density Function
- Scatter Plot
- Joint Excess Probability
- Tail Concentration Function
- Kendall Tau Correlation
- Coefficient of Tail Dependence
- Implied 'Gaussian' Correlation

There are three possible levels of data granularity:

- Comparisons made at an aggregate level e.g. total Economic Capital
- Comparisons made between a pair of risks e.g. "Scatter Plot "
- Comparisons made between all risk pairs e.g. "Tail Concentration Function "

The numerical exhibits that follow are based on simulated output assuming:

- 10 risk categories for ABC Insurance Company
- Pairwise correlation coefficients of 25% between all risks
- t copula with 5 d.f. unless otherwise stated e.g. in section 6.3 Economic Capital Aggregation where other copula assumptions are assumed
- 25,000 simulations

6.2 Economic Capital Aggregation

Description

The objective here is to calculate the total economic capital at the different percentiles using a number of different risk aggregation techniques. In the exhibit that follows capital numbers are shown for the (i) Gaussian copula, (ii) t Copula at 10, 5 and 2 d.f. and (iii) Variance-covariance matrix approach to capital aggregation.

The exhibit shown is the same as that shown in section 5.3.1.

Exhibit

Economic Capital - 25% Correlation						
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	1,760	1,685	1,578	1,421	1,658
90%	10	3,688	3,610	3,582	3,418	3,763
95%	20	4,928	4,906	5,004	4,889	5,182
99%	100	7,423	7,916	8,177	9,049	8,212
99.5%	200	8,391	9,087	10,031	11,052	9,455
99.95%	2,000	11,082	13,926	14,929	18,544	13,468

% change cf Gaussian Copula						
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-4.2%	-10.3%	-19.3%	-5.8%
90%	10	0.0%	-2.1%	-2.9%	-7.3%	2.0%
95%	20	0.0%	-0.4%	1.6%	-0.8%	5.2%
99%	100	0.0%	6.6%	10.2%	21.9%	10.6%
99.5%	200	0.0%	8.3%	19.5%	31.7%	12.7%
99.95%	2,000	0.0%	25.7%	34.7%	67.3%	21.5%

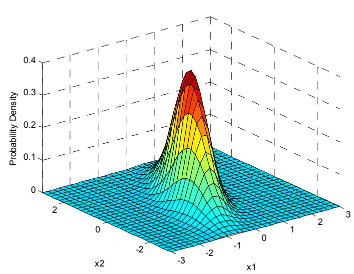
Advantages
- It is relatively simple to understand
- It is possible to directly measure the financial impact on a company

Disadvantages
- One has no information of what is happening at an individual risk category level at each percentile of interest
- The calculations are more computer intensive than those that will be discussed in the following sections.

6.3 Joint Probability Density Function

Description
The Joint Probability Density function is a 3 dimensional representation of the plot of values (u,v) of the risk factor distributions X and Y, where u and v are defined by the relationships $F_X(x) = u$ and $F_Y(y) = v$. A greater density of points represented by a larger value of the PDF. When there is tail dependency one would expect to see a greater density in the region of (1,1).

Exhibit



Advantages

- It is relatively simple to understand.
- The exhibits are relatively easy to create.

Disadvantages

- Sampling error may distort the presence or otherwise of 'tail' dependency strength
- There is no numerical measure that reflects the degree of dependency between risks
- One can only use this method for a pair of risks at a time

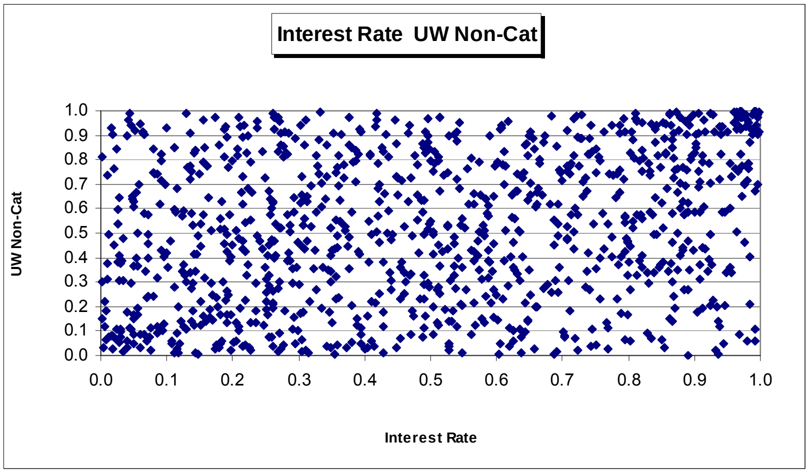
6.4 Scatter Plot

Description

A scatter plot involves a plot of values (u,v) of the risk factor distributions X and Y, where u and v are defined by the relationships $F_X(x) = u$ and $F_Y(y) = v$. Furthermore, x and y are values from X and Y respectively and u and v are values on the interval [0,1].

The extent of the clustering of points in the region of (1,1) indicates the level of 'tail' dependency between 2 risks.

Exhibit



Advantages

- It is relatively simple to understand.
- The exhibits are very easy to create.

Disadvantages

- Sampling error may distort the presence or otherwise of 'tail' dependency strength
- There is no numerical measure that reflects the degree of dependency between risks
- It may be difficult to distinguish a pair or risks with higher tail dependence from a pair of risks with higher correlation but lower tail dependence.
- One can only use this method for a pair of risks at a time

6.5 Joint Excess Probability

Description
For a pair of risks, the Joint Excess Probability is the joint probability that 2 risks are either greater or lower than some deemed threshold. Notation wise:

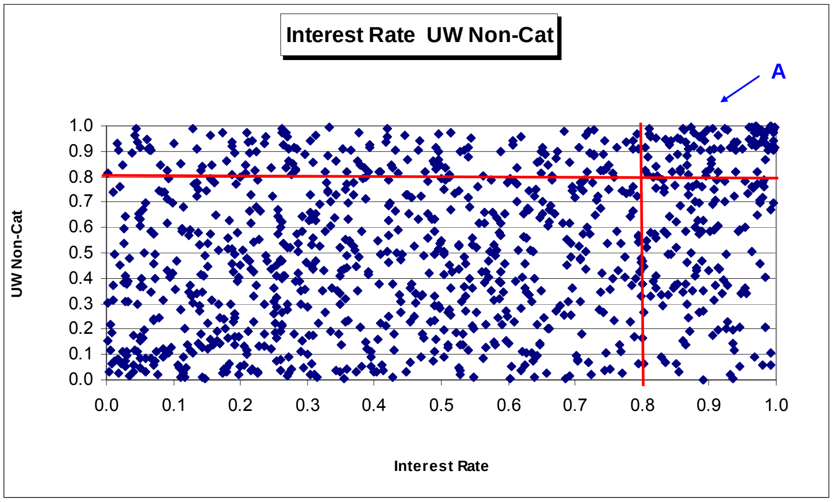
- RJEP(z) = P(u>z, v>z) [22]
- LJEP(z) = P(u<z, v<z)

where: u and v are defined by $F_X(x) = u$ and $F_Y(y) = v$; x and y are values from X and Y respectively and u and v are values on the unit interval [0,1].

For independence the values of RJEP(z) and LJEP(z) are $(1-z)^2$ and z^2 respectively.

Exhibit
The following is an illustration of the RJEP(z) concept using a scatter plot of the simulation output for two hypothetical risks X and Y.

e.g. RJEP(0.8) = No. of Points in A / Total No. of Points (in this case 1,000)



The following diagram is a matrix of values of the function RJEP(0.95) for each of the pairwise combinations of the 10 risks.[23]

For comparative purposes the value of RJEP(0.95) = 0.25% is shown where risks are independent of each other, i.e. there is 0% correlation. In this example the values between risk pairs should be identical however the presence of sampling error leads to small differences. With different pairwise correlations between risks the matrix of values becomes more meaningful.

[22] The functions RJEP(z), LJEP(z) are invented notation to distinguish between the top right hand and bottom left hand corners of the area [0,1] x [0,1].
[23] Even with 25,000 simulations sampling error is evident

RJEP(Z): t Copula 5 d.f. z 95.0%

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		0.87%	1.08%	1.07%	1.04%	0.98%	1.00%	0.89%	0.87%	0.98%
Property	2			1.04%	0.98%	0.97%	0.95%	0.98%	1.00%	0.95%	0.96%
Interest Rate	3				1.12%	1.06%	1.03%	1.10%	1.09%	1.04%	1.10%
Credit Spread	4					1.02%	0.99%	1.19%	1.01%	1.07%	1.17%
Credit Default	5						0.97%	1.05%	1.02%	1.00%	1.00%
UW - Cat	6							0.93%	0.91%	0.96%	1.07%
UW Non-Cat	7								0.97%	0.98%	1.09%
Reserve	8									1.04%	1.05%
Expenses	9										0.98%
Operational	10										
									Independence		0.25%

Advantages

- It is practical and the concept is relatively easy to understand
- The calculation is relatively easy to perform
- It allows the quantification of the level of dependence at a given percentile in a way which is both mathematically tractable, and simple to understand
- It provides a consistent methodology for comparing the relative strength of dependency between 2 or more risks whether the dependence between them is expressed using copulas or correlations, or in any other way
- For more than two risks it is possible to estimate RJEP(z) and LJEP(z) for each pair of risks and present the information as a matrix of values for all risks or a pair of risks

Disadvantages

- For most of practitioners used to linear correlations this would be a new concept and some confusion between the two numbers is possible. In particular, it could be mistaken to be a 'tail correlation', i.e. the level of correlation in the tail. In fact, the RJEP(z) and LJEP(z) functions are probabilities, i.e. take values between 0 and 1 whereas a correlation coefficient takes values between -1 and 1.
- It is difficult to translate a value of RJEP(z) or LJEP(z) into a number that is commonly understood e.g. linear correlation, or its equivalent at the 'tails'.
- The value of RJEP(z) or LJEP(z) depends on the marginal risk distributions, not only on the dependence structure between the risks.
- Sampling error may distort the presence or otherwise of 'tail' dependency strength

6.6 Right Tail Concentration Function

Description

For a pair of risks, the strength of 'tail' dependence between risk factors can be defined using the Right and Left Tail Concentration Functions R(z) and L(z) [10] respectively as follows:

- Right Tail Concentration Function: $R(z) = P(u>z / v>z) = P(u>z, v>z) / P (v>z)$
- Left Tail Concentration Function: $L(z) = P(u<z / v<z) = P(u<z, v<z) / P (v<z)$

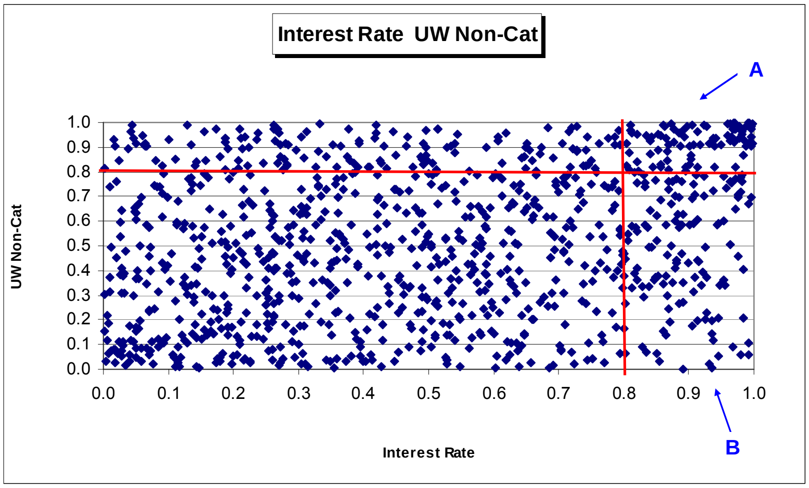
where: u and v are defined by $F_X(x) = u$ and $F_Y(y) = v$; x and y are values from X and Y respectively and u and v are values on the unit interval [0,1].

Technically speaking if we have more than two risks, e.g. three risks, then R(z) should be defined as $R(z) = P(u>z / v>z, w>z)$ etc.

Exhibit

The following is an illustration of the R(z) concept using a scatter plot of the simulation output for two hypothetical risks X and Y.

e.g. R(0.8) = No. of Points in A / (Total No. of Points (A + B))



The following is a matrix of values of the function R(0.95) for each of the pairwise combinations of the 10 risks. [24] In this example the values between risk pairs should be identical however the presence of sampling error leads to small differences. With different pairwise correlations between risks the matrix of values becomes more meaningful.

For comparative purposes the value of R(0.95) = 5.0% is shown where risks are independent of each other, i.e. there is 0% correlation.

R(Z): t Copula 5 d.f.				z	95.0%						
	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		17.58%	21.85%	21.61%	20.89%	19.68%	20.16%	17.98%	17.58%	19.68%
Property	2			21.18%	20.03%	19.87%	19.38%	20.03%	20.36%	19.38%	19.71%
Interest Rate	3				21.20%	19.98%	19.45%	20.82%	20.67%	19.68%	20.89%
Credit Spread	4					19.06%	18.46%	22.20%	18.91%	19.96%	21.82%
Credit Default	5						19.13%	20.71%	20.16%	19.61%	19.76%
UW - Cat	6							18.87%	18.46%	19.35%	21.62%
UW Non-Cat	7								19.27%	19.51%	21.74%
Reserve	8									21.06%	21.38%
Expenses	9										19.48%
Operational	10										
									Independence		5.0%

[24] Even with 25,000 simulations sampling error is evident

Advantages

- It is practical and the concept is relatively easy to understand
- The calculation is relatively easy to perform
- It allows the quantification of the level of dependence at a given percentile in a way which is both mathematically tractable, and simple to understand
- It is closely linked to another important copula parameter: "Coefficient of Tail Dependence" (section 6.8) which is a limiting case of the tail concentration function
- It provides a consistent methodology for comparing the relative strength of dependency between two or more risks whether the dependence between them is expressed using copulas or correlations, or in any other way
- For more than two risks it is possible to estimate R(z) and L(z) for each pair of risks and present the information as a matrix of values for all risks or a pair of risks

Disadvantages

- For most of practitioners used to linear correlations this would be a new concept and some confusion between the two numbers is possible. In particular, it could be misunderstood to be a 'tail correlation', i.e. the level of correlation in the tail. In fact, the tail concentration functions are different mathematical objects: they are probabilities, i.e. take values between 0 and 1 whereas correlation coefficient takes values between -1 and 1.
- It is difficult to translate a value of R(z) or L(z) into a number that is commonly understood i.e. linear correlation
- Sampling error may distort the presence or otherwise of 'tail' dependency strength

For comparative purposes the tables of matrices for R(0.95) and RJEP(0.95) are shown for both the T Copula with 5 d.f. and the Gaussian Copula in appendix 4.

6.7 Kendall Tau Correlation

The concept and definition of the Kendall tau rank correlation (or simply the Kendall tau) was discussed in section 2.4.

It is a type of rank correlation, i.e. a correlation coefficient which depends on the ranking of data points, not on their values. It's values lie between -1 and 1.

In this example the values between risk pairs should be identical however the presence of sampling error leads to small differences. With different pairwise correlations between risks the matrix of values becomes more meaningful.

Exhibit

Kendall Tau: t Copula 5 d.f.

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		17.62%	16.72%	17.74%	18.31%	16.29%	17.73%	16.72%	16.49%	17.50%
Property	2			15.38%	15.92%	16.39%	15.72%	17.20%	16.43%	16.29%	17.25%
Interest Rate	3				16.02%	16.67%	17.29%	17.66%	16.68%	16.71%	17.02%
Credit Spread	4					16.27%	17.17%	16.70%	15.94%	15.96%	16.02%
Credit Default	5						16.35%	17.22%	15.75%	15.68%	17.25%
UW - Cat	6							15.61%	15.16%	16.71%	17.90%
UW Non-Cat	7								15.56%	15.86%	17.63%
Reserve	8									16.73%	17.26%
Expenses	9										16.19%
Operational	10										

Advantages

- It is intuitively simple and the concept is relatively easy to understand.
- It is more intuitive to someone who is used to regular linear correlations than other measures such as tail dependence
- It does not depend on the absolute value of observations, which means it should be better dealing with data outliers.
- It provides a consistent methodology for comparing the relative strength of two or more different random variables with any type of dependency structure.
- It is possible to represent the information either as a matrix of values for all risks or a pair of risks

Disadvantages

- The calculation is slightly more challenging than with other methods
- It is difficult to translate values Kendall Tau into numbers that are commonly understood e.g. linear correlation, or its equivalent at the 'tails'.
- It does not identify trends like an ever-increasing "strength of relationship" with an increasing percentile, i.e. it is just a scalar measure, like the linear correlation.

Kendall Tau matrices are shown for both the t Copula with 5 d.f. and the Gaussian Copula in appendix 5.

6.8 Coefficient of Tail Dependence

Description

The Coefficient of Tail Dependence between two risks is an asymptotic measure of the dependence in the tails of the bivariate distribution (X,Y).

The mathematical definition of tail dependence was discussed in section 2.9.1.

For a multivariate distribution with a Gaussian copula, the tail dependence between any pair of risks is always zero. This is one of the important deficiencies of the Gaussian copula for modelling dependence.

For continuously distributed random variables with the t Copula the Coefficient of Tail Dependence is:

- $\lambda = 2t_{\nu+1}(-(\nu+1)^{0.5}(1-\rho)^{0.5}/ (1+\rho)^{0.5})$

where: ρ is the pairwise correlation coefficient between two risks

These coefficients can be viewed as the limiting conditional probabilities of the functions R(z) and L(z) respectively (see section 6.6)

Exhibit
For t copulas with 10, 5 and 2 d.f. respectively the following is a table of values for λ in the case of a 25% pairwise correlation between risks.

T d.f.	λ
10	2.6%
5	10.7%
2	27.2%

Advantages
- This is the most accurate mathematical measure of the "true" tail dependence between two risks
- It does not depend on the estimated percentile, it is just one single characteristic of a dependence structure, e.g. copula
- It provides a consistent methodology for comparing the relative strength of two or more different copulas
- It is possible to represent the information either as a matrix of values for all risks or a pair of risks

Disadvantages

- It is a relatively new concept and could be counter-intuitive to those who are just familiar with correlations. Moreover, it could actually be confused for 'tail correlation', because λ takes values between 0 and 1 (although correlation can take values between -1 and 1)
- The values of λ are limiting values and do not reflect an ever-increasing value with an increasing percentile.
- $\lambda = 0$ for Gaussian copula
- For the t copula λ is limited by the combination of a degrees of freedom parameter and correlation coefficient. Not all values of λ between 0 and 1 are achievable by fixing one of these two parameters and varying the other. E.g., if you are trying to calibrate a t Copula for two risks with correlation 0.5 and a Coefficient of Tail Dependence of 0.7 by choosing the degrees of freedom parameter, this might not be possible.
- The values use a closed-form solution which may, unless enough simulations are run, provide inconsistent values with values of R(z) from simulated output

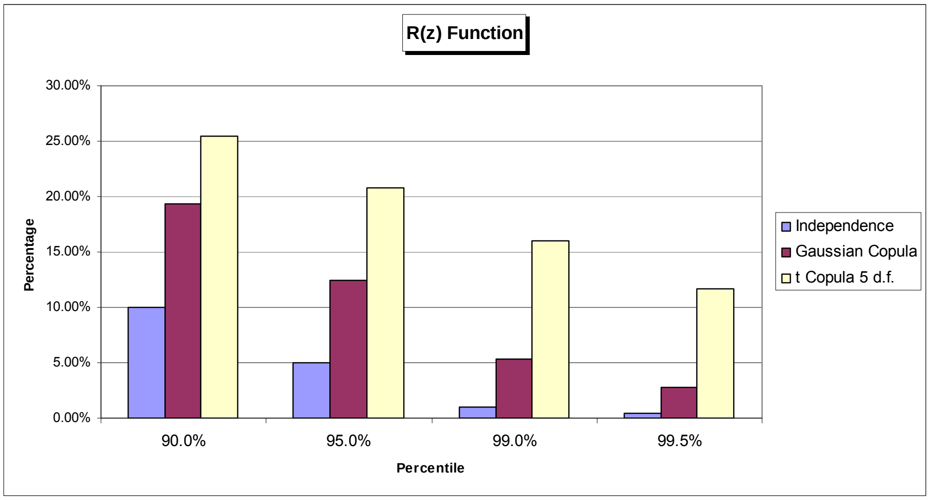
6.9 Implied 'Gaussian' Correlation

Description
This method is really a combination of the methods previously described in sections 6.5 and 6.6 with notable differences:
- R(z) and RJEP(z) values are shown for more than one percentile
- R(z) and RJEP(z) values are shown for 3 different scenarios, (i) Independence (where correlation is 0%) (ii) Gaussian copula and (iii) t copula with 5 d.f.. The latter two using the relevant correlation coefficients.

Exhibit

The graph below shows the values of R(z) for Interest rate risk vs UW Non-Cat risk assuming a correlation of 25%.

On inspection it can be seen that the values of R(z) are greater for the t copula in comparison with the Gaussian Copula; which in turn has larger values of R(z) compared to an assumption of independence between the risks. This observation holds for all percentiles. Furthermore the ratio of R(z) (t copula) / R(z) (Gaussian copula) increases with an increasing percentile.



The numbers behind this exhibit can also be used to determine a so-called 'Implied' Gaussian correlation between a pair of risks at each percentile. For example, at 99.0% the value of R(z) = 15.96% and 5.31% for the t copula with 5 d.f. and Gaussian copula respectively. However, if the linear correlation between these two risks is increased from 25% to 54% then the value of R(z) at 99.0% assuming the Gaussian copula now equals the same value of R(z) = 15.96% as was the case with a linear correlation of 25% and the t copula with 5 d.f. This approach is very sensitive to sampling error.

Advantages

- It is very useful to compare the t copula alongside the Gaussian Copula and the scenario of Independence. In this way one can get a feeling for the degree of tail dependency of the t Copula at each percentile
- Furthermore, showing values at different percentiles is a useful way of comparing how the relative values between the t Copula and the Gaussian copula change with percentile.
- It is easy to calculate and technically correct
- The exhibits are very easy to generate within a simulation model

Disadvantages

- It is difficult to translate values of R(z) and RJEP(z) into numbers that are commonly understood e.g. linear correlation.
- There are not many data points at the extreme percentiles and so the calculation is very sensitive to any sampling error in the joint distribution output.
- One can only use this method for a pair of risks at a time

Conclusions

Dependency is a very complex area of economic capital modelling allowing for a wide choice of various model types and approaches to parameterisation. Issues that arise over a typical 12-month modelling time horizon are compounded when we move into a multi-year model.

Even something intuitively simple as a correlation coefficient can cause serious practical difficulties, including spurious relationships, availability of data and technical constraints. As was mentioned earlier on in the paper a simple scatter plot is likely to lead to a different interpretation than an historical time series representation of the same information.

A few general messages coming out of our work are as follows:

1. A single correlation coefficient is often not enough to describe the dependency between risks in more extreme scenarios, the distribution-based copula approach to modelling dependency can be more meaningful

2. If copulas are used then the selection of an appropriate copula and its parameters should be based on sound analysis and judgement. However, there are considerable issues in trying to parameterise heavy-tailed copulas and so a pragmatic approach is often called, which is touched upon in section 6.

3. A company needs to be extremely careful if it is using higher correlations within the variance-covariance framework as a substitute for tail dependence and copulas. The choice of correlations should not be based on the notion of a prudent margin in the absence of any analytical work underpinning the assumptions made. One needs to remember that a correlation matrix can be calibrated to reflect the average level of dependence related to a certain confidence level, but it will not work with a different probability of loss.

4. Copulas do not model the change of dependency structure over time, in particular at different points in the economic cycle, and in the case of non-life insurance companies the underwriting cycle.

5. Even a simple correlation matrix can cause quite a lot of issues including positive semi-definiteness, high dimensionality and filling in the missing terms.

In our research we have touched upon a number of different topics, some more complex than others. Whilst much valuable work within the sphere of dependency modelling has been accomplished over the last few years more still needs to be done in advance of the implementation of robust models and credible parameters within the Solvency II framework. We would like to think that the actuarial profession will be at the forefront of such developments as they affect insurance organisations.

References

1. Shaw, J.; “Beyond VAR and Stress Testing”; in *VAR: Understanding and Applying Value at Risk,* pp. 221-224; (1997) Risk Publications, London

2. Embrechts P, McNeil A., Straumann A.; “Correlation and Dependence In Risk Management: Properties And Pitfalls”; (1999)

3. Barnett J., Kreps R., Major J.,Venter G.;“Multivariate Copulas for Financial Modelling”; Casualty Actuarial Society, Volume 1/Issue 1; (2007)

4. “QIS4 Technical Specifications (MARKT/2505/08)”; Brussels, (31 March 2008)

5. Tufte, Edward R.;”The Cognitive Style of PowerPoint: Pitching Out Corrupts Within”; Cheshire, Connecticut: Graphics Press (2006)

6. “QIS3: Calibration of the underwriting risk, market risk and MCR”; (April 2007)

7. “Diversification”; Technical paper by the Solvency II Groupe Consultatif Working Group, (2005)

8. Budden M., Hadavas P., Hoffman L.; “On the Generation of Correlation Matrices”; Applied Mathematics E-Notes 8, pp 279-282 (2008)

9. Embrechts, P., Frey, R., McNeil, A.; “Quantitative Risk Management”; (2005)

10. Venter, G.;“Tails of Copulas”; Proceedings of CAS LXXXIX pp. 68 – 113 (2002)

11. Anscombe F.J.; “Graphs in Statistical Analysis”, American Statistician (1973)

12. Kat H.M.;”The Dangers of Using Correlation to Measure Dependence”; (2002)

13. Noether G.E.; “Why Kendall Tau?”; (1986)

14. Embrechts, P., Lindskog, F., McNeil, A. J. ;”Modelling Dependence with Copulas and Applications to Risk Management”; (2001)

15. Joe, H.; “Multivariate Models and Dependence Concepts”; (1997); Chapman & Hall, London.

16. “A Global Framework for Insurer Solvency Assessment”; Report by the Insurer Solvency Assessment Working Party of the International Actuarial Association; (2004)

17. “Enterprise Risk Analysis for Property & Liability Insurance Companies”; (2007); Guy Carpenter

18. “Trying to find the sting in the tail”; www.insuranceriskandcapital.com; (March 2009); Brands Business Ltd 2009

19. “Actuarial Aspects of Internal Models for Solvency II”; Brooks D. et al, Draft paper Presented to the Institute of Actuaries; (23 February 2009)

Appendix 1

Positive Semi-Definite Matrices

Definition

A positive semi-definite (PSD) matrix is similar in many ways to a non-negative real number.

An n × n real symmetric matrix M is positive definite if $z^TMz \geq 0$ for all non-zero vectors z with real entries, where z^T denotes the transpose of z.

Cholesky Decomposition

The importance of PSD matrices in statistics derives from the following property:

If a square matrix A has real entries and is PSD, then A can be decomposed as $A=LL^*$, where L is a lower triangular matrix with real entries, and L^* denotes the transpose of L. This is the Cholesky decomposition.

Any square matrix A with non-zero pivots can be written as the product of a lower triangular matrix L and an upper triangular matrix U; this is called the LU decomposition. However, if A is PSD, we can choose the factors such that U is the transpose of L.

One can intuitively think of Cholesky decomposition as an operation similar to taking a square root in a matrix world.

Example 1

The following diagrams show what at first appearances appear to be perfectly normal matrices. Yet, the first one is PSD and the other is not.

PSD Matrix

Risk	1	2	3	4	5	6
1	1.00	0.20	0.50	0.10	0.20	0.60
2	0.20	1.00	0.20	0.70	0.40	0.10
3	0.50	0.20	1.00	0.50	0.25	0.30
4	0.10	0.70	0.50	1.00	0.10	0.20
5	0.20	0.40	0.25	0.10	1.00	-0.25
6	0.60	0.10	0.30	0.20	-0.25	1.00

Non-PSD Matrix

Risk	1	2	3	4	5	6
1	1.00	0.20	0.50	0.10	0.20	0.60
2	0.20	1.00	-0.20	0.70	0.40	0.10
3	0.50	-0.20	1.00	0.50	0.25	0.30
4	0.10	0.70	0.50	1.00	0.10	0.20
5	0.20	0.40	0.25	0.10	1.00	-0.25
6	0.60	0.10	0.30	0.20	-0.25	1.00

Example 2

The following illustrates the computation of the cholesky decomposition matrix (Table 2) from the original correlation matrix in Table 1. Table 4 is the matrix computed through a multiplication of the cholesky matrix and its transpose (Table 3). As can be seen, and agrees with theory, this calculation gives us the original correlation matrix.

Table 1 - CORRELATION MATRIX

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Property	2	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Interest Rate	3	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Credit Spread	4	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25
Credit Default	5	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25
NL UW - Catastrophe	6	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25
NL UW Non-Catastrophe	7	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25
NL Reserving	8	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25
Expenses	9	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25
Operational	10	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00

Table 2 - CHOLESKY MATRIX

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Property	2	0.25	0.97	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Interest Rate	3	0.25	0.19	0.95	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Credit Spread	4	0.25	0.19	0.16	0.94	0.00	0.00	0.00	0.00	0.00	0.00
Credit Default	5	0.25	0.19	0.16	0.13	0.93	0.00	0.00	0.00	0.00	0.00
NL UW - Catastrophe	6	0.25	0.19	0.16	0.13	0.12	0.92	0.00	0.00	0.00	0.00
NL UW Non-Catastrophe	7	0.25	0.19	0.16	0.13	0.12	0.10	0.91	0.00	0.00	0.00
NL Reserving	8	0.25	0.19	0.16	0.13	0.12	0.10	0.09	0.91	0.00	0.00
Expenses	9	0.25	0.19	0.16	0.13	0.12	0.10	0.09	0.08	0.90	0.00
Operational	10	0.25	0.19	0.16	0.13	0.12	0.10	0.09	0.08	0.08	0.90

Table 3 - TRANSPOSE CHOLESKY MATRIX

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Property	2	0.00	0.97	0.19	0.19	0.19	0.19	0.19	0.19	0.19	0.19
Interest Rate	3	0.00	0.00	0.95	0.16	0.16	0.16	0.16	0.16	0.16	0.16
Credit Spread	4	0.00	0.00	0.00	0.94	0.13	0.13	0.13	0.13	0.13	0.13
Credit Default	5	0.00	0.00	0.00	0.00	0.93	0.12	0.12	0.12	0.12	0.12
NL UW - Catastrophe	6	0.00	0.00	0.00	0.00	0.00	0.92	0.10	0.10	0.10	0.10
NL UW Non-Catastrophe	7	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.09	0.09	0.09
NL Reserving	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.91	0.08	0.08
Expenses	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90	0.08
Operational	10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.90

Table 4 - ORIGINAL MATRIX - CHECK

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Property	2	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Interest Rate	3	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Credit Spread	4	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25	0.25
Credit Default	5	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25	0.25
NL UW - Catastrophe	6	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25	0.25
NL UW Non-Catastrophe	7	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25	0.25
NL Reserving	8	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25	0.25
Expenses	9	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00	0.25
Operational	10	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25	1.00

Appendix 2

Anscombe's Quartet

Description

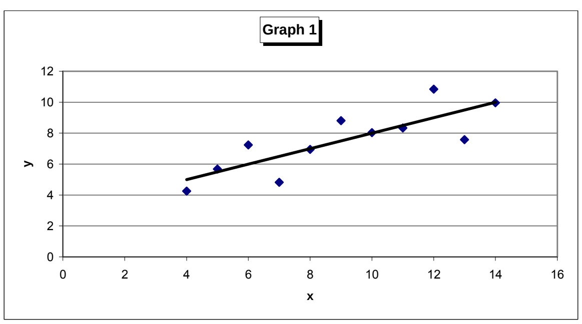
Anscombe's quartet consists of four data sets which have the same identical statistical properties but which are very different to each other when viewed graphically. The different graphs are labelled 1 through to 4. The linear regression line for each set of points is given by y = 3 + 0.5x.

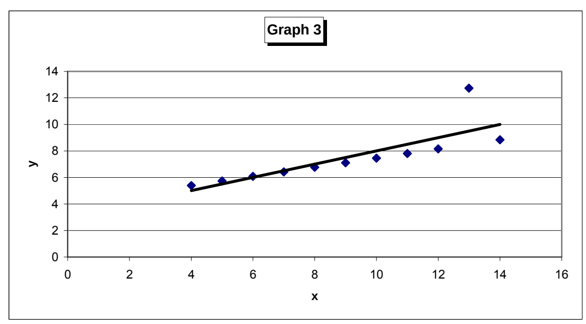
The statistics for all 4 data sets are shown in the following table:

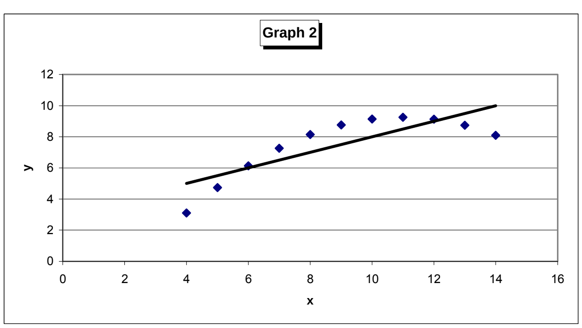
Property	Value
Mean of each x variable	9.0
Variance of each x variable	10.0
Mean of each y variable	7.5
Variance of each y variable	3.75
Correlation between each x and y variable	0.816
Linear regression line	y = 3 + 0.5*x*

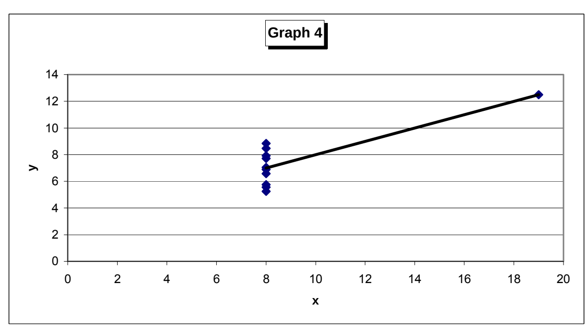
Comments on the Graphs

Graph 1 – What one would expect when considering two correlated variable that follow the assumption of normality.
Graph 2 – The relationship is not linear but an obvious non-linear relationship exists.
Graph 3 – The linear relationship is perfect except for one outlier.
Graph 4 – The relationship between variables is not linear but one outlier is enough to give a correlation of 0.81 and make it appear as though there is one.









Appendix 3

Ranges for estimates of missing correlation

Given two correlation coefficients $\rho_{X,Y}$ and $\rho_{Y,Z}$, the valid range for $\rho_{X,Z}$ is given by the following double inequality:

$$\rho_{X,Y}\rho_{Y,Z} - \sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)} \le \rho_{X,Z} \le \rho_{X,Y}\rho_{Y,Z} + \sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)} .$$

This double inequality can be deduced by the following reasoning.

The partial correlation between X and Z dependent on Y is:

$$\rho_{X,Z;Y} = \frac{\rho_{X,Z} - \rho_{X,Y}\rho_{Y,Z}}{\sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)}}$$

This partial correlation must be between -1 and 1:

$$-1 \le \frac{\rho_{X,Z} - \rho_{X,Y}\rho_{Y,Z}}{\sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)}} \le 1$$

Simplifying we get our required double inequality:

$$\rho_{X,Y}\rho_{Y,Z} - \sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)} \le \rho_{X,Z} \le \rho_{X,Y}\rho_{Y,Z} + \sqrt{(1-\rho_{X,Y}^2)(1-\rho_{Y,Z}^2)}$$

Appendix 4

RJEP(0.95) & R(0.95) – t Copula 5 df, Gaussian Copula and Independence

R(Z): t Copula 5 d.f.

z	95.0%

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		17.58%	21.85%	21.61%	20.89%	19.68%	20.16%	17.98%	17.58%	19.68%
Property	2			21.18%	20.03%	19.87%	19.38%	20.03%	20.36%	19.38%	19.71%
Interest Rate	3				21.20%	19.98%	19.45%	20.82%	20.67%	19.68%	20.89%
Credit Spread	4					19.06%	18.46%	22.20%	18.91%	19.96%	21.82%
Credit Default	5						19.13%	20.71%	20.16%	19.61%	19.76%
UW - Cat	6							18.87%	18.46%	19.35%	21.62%
UW Non-Cat	7								19.27%	19.51%	21.74%
Reserve	8									21.06%	21.38%
Expenses	9										19.48%
Operational	10										

RJEP(Z): t Copula 5 d.f.

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		0.87%	1.08%	1.07%	1.04%	0.98%	1.00%	0.89%	0.87%	0.98%
Property	2			1.04%	0.98%	0.97%	0.95%	0.98%	1.00%	0.95%	0.96%
Interest Rate	3				1.12%	1.06%	1.03%	1.10%	1.09%	1.04%	1.10%
Credit Spread	4					1.02%	0.99%	1.19%	1.01%	1.07%	1.17%
Credit Default	5						0.97%	1.05%	1.02%	1.00%	1.00%
UW - Cat	6							0.93%	0.91%	0.96%	1.07%
UW Non-Cat	7								0.97%	0.98%	1.09%
Reserve	8									1.04%	1.05%
Expenses	9										0.98%
Operational	10										

R(Z): Gaussian Copula

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		11.14%	11.22%	11.88%	12.12%	11.06%	10.89%	11.47%	11.55%	11.71%
Property	2			11.31%	11.63%	12.68%	11.55%	11.47%	13.17%	13.09%	11.71%
Interest Rate	3				11.80%	10.98%	10.82%	12.45%	10.90%	10.66%	12.45%
Credit Spread	4					11.22%	11.85%	11.46%	11.69%	12.49%	13.68%
Credit Default	5						13.19%	12.30%	11.33%	11.89%	11.65%
UW - Cat	6							10.67%	11.99%	12.41%	12.82%
UW Non-Cat	7								14.12%	11.46%	12.19%
Reserve	8									11.88%	12.52%
Expenses	9										11.01%
Operational	10										

RJEP(Z): Gaussian Copula

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		0.54%	0.55%	0.58%	0.59%	0.54%	0.53%	0.56%	0.56%	0.57%
Property	2			0.56%	0.58%	0.63%	0.57%	0.57%	0.65%	0.65%	0.58%
Interest Rate	3				0.58%	0.54%	0.53%	0.61%	0.54%	0.52%	0.61%
Credit Spread	4					0.56%	0.60%	0.58%	0.59%	0.63%	0.69%
Credit Default	5						0.65%	0.61%	0.56%	0.59%	0.58%
UW - Cat	6							0.52%	0.58%	0.60%	0.62%
UW Non-Cat	7								0.70%	0.57%	0.60%
Reserve	8									0.59%	0.62%
Expenses	9										0.56%
Operational	10										

R(Z): Independence	5.0%
RJEP(Z): Independence	0.25%

Kendall Tau – t Copula 5 d.f. and Gaussian Copula

Kendall Tau: t Copula 5 d.f.

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		17.62%	16.72%	17.74%	18.31%	16.29%	17.73%	16.72%	16.49%	17.50%
Property	2			15.38%	15.92%	16.39%	15.72%	17.20%	16.43%	16.29%	17.25%
Interest Rate	3				16.02%	16.67%	17.29%	17.66%	16.68%	16.71%	17.02%
Credit Spread	4					16.27%	17.17%	16.70%	15.94%	15.96%	16.02%
Credit Default	5						16.35%	17.22%	15.75%	15.68%	17.25%
UW - Cat	6							15.61%	15.16%	16.71%	17.90%
UW Non-Cat	7								15.56%	15.86%	17.63%
Reserve	8									16.73%	17.26%
Expenses	9										16.19%
Operational	10										

Kendall Tau: Gaussian Copula

	No.	1	2	3	4	5	6	7	8	9	10
Equity	1		16.25%	16.24%	15.64%	16.29%	15.97%	16.24%	15.53%	16.90%	15.58%
Property	2			16.13%	16.37%	15.35%	14.62%	16.00%	15.24%	17.53%	15.59%
Interest Rate	3				16.33%	15.90%	15.23%	14.50%	15.66%	16.78%	16.31%
Credit Spread	4					15.21%	15.30%	15.54%	15.16%	17.13%	16.62%
Credit Default	5						16.17%	13.69%	14.71%	14.85%	16.29%
UW - Cat	6							13.38%	16.68%	15.72%	16.85%
UW Non-Cat	7								14.64%	14.75%	15.54%
Reserve	8									16.97%	15.78%
Expenses	9										15.71%
Operational	10										

Appendix 6

ABC Insurance Company – Other distribution parameters

Risk Type	Distribution	Mu	Sigma	E(X)	SD(X)	CV(X)
Equity	Lognormal	7.4893	0.4724	2,000	1,000	50%
Property	Lognormal	7.4893	0.4724	2,000	1,000	50%
Interest Rate	Lognormal	7.4893	0.4724	2,000	1,000	50%
Credit Spread	Lognormal	7.4893	0.4724	2,000	1,000	50%
Credit Default	Lognormal	7.4893	0.4724	2,000	1,000	50%
UW - Cat	Lognormal	7.4893	0.4724	2,000	1,000	50%
UW Non-Cat	Lognormal	7.4893	0.4724	2,000	1,000	50%
Reserve	Lognormal	7.4893	0.4724	2,000	1,000	50%
Expenses	Lognormal	7.4893	0.4724	2,000	1,000	50%
Operational	Lognormal	7.4893	0.4724	2,000	1,000	50%

Risk Type	Distribution	Mu	Sigma	E(X)	SD(X)	CV(X)
Equity	Normal	2,000	500	2,000	500	25%
Property	Normal	2,000	500	2,000	500	25%
Interest Rate	Normal	2,000	500	2,000	500	25%
Credit Spread	Normal	2,000	500	2,000	500	25%
Credit Default	Normal	2,000	500	2,000	500	25%
UW - Cat	Normal	2,000	500	2,000	500	25%
UW Non-Cat	Normal	2,000	500	2,000	500	25%
Reserve	Normal	2,000	500	2,000	500	25%
Expenses	Normal	2,000	500	2,000	500	25%
Operational	Normal	2,000	500	2,000	500	25%

ABC Insurance Co. – Lognormal risks (CV = 25%, Corr = 10%, 25%, 50%)

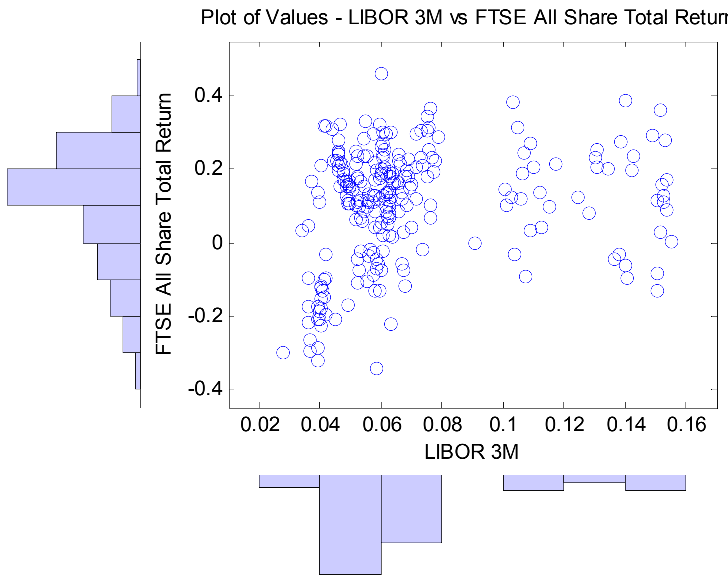
Economic Capital - 10% Correlation					% change cf Gaussian	
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-7.6%	-10.4%	-23.1%	-9.3%
90%	10	0.0%	-1.4%	-2.3%	-6.4%	1.9%
95%	20	0.0%	2.0%	2.1%	4.4%	6.2%
99%	100	0.0%	7.7%	13.0%	23.5%	13.1%
99.5%	200	0.0%	11.2%	18.5%	31.7%	15.5%
99.95%	2,000	0.0%	21.8%	32.6%	62.9%	25.5%

Economic Capital - 25% Correlation					% change cf Gaussian	
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-4.2%	-10.3%	-19.3%	-5.8%
90%	10	0.0%	-2.1%	-2.9%	-7.3%	2.0%
95%	20	0.0%	-0.4%	1.6%	-0.8%	5.2%
99%	100	0.0%	6.6%	10.2%	21.9%	10.6%
99.5%	200	0.0%	8.3%	19.5%	31.7%	12.7%
99.95%	2,000	0.0%	25.7%	34.7%	67.3%	21.5%

Economic Capital - 50% Correlation					% change cf Gaussian	
Percentile	**Return**	**Gaussian**	**t - 10 df**	**t - 5 df**	**t - 2 df**	**V CV**
75.0%	4	0.0%	-1.0%	-5.9%	-12.8%	-2.2%
90%	10	0.0%	-2.5%	-2.4%	-7.2%	1.1%
95%	20	0.0%	-1.7%	0.2%	-2.1%	2.6%
99%	100	0.0%	3.7%	7.6%	14.3%	6.6%
99.5%	200	0.0%	6.7%	10.6%	18.5%	8.7%
99.95%	2,000	0.0%	11.7%	11.7%	32.7%	7.9%

Canonical Maximum Likelihood – LIBOR 3M vs FTSE All Share Total Return

Scatter Plot of values for LIBOR 3M vs FTSE All Share Total Return



Scatter Plot of [0,1] values for LIBOR 3M vs FTSE All Share Total Return