Dependent Tails

Andrew D Smith, UK
August 2002

Abstract: Both investment and insurance markets can show phenomena known as tail dependency. Two lines of business can appear to operate independently most of the time, but in adversity all lines deteriorate together. Even if we can estimate well the loss ratio for each class of business, how can we capture the correlation structure? This note proves some background on techniques for introducing correlations between lines of business. It answers the key questions:

- How can a user introduce correlations between arbitrary distributions?
- Does the chosen methodology matter?

Workshop attendees will have the opportunity to run simulations live, examining the effect of different correlation structures on the distributions of losses, capital required, and allocation of capital between lines. The workshop will also consider ways of calibrating correlation structures to historic loss information.

Keywords: Correlation, Copula, Tail dependency, Dynamic Financial Analysis, Calibration.

Contact details:
Andrew D Smith, B&W Deloitte, Horizon House, 28 Upper High Street, Epsom, Surrey KT18 4RS, United Kingdom.
email: andrewdsmith8@bw-deloitte.com

Introduction – Two Examples

In this introduction, we look at two examples that illustrate the main points in the paper. The first example illustrates the problem of generating jointly correlated variables. The second example illustrates how the choice of correlation structure can have an impact on calculated risk measures – we give an example where the economic capital required to support a business becomes 5 times larger when a correct structure is used.

Introductory Example I: Correlated Exponential Distributions

An insurance company considers total claims on two lines of business. In each case, the line of business loss is assumed to have an exponential distribution with mean €1000. The two distributions are also required to have a 50% correlation.

And this is the tricky bit. How can we generate exponentially distributed variables with a chosen correlation? Are there general algorithms, which work for a wide range of distributions and correlations?

After much thought, we remember that an exponential distribution with mean 2 is the same as a χ_2^2 distribution, that is, the sum of two squared normal distributions. Our two loss distributions L_1 and L_2 could therefore be generated as follows:

L_1 = €500 * [$Z_1^2 + Z_3^2$]
L_2 = €500 * [$Z_2^2 + Z_3^2$]

where Z_1, Z_2 and Z_3 are independent standard normal variates.

This solution is a special case, limited only to exponential distributions and to correlations of 50%. We made some arbitrary choices – would other ways of putting the variables together give different results for the combined claims across the two lines?

What we need is a generic approach, which will work for arbitrary distributions and correlation structures. It is this, which we set out in the remainder of the paper.

Introductory Example II: Uncorrelated Normal Profits

Let us consider a business wishing to evaluate its need for capital. The business has two operating units. Each operating unit generates a profit (positive or negative) with respective means of €100 and €120.

The profits are believed to enjoy normal distributions, with zero correlation and standard deviations of €60 and €80 respectively. The business is interested in evaluating how much capital it needs, in order to have a 99% chance of staying solvent.

On the face of it, this is a simple value at risk calculation. If the two operating unit profits are independent normal with zero correlation, then by the usual Pythagoras result, the sum will also be normal with mean €220 and standard deviation €100. According to the normal distribution, the 1%-ile level is 2.33 standard deviations below the mean. One year in 100, we might find a net profit of €220 - €233, that is -€13. Therefore, to keep the probability of failure down to 1% the business needs to start the year with an additional €13 of capital.

The business decides to back-test its capital requirement against its own data. After extensive experience and back testing, the insurer is able to validate the two normal distributions, and also the zero correlation. However, with €13 of capital, the business finds that historical failure rate would have been 2.4%, not 1% as predicted. In order to reduce the probability to 1% they would need to find €67 of capital – more than 5 times larger than originally budgeted.

We can explain the discrepancy in terms of how the normal distributions are put together. On further investigation, the two lines are indeed uncorrelated, but are not independent. The quants tried to improve on their original model, and this was their alternative, which better fitted the historic data.

line 1 profit = €100 + €60 * Z
line 2 profit = €120 ± €80 * Z

where Z is a standard normal variate, and the choice of ± is made independently of Z. This plainly gives two uncorrelated normal distributions with the stated means and standard deviations. However, the two lines of business are certainly not independent, because the same value of Z applies to each line. For example, if we observe that line 1 has a profit of €130, then we can see that Z= ½ and so the line 2 profit must be either €80 or €160.

When line 1 has an extreme profit, the line 2 will also show an extreme profit. The combined distribution is a mixture of

N(€220, €140^2) if "+" applied
N(€220, €20^2) if "-" applied

It is the first of these distributions, with the €140 standard deviation, which gave rise to the worst loss events, and which accounts for most of the capital requirement.

This example has illustrated how correlation may be an inadequate measure of the relationship between two variables. There are a number of questions we have left unanswered. One of these is whether the independent normal model gives the most optimistic capital requirements, or whether it lies in the middle of a plausible range of dependency structures.

Ways of Introducing Correlations

There are many direct ways of introducing correlations into a model of an enterprise. It is preferable to explain correlations in terms of business drivers and known causal mechanisms. For example, the correlation between the price of a bond and market interest rates, arises because the bond price is the present value of its future cash flows. It is usually better to model this known link directly, rather than trying to piece together a historic correlation structure.

In an insurance context, many mechanisms can give rise to correlations between different lines of business. Three of the more important effects are:

- Perils, which can generate losses on many policies at once, covering more than one class of business. Examples include natural hazards such as storms, man-made events such as the World Trade Centre, and wide ranging legal or political changes.
- Cycles in the market price of insurance, reflecting the availability of capital in the industry or in the competitive environment.
- Macro economic effects such as movements in currency exchange rates or changes in inflation trends, which can affect the size of claims when converted to a single accounting currency.

It is wise to seek first the known business correlations. This frequently accounts for most if not all the correlations seen in historic loss ratios. Where all the known effects have been stripped out, and a residual correlation remains, we need to resort to statistical correlation methods to calibrate a model.

A Class of Algorithms for Correlated Variables

After having exhausted those correlations we can explain, we will typically have a data set adjusted for known effects. If this adjusted set still contains significant correlations, we must resort to data-based modelling tools.

In this section, we outline an algorithm for generating a correlated pair of variables. For illustrative purposes, we consider generating 1000 simulations from a bivariate distribution. We specify two required marginal distributions and a required correlation ρ.

The algorithm:

- Simulate X1(n) independently from marginal distribution 1, for n = 1 to 1000
- Simulate X2(n) independently from marginal distribution 2, for n = 1 to 1000
- Sort X1 into increasing order, so X1(1) ≤ X1(2) … ≤ X1(1000)
- Likewise, sort X2 likewise into increasing order
- Generate pairs {Z1(n), Z2(n)} for n=1 to 1000 independently from a “suitable” bivariate distribution with correlation ρ. This is sometimes (but not strictly correctly) called the “copula distribution”.
- Fill an array R1(1) … R1(1000) with the integers 1 through 1000, such that R(n)=n for each n
- Likewise, create an array R2(1 to 1000) with R2(n) = n
- Sort the array Z1 while simultaneously rearranging the array R1 to follow the reordering of Z1. After this sort, the array Z1 will be in increasing order, but the elements of R1 will point to the original sequence of Z1 prior to sorting. Thus, for example, R1(1) will contain the original simulation number, which produced the smallest value of Z1. The values R1(1) through R1(1000) are now some permutation of the integers from 1 to 1000.
- Likewise, sort Z2 while simultaneously rearranging R2
- Define an array Y1 by Y1(R1(n)) = X1(n) for each n from 1 to 1000
- Likewise, define by Y2(R2(n)) = X2(n)
- The pairs {Y1(n), Y2(n) } now have the required marginal distributions, and are correlated as required.

Some observations on the Algorithm

Let us first verify whether the algorithm works.

The output array Y1 is a permutation of the generated array X1. Y1 therefore has the same marginal distribution as X1, as required. Likewise Y2 has the required marginal distribution.

Furthermore, we can see that the correlation of {Y1, Y2} reflects the correlation of the copula distribution {Z1, Z2}. To see how this works, lets look at some extreme examples.

If copula variables Z1 and Z2 are 100% correlated, then the largest observation of Z1 happens in the same simulation as the largest Z2. Suppose this happens in original simulation 567. Before sorting, Z1(567) and Z2(567) are the largest observations of Z1 and Z2 respectively. And of course, before sorting, R1(567)=R2(567)=567. Now we do the sorting. After sorting Z1 with corresponding changes in R1, Z1(1000) becomes the largest observation of Z1, with R1(1000) = 567 moving with it. By the same logic, we also see R2(1000) = 567. Then finally, we can see that Y1(567) = Y1(R1(1000)) = Z1(1000) and Y2(567) = X2(1000). This means that the largest observation of Y1 is still associated with the largest Y2. The same argument applies for the second largest Y1, right down to the smallest Y1. We have obtained a set of simulations where Y1 is an increasing function of Y2.

In the same way, if the copula variables Z1 and Z2 are -100% correlated then we will see that Y2 is a decreasing function of Y1. If Z1 and Z2 are independent, then so are Y1 and Y2.

We can also see a limitation of the method. Even if the copula variables are 100% correlated, the final variables Y1 and Y2 might not be. For example, suppose Y1 has a normal distribution and Y2 has a gamma distribution. Our algorithm could express Y2 as an increasing function of Y1. But to get 100% correlation between Y1 and Y2, we would need a linear relationship. This we cannot have – if Y1 were normal and Y2 were a linear function of Y2, then Y2 would have to be normal, instead of gamma.

This highlights the most important limitation of the technique so far. We have a tool, which generates correlated random variables with arbitrary distributions. What we need is a way of predicting the correlations that will come out of the algorithm. The trouble is that that the out-coming correlations are not the same as the input correlations. We will revisit this issue when we look at alternative measures of association.

Suitable Multivariate Copula Distributions

There are a number of standard copula distributions in the literature, including those by Cook and Johnson (1981) and many more references in Embrechts, Lindskog and McNeil (2001) and in Wang (2001). Genest and McKay (1986) describe a family of copulas, which they name as “Archimedean”. Many of these families possess a single (scalar) parameter, which determines the degree of dependence between all pairs of components. For a two dimensional distribution this is fine- the parameter is a measure of correlation or something like it. For multivariate distributions, we have more of a problem, because the same parameter must apply to every pair. This imposes symmetry, which may not be desirable. For example, an insurer might want to introduce a high correlation between motor and household loss ratios, with a smaller correlation between these and employers' liability.

Cambanis, Huang and Simons (1981) investigate a broader family of copulas, the so-called elliptically contoured distributions. A multivariate elliptically contoured distribution is characterised by a density function of the form:

$$f(x) = \frac{1}{\sqrt{|\Sigma|}} \phi\left[(x-\mu)^T \Sigma^{-1} (x-\mu)\right]$$

The function φ must satisfy certain conditions, under which the distribution has mean μ and variance-covariance matrix Σ.

Special cases of elliptically contoured distributions include the normal distribution and multivariate *t* distribution. The advantage of elliptically contoured distributions is that we can specify a full variance-covariance matrix. We can therefore accommodate situations where users wish to specify pair wise correlations between several lines of business.

Elliptically contoured distributions have a number of elegant properties, and are generally easy to simulate from.

A General Algorithm

This section provides a construction for creating a random vector *X*, whose components have zero mean, unit standard deviation and a given correlation matrix ρ. To generate a *k*-dimensional vector *X*, we need the following parameters:

(i) a scalar random variable H distributed on the real line, with mean 1 and standard deviation σ_H
(ii) a symmetric positive definite k-by-k correlation matrix ρ whose diagonal elements are 1.
(iii) a non-centrality k-vector u with $u^T\rho^{-1}u \le 1$. This necessarily implies that $|u_i| \le 1$ for each i.

Our algorithm is then as follows:
(iv) generate H and a multivariate normal Z independent of H with mean 0 and variance-covariance matrix $\rho\text{-}uu^T$
(v) define $X = (H\text{-}1)\ \sigma_H^{-1}u + H^{1/2}.Z$

It is straightforward to verify that this algorithm does indeed produce a vector X with zero mean, unit standard deviation and correlation matrix ρ. If u=0 we have an elliptically contoured distribution.

Measures of Association

To specify a multivariate distribution, we need a way of capturing how related a pair of distributions are.

The classical method of measuring association is the linear (pearson) correlation, defined using variances and covariances as:

$$\text{pearson correlation} = \frac{\mathbf{Cov}(X,Y)}{\sqrt{\mathbf{Var}(X)\mathbf{Var}(Y)}}$$

We have already seen than that the copula algorithm does not preserve pearson correlations. Models would be easier to construct if we used association measures which monotone transformations preserve.

One such measure is the spearman rank correlation, sometimes called spearman's rho. This is defined in a similar way to the pearson correlation, except that each observation of X or Y in a sample is first replaced by its rank. For example, we would replace the smallest observation of X by 1, then next smallest by 2 and so on, applying this also to Y. The resulting correlation measure is invariant under monotone transformations of X and Y. It is only the order of the observations that matters.

It is useful to consider a transformation of spearman's correlation as follows. Let F_X denote the marginal cumulative distribution function of X, and F_Y that of Y. Then both $F_X(X)$ and $F_Y(Y)$ have uniform distributions, with mean 1/2 and variance 1/12. The pearson correlation of $F_X(X)$ and $F_Y(Y)$ is precisely the spearman correlation of X and Y.

Now let (X_1, Y_1), (X_2, Y_2) and (X_3, Y_3) be three independent observations from the bivariate distribution (X, Y). The spearman correlation of X and Y is:

$$\begin{aligned}
\text{spearman}(X, Y) &= 12\mathbf{E}\{F_X(X_1)F_Y(Y_1)\} - 3 \\
&= 12\mathbf{E}\{\mathbf{P}[X_2 \le X_1 \text{and} Y_3 \le Y_1 | X_1, X_2]\} - 3 \\
&= 12\mathbf{P}[X_2 \le X_1 \text{and} Y_3 \le Y_1] - 3
\end{aligned}$$

We can deduce similar expressions for the other quadrants of possible combinations of X_1, X_2, Y_1 and Y_3, making use of such identities as $\mathbf{P}\{X_2 < X_1\} = \frac{1}{2}$. Denoting the spearman correlation by ρ, we can deduce that:

$$\mathbf{P}\{X_2 \le X_1, Y_3 \le Y_1\} = \frac{1}{4} + \frac{\rho}{12}$$

$$\mathbf{P}\{X_2 \le X_1, Y_3 > Y_1\} = \frac{1}{4} - \frac{\rho}{12}$$

$$\mathbf{P}\{X_2 > X_1, Y_3 \le Y_1\} = \frac{1}{4} - \frac{\rho}{12}$$

$$\mathbf{P}\{X_2 > X_1, Y_3 > Y_1\} = \frac{1}{4} + \frac{\rho}{12}$$

An alternative measure of association is kendall’s tau. This is defined by looking at the ordering possibilities for the two pairs (X_1, Y_1), (X_2, Y_2). Denoting kendalls tau by τ, the corresponding definitions (based on Kendall and Stuart, 1979) are:

$$\mathbf{P}\{X_2 \le X_1, Y_2 \le Y_1\} = \frac{1}{4} + \frac{\tau}{4}$$

$$\mathbf{P}\{X_2 \le X_1, Y_2 > Y_1\} = \frac{1}{4} - \frac{\tau}{4}$$

$$\mathbf{P}\{X_2 > X_1, Y_2 \le Y_1\} = \frac{1}{4} - \frac{\tau}{4}$$

$$\mathbf{P}\{X_2 > X_1, Y_2 > Y_1\} = \frac{1}{4} + \frac{\tau}{4}$$

Spearman’s rho and kendall’s tau are *concordance* measures. This means they have a number of useful properties as follows (selected from a longer list in Scarsini, 1984, defining concordance measures):

- the concordance lies between –1 and 1
- the concordance of X with itself is +1, and with $-X$ is –1

- the concordance of X with Y is the same as the concordance of Y with X
- independent variables have zero concordance
- the concordance of X with $-Y$ is minus the concordance of X with Y
- the concordance is unaffected by monotone transformations of X and Y

In addition, with some ingenuity we can identify rho and tau analytically for multivariate normal distributions. The results for kendall's tau also extend to elliptically contoured distributions.

Calibration Issues

We now consider some alternative ways of calibrating multivariate dependency structures. We consider first the problem assuming class of copula functions is given, and then move on to examine how a copula class might b chosen for calibration.

Let us suppose first then than a class of copula functions has been given. For each pair of variables, we wish to specify a measure of association, and then back-solve to give a distribution with that property. The contenders for the measure are pearson correlation, spearman's rho and kendall's tau. Pearson correlation is the best known, but is also the most awkward as it is not preserved by the monotone transformations. Next in line is spearman's rho – moderately well known but also invariant under monotone transformations. This is still difficult to calculate for some copula distributions, so a simulation based trial and error process may be required. Kendall's tau is least familiar but also the most tractable, in that it is known analytically for a wide class of copula distributions, including all elliptically contoured distributions.

In a multivariate context, we can face the problem that the chosen measures of association between many pairs of variables cannot be assembled into a single copula of the chosen family. This is similar to the problem with linear correlations, estimating correlation matrices that are not positive definite. However, in the case of linear correlations, we are at least assured that an attempt to estimate correlations from a single data sample (as opposed to correlations taken from series of different lengths or otherwise different sources) will result in a consistent matrix. Alas, this is not guaranteed in the case of spearman's rho or kendall's tau – experience shows that infeasible matrices crop up often with real data, especially with high dimensional problems. I can offer no magic solution to this, other than ad-hoc adjustments until the matrix is feasible.. The range of feasible matrices depends on the form of copula chosen – and it may be that some forms are more forgiving than others of extreme input matrices. Further work is needed in this area.

We now move on to the tricky aspect – choosing a copula function. In practice, the normal copula serves as a convenient default. So we need to look for reasons not to choose the normal copula.

The most fashionable aspect of non-normality is the measure defined by Joe (1997) as *tail dependence*. Tail dependence is defined in terms of limiting behaviour of probability functions close to 1. The normal distribution has zero tail dependence, even if the correlation is positive (but less than 1). If we could demonstrate positive tail dependence for a data set, this could be a good reason to reject normal copula structures. Unfortunately, such limits are difficult to estimate robustly from finite data sets. This is because of the limiting operations involved.

Instead, I propose examining a set of measures I call *quadrant correlations*. Unless someone else has defined these independently, I believe this is the first time quadrant correlations have been described in the literature.

For simplicity, let us assume that *X* and *Y* have been scaled to have zero mean and unit standard deviation. Then the pearson correlation is defined by

$$\rho_{XY} = \mathbf{E}[XY]$$

I define the four quadrant correlations as:

$$\rho_{XY}^{++} = \mathbf{E}[\max\{X,0\}\max\{Y,0\}]$$
$$\rho_{XY}^{+-} = \mathbf{E}[\max\{X,0\}\min\{Y,0\}]$$
$$\rho_{XY}^{-+} = \mathbf{E}[\min\{X,0\}\max\{Y,0\}]$$
$$\rho_{XY}^{--} = \mathbf{E}[\min\{X,0\}\min\{Y,0\}]$$

These plainly add up to the pearson correlation. Both ρ^{++} and ρ^{--} are positive, the other two quadrant correlations being negative. For a given distribution family (for example, normal) these will all be determined by the pearson correlation. A test for a normal correlation structure is therefore whether the quadrant correlations bear the correct relation to the overall correlation, consistent with the normal distribution.

For problems involving copulas, it is more convenient to deal with spearman rank correlations. The same concepts of quadrant correlations carry over. We can define the first quadrant spearman rank correlation by the expectation:

$$\rho_{XY}^{++} = 12\mathbf{E}\left[\max\left\{F_X(X) - \tfrac{1}{2}, 0\right\}, \max\left\{F_Y(Y) - \tfrac{1}{2}, 0\right\}\right]$$

Other quadrant spearman rank correlations are similarly defined. The constants of 1/2 and 1/12, which appear here, are required in order to transform the cumulative distribution functions to mean zero and unit standard deviation. Note that in addition to testing normality, we can test for elliptically contoured distributions, for which correlations in opposite quadrants are equal.

As before, we can express quadrant spearman rank correlations in terms of probabilities of independent samples. Specifically, suppose that X_1 and X_2 are independent and identically distributed with cumulative distribution function F. Suppose also that each X has median zero, so that $F(0) = 1/2$. Then we can see that

$$\max\{F(X_1) - ½ , 0\} = \mathbf{Prob}\{ 0 \le X_2 \le X_1 \mid X_1\}$$
$$\min\{F(X_1) - ½ , 0\} = -\mathbf{Prob}\{ 0 \le X_1 < X_2 \mid X_1\}$$

Taking this and similar results, we can now identify the four quadrant spearman rank correlations:

$$\begin{aligned}
\rho_{XY}^{++} &= 12\mathbf{P}\{0 \le X_2 \le X_1, 0 \le Y_3 \le Y_1\} \\
\rho_{XY}^{+-} &= -12\mathbf{P}\{0 \le X_2 \le X_1, 0 \le Y_1 < Y_3\} \\
\rho_{XY}^{-+} &= -12\mathbf{P}\{0 \le X_1 < X_2, 0 \le Y_3 \le Y_1\} \\
\rho_{XY}^{++} &= 12\mathbf{P}\{0 \le X_1 < X_2, 0 \le Y_1 < Y_3\}
\end{aligned}$$

These spearman correlations plainly have parallels in the world of kendall's tau. We can therefore define the quadrant kendall's taus as:

$$\begin{aligned}
\tau_{XY}^{++} &= 4\mathbf{P}\{0 \le X_2 \le X_1, 0 \le Y_2 \le Y_1\} \\
\tau_{XY}^{+-} &= -4\mathbf{P}\{0 \le X_2 \le X_1, 0 \le Y_1 < Y_2\} \\
\tau_{XY}^{-+} &= -4\mathbf{P}\{0 \le X_1 < X_2, 0 \le Y_2 \le Y_1\} \\
\tau_{XY}^{++} &= 4\mathbf{P}\{0 \le X_1 < X_2, 0 \le Y_1 < Y_2\}
\end{aligned}$$

For multivariate normal distributions, we can evaluate both the quadrant rank correlations and the quadrant kendalls taus analytically using trigonometric functions.

Conclusions

There are several steps in determining appropriate dependency structures for correlated random variables. We have found the following procedure to be workable in practice:

- Investigate known economic or business causal links, which could account for observed dependencies
- Adjust the data for known causal links, and calibrate marginal distributions for the adjusted data
- Examine the remaining residuals for evidence of dependencies; if there is little evidence then assume independence as a null hypothesis
- If there is evidence of dependency, use quadrant rank correlations to examine evidence for non-normal copulas
- Use normal copulas in the absence of evidence to the contrary
- Otherwise, examine quadrant rank correlations for evidence of non-ellipticity
- Calibrate an appropriate copula function to replicate observed quadrant rank correlations
- Some overall adjustments may be required to ensure the desired correlation matrix is consistent.
- Simulate dependency structures using the algorithm outlined in this paper.

Acknowledgements

My interest in this subject, and some of the ideas in this paper, arose following a GIRO workshop on a similar subject, led by Dmitris Papachristou and Andrzej Czernuszewicz. Unfortunately, I have been unable to find any paper from this session. I am grateful to Andreas Tsanakas, Peter Tavner and Karen Seidl for useful discussions. All errors and omission are entirely the author's.

Bibliography

Cambanis, S, Huang, S and Simons, G (1981) *On the theory of elliptically contoured distributions*, Journal of multivariate analysis 11, 368-85.

Cook, R and Johnson, M (1981) *A family of distributions for modelling non-elliptically symmetric multivariate data*, Journal of the Royal Statistical Society B43, 210-218.

Embrechts, P; Lindskog, F and McNeil A (2001). *Modelling Dependence with Copulas and Applications to Risk Management*. ETH working paper.

Genest, C and Mackay, J (1986). *The Joy of Copulas – Bivariate distributions with uniform marginals,* The American Statistician 40 #4, 280-283.

Joe, H (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.

Kendall, M and Stuart, A (1979) *Handbook of Statistics*. Griffin and Company, London.

Nelsen R(1999) *An Introduction to Copulas*. Springer, New York.

Scarsini, M (1984) *On Measures of Concordance*, Stochastica 8:201-218

Wang, S. (2001) *Aggregation of Correlated Risk Portfolios: Models and Algorithms*. Casualty Actuarial Society