Institute of Actuaries of Australia

LINEAR CORRELATION AS A MEASURE OF DEPENDENCY

Prepared by
Stephen Britt
Albert Napoli

Presented to the Institute of Actuaries of Australia
XVth General Insurance Seminar 16-19 October 2005

This paper has been prepared for the Institute of Actuaries of Australia's (Institute) XVth General Insurance Seminar 2005. The Institute Council wishes it to be understood that opinions put forward herein are not necessarily those of the Institute and the Council is not responsible for those opinions.

© 2005 Stephen Britt, Albert Napoli

The Institute will ensure that all reproductions of the paper acknowledge the Author/s as the author/s, and include the above copyright statement:

The Institute of Actuaries of Australia
Level 7 Challis House 4 Martin Place
Sydney NSW Australia 2000
Telephone: +61 2 9233 3466 Facsimile: +61 2 9233 3446
Email: actuaries@actuaries.asn.au **Website:** www.actuaries.asn.au

LINEAR CORRELATION AS A MEASURE OF DEPENDENCY

Stephen Britt
Albert Napoli

Abstract

Pearson (or linear) correlation is one of the most commonly used measures of dependency. However, linear correlation has a number of drawbacks such as failing to show perfect dependence if the relationship between two variables is not linear.

In this paper, the authors provide an outline of the desirable properties of a measure of dependency and an intuitive interpretation of Pearson correlation which highlights the shortcomings of the measure as a dependency measure. Simulation is used to illustrate these shortcomings and examine some alternatives which satisfy the desirable properties of a measure of dependency.

Keywords: dependency; correlation; copula

LINEAR CORRELATION AS A MEASURE OF DEPENDENCY

Stephen Britt
Albert Napoli

1 Introduction

Australian general insurance actuaries have for many years had a requirement when reserving to report the best estimate of the claims expected to arise from insurance premiums received. Prudential Standard GPS 210, issued in November 2001, added to this requirement by specifying that the insurance liability should be decomposed into a central estimate plus a quantifiable risk margin such that there is a 75 percent probability that the reserves will be sufficient to meet claims. Actuaries working for listed insurers have the added responsibility of quantifying the size of the prudential margin they actually hold by identifying the probability of sufficiency of their total portfolio of insurance liabilities.

The approach almost universally suggested/used is quite simple (!):

- determine the distribution of reserves for each line of business on a stand-alone basis; and
- 'combine' these using a matrix of (Pearson) correlation coefficients.

The former is well understood, at least in theory, and there is a history of the literature going back at least 20 years to illustrate some statistical techniques.[1] The latter appears to be less well understood despite occasional attempts[2] to discuss the shortcomings of Pearson correlations for this purpose.

The latter relies on Pearson correlation coefficients to estimate the dependency structure of the reserve estimates. There are a number of issues which arise from this approach:

- the approximations actually used are not clear to management or (we suspect) to all the actuaries using the techniques;
- Pearson correlation is often simply used as the measure of dependency structure, rather than as a choice of a number of different choices;
- given the size of the data sets available, estimates of correlation coefficient have significant statistical error. This may not be too problematic, as correlations are usually not derived from the data but rather from general reasoning (not necessarily a bad thing);
- for the sorts of distributions that actuaries usually assume in setting reserves (ie those that do not allow negative reserves):
    - Perfect dependence does not mean a correlation of 1; and
    - Perfect negative dependence does not mean a correlation of -1.

The authors would love to set out a theoretically correct, easy, comprehensive, intuitive and robust way to estimate the diversification benefits that arise from estimating the insurance reserves of portfolios of insurance liabilities. Unfortunately we have not come across such a method and have every reason to believe that, like the Mirror of Erised in Harry Potter, such a technique lies in the realms of wishful thinking.

What the authors hope to achieve with this paper is far more modest:

- Explain in an intuitive manner exactly what Pearson correlation is, and what its shortcomings are when applied in traditional general insurance applications;
- Introduce some other dependency measures, so that actuaries lucky enough to have access to enough data can estimate robust dependency structures;

[1] See for example Taylor (1988), Ashe (1986), Bateup & Reed (2001), and Collings & White (2001).
[2] Priest (2003) is a recent Australian example.

- Give some pointers to various techniques (structural models and copulas) that may be used to model dependency in insurance contexts.

2 Examining Pearson Correlation

Pearson correlation (or linear correlation as it is also called) is the most commonly (mis-) used dependency measure. It is uncertain whether it is mis-used more commonly than it is correctly used.

This section is predominantly graphical, and based on simulation rather than algebraic manipulation. In this way we sacrifice rigor for simplicity of presentation.

2.1 What is Pearson (Linear) Correlation?

Although not defined in this way linear correlation is perhaps best thought of as a measure of the degree to which the two data sets can be fit by a straight line via linear regression.

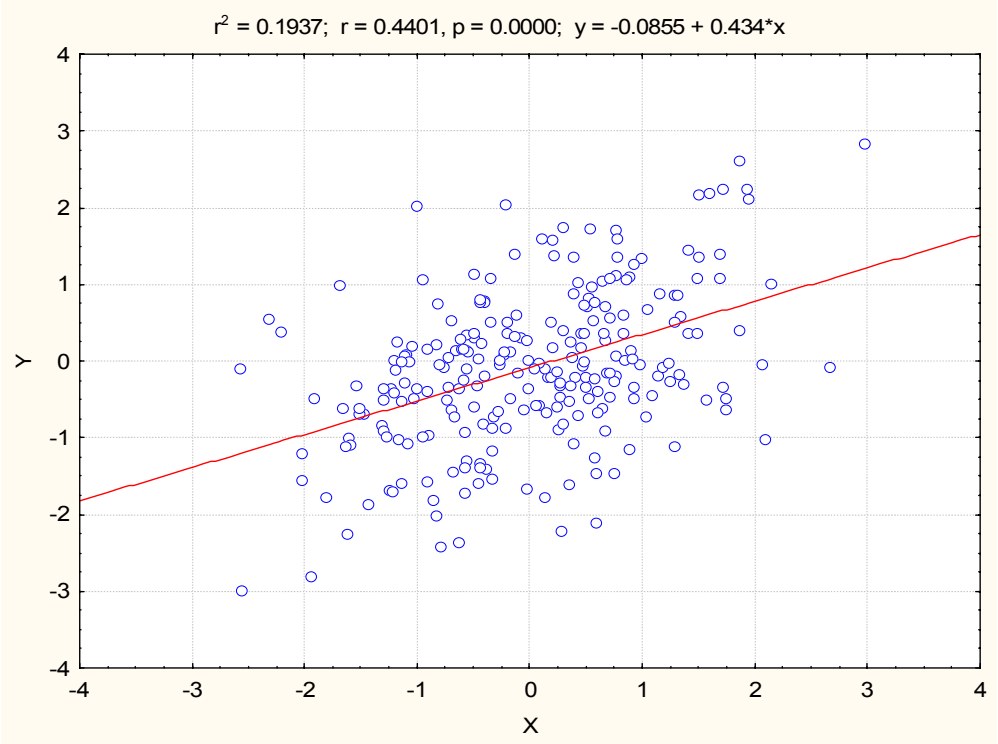

Exhibit 1: Normally Distributed Variables, Correlation = 0.44

Exhibit 1 shows 250 pairs of two normally distributed variables with correlation coefficient 0.44.[3] See how the observations cluster around the straight line, with some reasonable variation about it.

Exhibit 2 shows the same two variables with a correlation of 0.89 rather than 0.44. The observations lie more closely on the line.

[3] We tried for 0.5, but correlation is always generated with significant error even with a sample size of 250. DFA users should note, and see the appendix for a discussion.

The statistical measure of the degree to which the regression line fits the data is the r^2, which is the proportion of the total variance of the data which can be explained by a straight line. r^2 is related to Pearson correlation – it is the square of the correlation (times the sign of the slope of the regression coefficient).

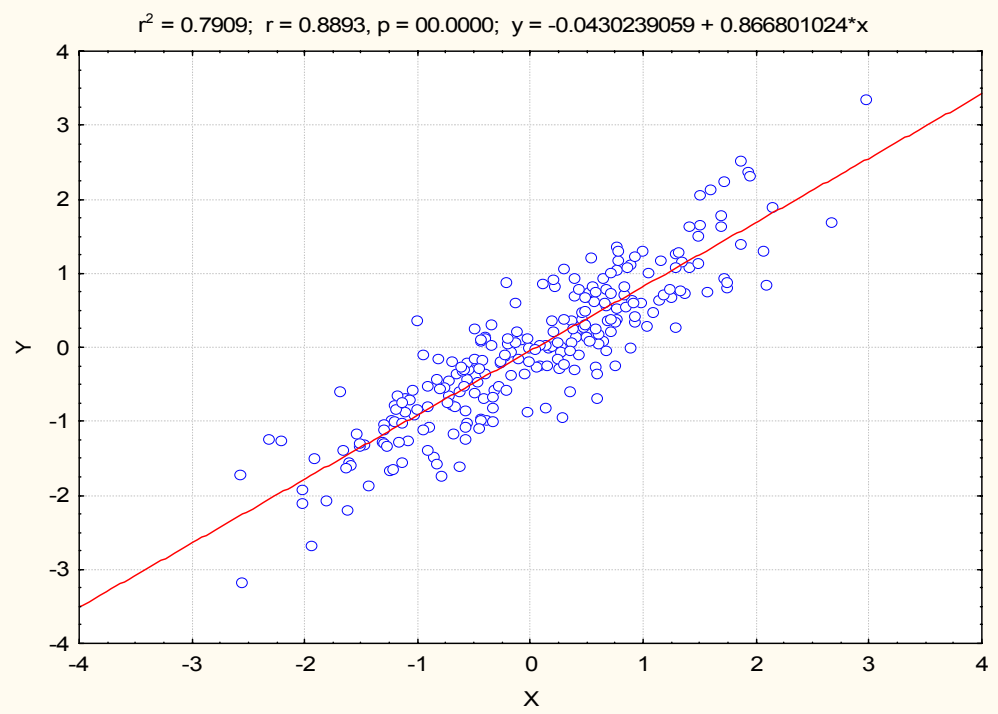

Exhibit 2: Normally Distributed Variables, Correlation = 0.89

Having defined geometrically what we mean by (Pearson or linear) correlation, let's misapply it to the following. Set X~N(0, 1) and Y such that log(Y) = X. Y is distributed as log-normal with underlying parameters mu = zero, and sigma = 1. Furthermore, if we know $X = X_1$ we know for certain what Y is. $Y=\exp(X_1)$. This is sometimes referred to as perfectly 'adapted to the same process' in finance. Exhibit 3 below shows the relevant graph.

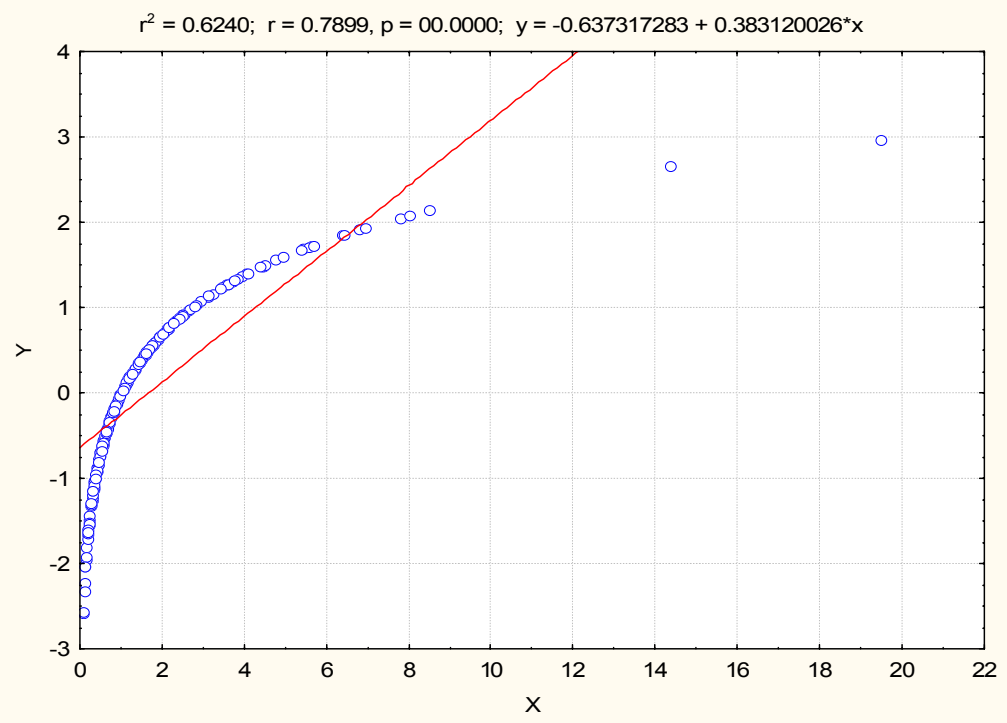

Exhibit 3: Normally Distributed and its Exponent

The correlation is not 1, as you might expect if you considered correlation as a measure of dependency, but is 0.79. Thinking of it in terms of closeness to a linear regression line though, the result is perfectly sensible (although perhaps the use of linear correlation in this case isn't!)

This chart shows another interesting fact about linear correlation – it is not possible to get a correlation above 0.79 if your data are drawn from these marginal distributions – linear correlation is not bounded on [-1, 1] in this case!

2.1.1 Correlation of Lognormally Distributed Pairs of Data

Perhaps the most relevant aspect for us is the correlation between pairs of variables when both are lognormally distributed. This is a very common assumption for the distribution of outstanding claims, and the broad conclusions also apply to similar distributions (Weibull and Gamma, for example).

Before proceeding, recall that with normally distributed random variables:

- μ is the location parameter (the mean); and
- σ is the scale parameter (standard deviation).

With log-normally distributed random variables however:

- μ is the scale parameter (the size of the distribution); and
- σ is the shape parameter. Increasing σ not only increases the variance of the distribution but also its skewness.

Standardising on $\mu = 0$, different σ means different shapes.

This is illustrated in Exhibit 4 which shows the distribution of two lognormally distributed variables with $\mu = 0$, and $\sigma = 0.25$ and 0.5.

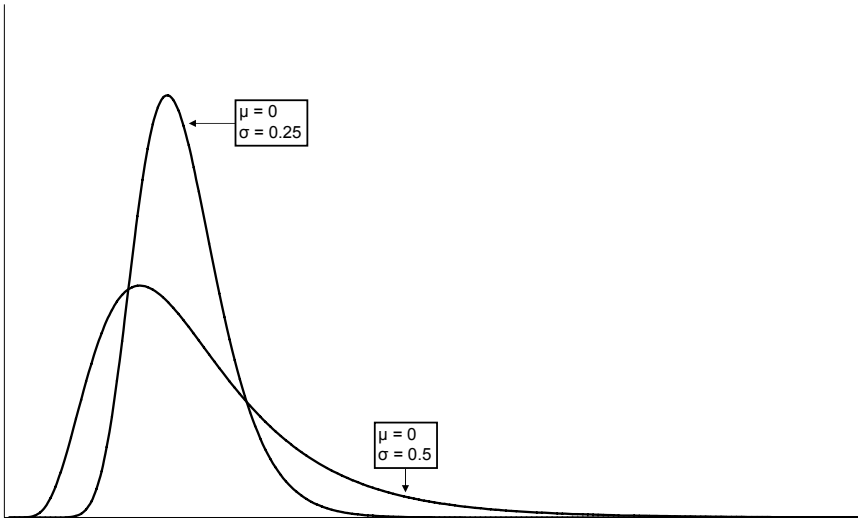

Exhibit 4 : Two Lognormal Random Variables, Different Shape Parameters

There are two different cases (in each case the underlying distribution of both variables is log-normal):

- Two different underlying variables, with the same shape parameter; and
- Two different underlying variables, with different shape parameters.

2.1.2 Different Variables, Same Shape

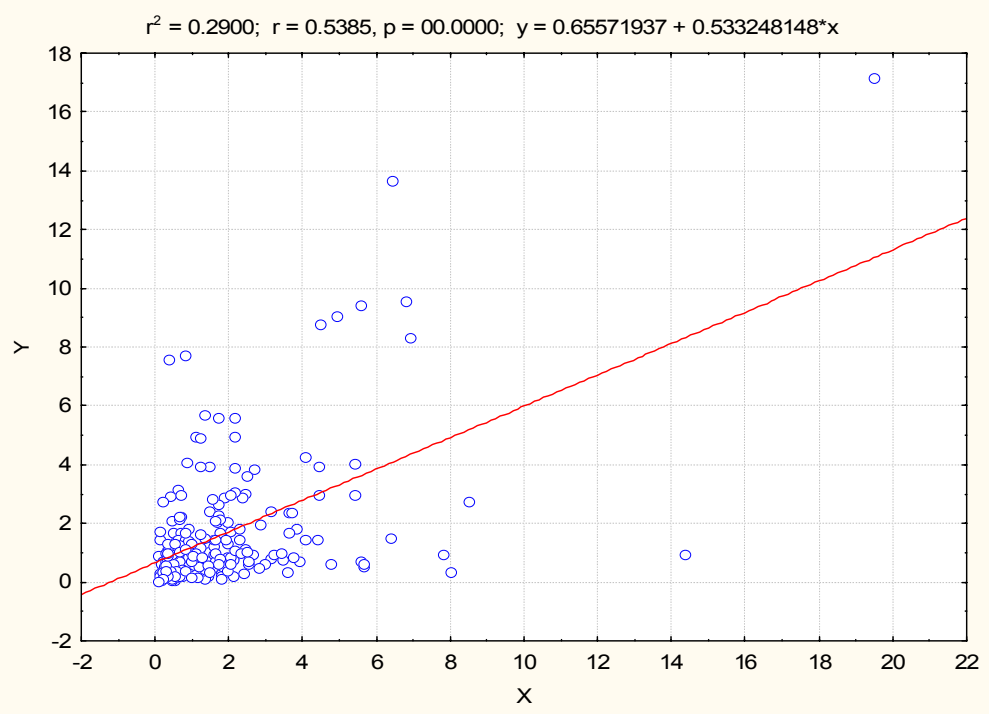

Exhibit 5: Two Lognormally Distributed Random Variables, Same Shape

In Exhibit 5 the two variables have the same shape parameter (1). The correlation between the underlying normals is still 0.44.

The observed correlation is 0.54, compared with the correlation of the underlying normal distributions which is 0.44 (see Exhibit 1). It is not hard to imagine that the observation on the top right corner of the chart may have significant effect on the regression line, unduly influencing the correlation estimate[4].

2.1.3 Different Variables, Different Shape

Finally, we look at the most common case – two variables with different underlying normally distributed random variables (with correlation 0.44) and different shapes (sigma = 1 and 2 respectively). Exhibit 6 below shows the case.

[4] Statistically speaking, the observation has significant leverage in the linear regression.

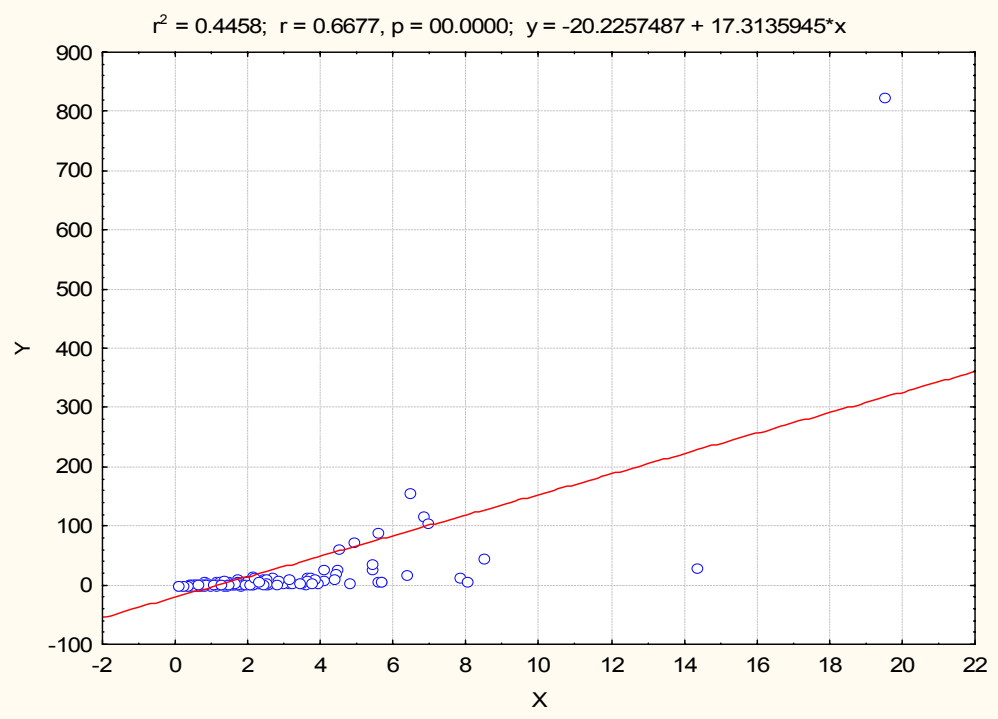

Exhibit 6 : Lognormally Distributed Variables with Different Shapes

Here correlation is 0.67, rather than 0.44.

3 Dependency Concepts

Thus far the paper has used a graphical approach to presenting dependency concepts, and illustrating the shortcoming of linear correlation. This section introduces some vocabulary which we then use to discuss some alternative dependency measures.

Section 4 of Embrechts, McNeil and Straumann, sets out these concepts in a more rigorous mathematical fashion using copula theory.

3.1 Perfect Dependence

The concept of independence is intuitively clear to most actuaries. Two variables X and Y are independent if knowing one value (say, X = 5) gives no information at all regarding Y.

Formally,

$$P(Y \le y | X = x) = P(Y \le y)$$

Equation 1

We would like to define perfect dependence as 'the opposite' of independence' and seek a statistical definition that fits this view. Intuitively, the two variables defined in Exhibit 3 exhibit perfect dependence, in that knowing one variable is all that is needed to determine the other. Mathematically we could probably recognise that all we have done to generate perfectly dependant variables is to make one variable a (deterministic) function of the other.

Mathematically, two variables X and Y are perfectly dependent if there exists some deterministic function G() such that Y = G(X). In Exhibit 3 that function was the exponential function. This is serviceable, but not in line with our probabilistic definition of independence. In addition, there are two different types of perfect dependence that we may like to illustrate.

Using our pair of random variables (X, Y) again, there are two cases of perfect dependence:

1. 'big' values of X are always associated with 'big' values of Y;
2. 'big' values of X are always associated with 'small' values of Y.

In the first case the perfect association is referred to as *comonotonicity*. In the second case the perfect association is referred to as *countermonotonicity*. Formally, these can be defined in terms of probability statements, thus comonotonicity is defined as (for all pairs X and Y):

$$P(X \leq x) = P(Y \leq y)$$

Equation 2

Countermonotonicity is defined as:

$$P(X \leq x) = 1 - P(Y \leq y)$$

Equation 3

3.2 Desirable Properties of Dependence Measures

Embrechts, McNeil and Straumann define a number of properties they believe should be met by a reasonable dependency measure. We agree and discuss them here. Define δ(X, Y) as a dependency measure between a pair of random variables X and Y.

P1: $\delta(X, Y) = \delta(Y, X)$ symmetry

P2: $-1 \leq \delta(X, Y) \leq 1$ normalisation

P3: $\delta(X, Y) = 1$ implies X, Y are comonotonic
$\delta(X, Y) = -1$ implies X, Y are countermonotonic

P4: For some function T(), which is either monotonic[5] increasing or monotonic decreasing,

$\delta(T(X), Y) = \delta(X, Y)$ for T monotonic increasing
$\delta(T(X), Y) = -\delta(X, Y)$ for T monotonic decreasing.

P1 simply implies that the order of calculation is irrelevant.

P2 is nice in that it separates the strength of association from the units of measurement. (Dependence between weight and height should be the same whether height is measured in centimetres or metres).

P3 fits in with the intuition of dependency discussed earlier.

P4 would be a very useful feature of a dependency structure. It suggests that you could take any monotonic function (eg the log function) of X and Y, and get the same value of the dependency structure. This would get rid of the problems with Pearson correlation.

[5] A monotonic function is either continuously increasing (such as the log function) or decreasing (such as 1/x for x > 0).

3.3 Pearson Correlation

Mathematically Pearson correlation is defined as

$$\rho_p = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

Equation 4

where σ_X, σ_Y are the standard deviations of X and Y (which scale the measure so as to be between -1 and 1), and Cov(X, Y) is the covariance between X and Y.

$$Cov(X,Y) = E\left[\left(X - \overline{X}\right)\left(Y - \overline{Y}\right)\right]$$

Equation 5

The expectation is based on the difference from the mean of the two variables, and this is its downfall when distributions are not symmetric. This is possibly best shown by way of example. Consider the distributions we have been considering (lognormal with parameters $\mu = 0$, $\sigma = 1$):

- The mean is about 1.65;
- The 1st percentile is about 0.01 (1.64 from the mean);
- The 99th percentile is 10.24 (8.59 from the mean).

The covariance term will be affected by observations in the right tail which will have a much greater impact than the left tail – each of which is equally likely.

Considering the four criteria above:

- P1 (symmetry) is met – order of the calculation does not affect the result;
- P2 (normalization) fails – it can be shown that correlation can reach the extremes of (-1, 1) only when the distribution is symmetric;
- P3 fails, as illustrated in Exhibit 3 above which has the perfect dependence and correlation 0.79.
- P4 also fails – Exhibit 4 is Exhibit 1under an exponential transformation and the correlation changes from 0.44 to 0.53.

Enough said about Pearson correlation, lets examine some more robust alternatives.

3.4 Rank correlation

Given a pair of variables X and Y, Equations 2 and 3 are two conditions, one of which needs to be met for perfect dependence. If the two variables are partially (positively) dependent then an appropriate statement is that Equation 2 above is 'approximately' true. That is, the extent to which Equation 6 below is true, could be used as a measure of dependency:

$$P(X \le x) \approx P(Y \le y)$$

Equation 6

Assume that we have N observations of pairs of values, one from X and one from Y. Let $\{x_i, y_i\}$ be the i'th observation. Replace x_i with the $P(X < x_i)$ and y_i with $P(Y < y_i)$. Our pairs of vectors are now $\{P(X < x_i), P(Y < y_i)\}$ and if the two variables had perfect dependence these vectors would be identical.

As these values are now from a uniform (and hence symmetric) distribution, we can use the correlation function to obtain a measure of dependency. Finally, instead of $P(X < x_i)$ we can use the rank of x_i (as $P(X < x_i) = \text{Rank}(x_i) / N$). This measure is called Spearman's Rho (or rank correlation, or ρ_s). Thus to calculate Spearman's Rho:

1. Replace each value in the vectors X and Y with their rank[6];
2. Calculate the Pearson correlation coefficient of the rank.

Assessing Spearman's Rho against the criteria P1 – P4 above:

- Criteria of symmetry (P1) and normalcy (P2) are inherited from the correlation function;
- If the two variables are comonotonic the rank of each pair $\{x_i, y_i\}$ will be the same, and the correlation will be 1. If they are counter monotonic the correlation will be -1. Criteria P3 is satisfied.
- For any monotonic increasing or decreasing function T() the rank of the observations will remain unchanged, so criteria P4 is satisfied.

Spearman's Rho is a very popular and intuitive alternative to correlation.

3.5 Concordance

Consider two observations $\{x_i, y_i\}$ and $\{x_k, y_k\}$ from our two distributions X and Y. These two pairs of values are said to be:

- *concordant* if $(x_i - x_k)(y_i - y_k) > 0$;
- *discordant* if $(x_i - x_k)(y_i - y_k) < 0$.

Concordance is a natural extension of the positive dependency concept that x_i being 'big' implies that y_i is probably 'big' and visa versa. A measure of dependency which exploits concordance is the probability that a pair of observations is concordant, less the probability that a pair of observations is discordant. This defines Kendall's Tau[7] (or τ). For the purpose of the calculation assume that the $\{x_k, y_k\}$ are the medians, in this case (0, 0).

[6] If two observations have the same value, give them the same rank.
[7] See Johnson, Kotz and Balakrishnan (1995) or the Statsoft Electronic Manual (2005) for adjustments should points lie on the median.

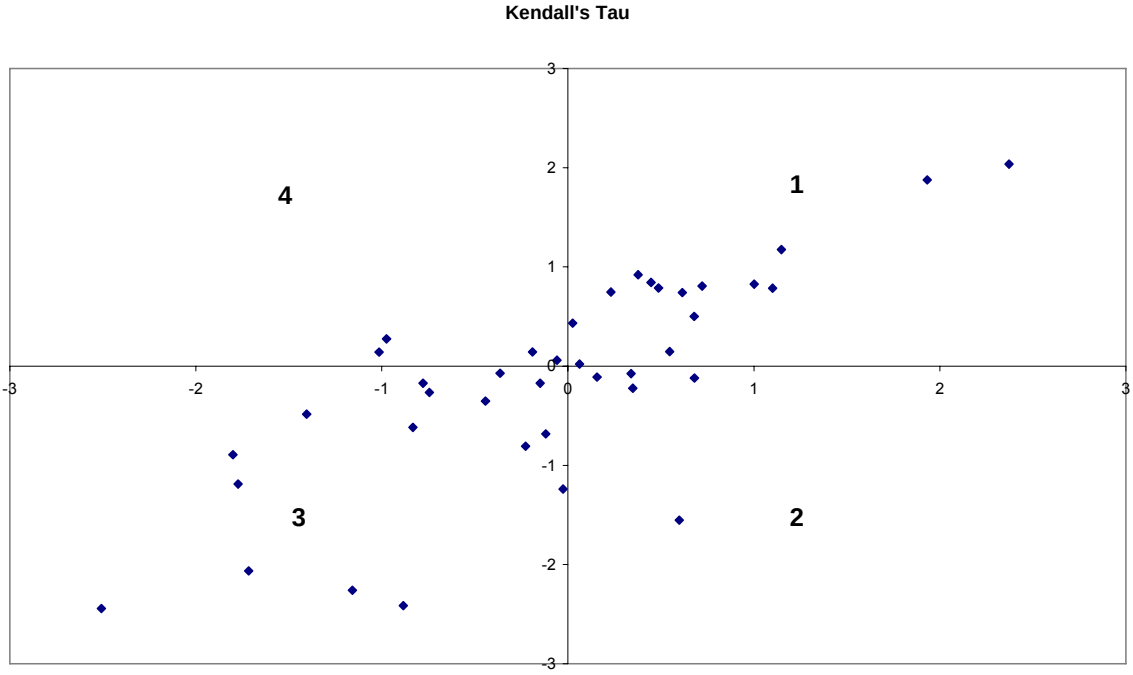

Exhibit 7 : paired unit normal random variables

Calculation is best achieved via an example. Exhibit 7 shows two variables, with the axis set at the median:

- Values in quadrants 1 and 3 are concordant (26 of these – 65%);
- Values in quadrants 2 and 4 are discordant (14 of these – 35%).

Kendall's Tau = 0.65 – 0.35 = 0.3.

Comparing to our assessment criteria:

Symmetry (P1) is preserved – swapping X and Y is equivalent to rotating the chart above through 180 degrees, wherein the top-right quadrant becomes the bottom left quadrant maintaining tau;

Normalcy (P2) is also preserved. Being the difference between two probabilities (which sum to 1) the measure is between -1 and 1.

P3 is maintained. If the two pairs of variables are comonotonic they will lie on a line passing from quadrant 3 through the origin to quadrant 1 (implies tau = 1) and if countercomonotonic they will lie on a line from quadrant 4 to quadrant 2 (implies tau = -1).

P4 is also maintained. A monotonic increasing function T() will preserve the rank of the pairs and hence all observations will remain in the same quadrant. A monotonic decreasing function will swap quadrants 1 and 2 and quadrants 3 and 4. Tau will have the same value with a different sign. (Visual thinkers will recognize that applying a monotonic increasing function to the Y values will stretch the Y axis. Once the chart x axis is realigned with the new median, each observation will remain in the same quadrant as before).

Kendall's Tau is less well known and less popular than rank correlation. It can be used as a robust estimator of linear correlation[8] in some circumstances, and is useful in parameterising certain types of copula.

3.6 Comparison of Alternatives to Linear Correlation

Exhibit 8 below shows Pearson correlation, Kendall Tau and Spearman Rho for the data graphed in Exhibits 1-3 and 5-6 above.

Exhibit	*Distribution 1*	*Distribution 2*	*Pearson*	*Spearman Rho*	*Kendall Tau*
1	N(0,1)	N(0,1)	0.44	0.39	0.27
2	N(0,1)	N(0,1)	0.89	0.89	0.71
3	N(0,1)	LN(0,1)	0.79	1	1
5	LN(0,1)	LN(0,1)	0.54	0.39	0.27
6	LN(0,1)	LN(0,2)	0.67	0.39	0.27

Exhibit 8 : Various measures of dependence

By their construction, Spearman Rho and Kendall Tau are unaffected by the underlying distributions, and can tell when there is perfect dependence in the data (Exhibit 3). Pearson is way out, suggesting a 'correlation' of only 0.79.

Exhibit 5 represents Exhibit 1 where both X and Y have undergone a monotonic increasing function change (Exhibit 1 is the log of Exhibit 5). That is, they have undergone the 'T()' transformation described in criteria P4 in Section 3.2. The table shows that Pearson correlation changes its value under this transformation, whereas Spearman's Rho and Kendall's Tau maintain their values.

Pearson correlation fails the tests for a dependency measure.

4 Copulas

Copulas are an emerging yet vast area of dependency modelling in actuarial science, with a number of excellent papers that introduce copulas[9] and give samples of applications[10]. For a more detailed analysis the book by Nelsen (1999) is recommended.

Rather than give a thorough mathematical introduction to the topic[11] we use examples to illustrate the key characteristics of copulas which makes them so potentially useful to actuaries.

Consider a bet based on two games of chance – a toss of an unbiased coin and a roll of an unbiased die. We wish to evaluate the probability that the coin toss yields heads and that the die roll is less than 3. The answer is 1/6. To derive this is a three step process.

[8] See Lindskog (2000).
[9] Embrechts et al (2001)
[10] Faivre (2003) and Isaacs (2003)
[11] Embrechts et al (2001) does this far better than I could – why try to reinvent the wheel?

Step 1: Determine the marginal distributions[12] F_1 *and* F_2 of the two variables. In the coin toss (F_1, probability of a head) is a Bernoulli with parameter 1/2. The die is discrete uniform with probability function of observing integer value y equal to 1/6 (for y between 1 and 6 inclusive). Call this F_2.

Step 2: Evaluate the marginal probabilities. Chance of a head is 1/2 and chance of a die outcome less than three is 1/3.

Step 3: Find a function C(F_1, F_2) which takes (in this case) two arguments, bounded on [0, 1] (i.e. both between 0 an 1 inclusive) and returns the required probability. This function needs to fully reflect the dependency structure of the joint distribution. One simple function is the product of the two values, and in this case it gives the correct answer – 1/6.

Three important points are raised here.

- There was nothing special about the marginal distributions we used, and in this instance the distributions used were different. We want to be able to specify the joint distribution in terms of the marginal distributions and the function C. Furthermore, we would like the function C to not be specific to the choice of the marginal distributions used.
- All distribution functions return probabilities (which are bounded on [0, 1]) so provided the function C takes as inputs probabilities (or any values between 0 and 1 inclusive) it can apply to any marginal distribution. That is, C needs only deal with the dependency structure of the distributions and is agnostic as to which distributions were used.
- The function C, if it is to potentially apply in all cases and return a valid probability, must also be bounded in its range on [0, 1].

Before formally defining C as a copula, we examine one further useful property of some copulas.

4.1 Tail Dependence

This example uses two random variables with marginal distributions N(0, 1) and correlation 0.5. We show that the distribution is not uniquely defined by providing three answers to the question 'What is the probability that both X and Y are greater than 1?

We examine three different formats for the copula C. First make the most common assumption that the joint distribution is bivariate normal. This implies the use of the Gaussian copula, which is pictured below.

[12] The marginal distributions of a multivariate distribution are simply the unconditional distributions of each variable.

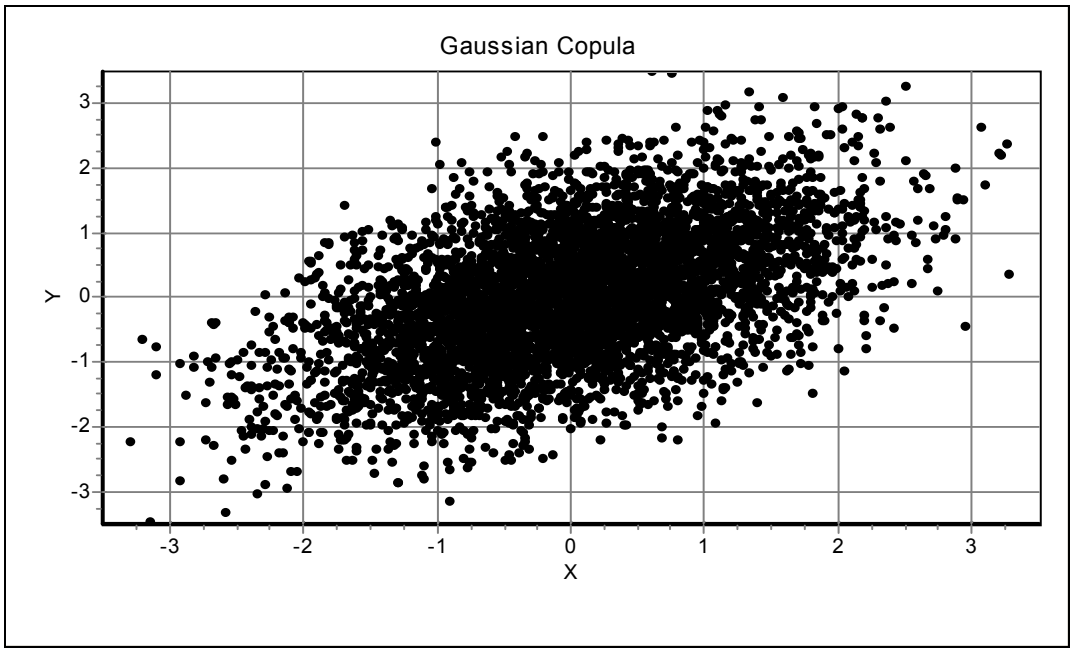

Exhibit 9 Gaussian copula

The required probability is about 6.3%. This is derived by simulating 5000 points for X and Y, and counting to evaluate the joint probability.

Next, assume that while the correlation is 0.5 that at the extreme ends of the distribution the dependency is actually higher. One form of copula that would help is the Student's t copula pictured below. The strength of the dependency in the tail is defined by the degrees of freedom of the underlying t distribution. We use a value of 4 here.

Note that the underlying marginal distribution is unchanged (if we just looked at the X values on their own they are distributed N(0, 1)). What has changed is the conditional distribution of Y, given X.

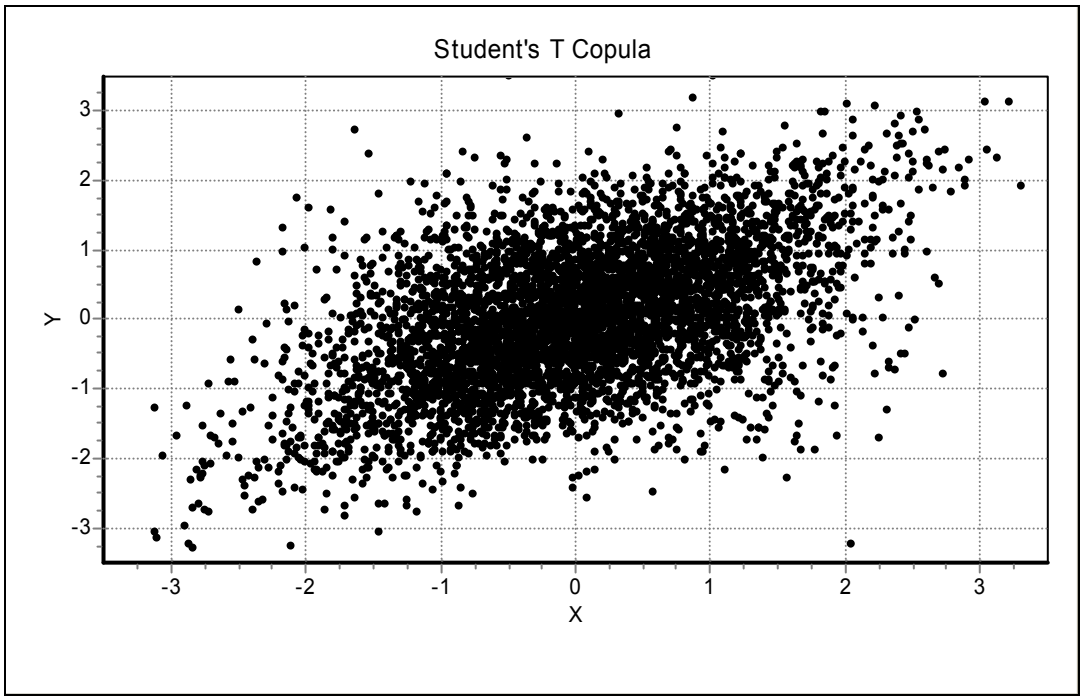

Exhibit 10 : Students T Copula

In both the upper and lower tail of the distribution, the values tend to cluster more than in the previous picture. The probability that X and Y both exceed 1 is higher at 7%. Note that by symmetry the probability that both X and Y are below 1 is also 7%.

Finally, the picture below shows the case where the right tail seems to be more strongly dependent than the left tail. This uses the so called Gumbel copula. Again the marginal distributions are both N(0, 1) but the dependency structure is stronger in the right tail than in the left tail.

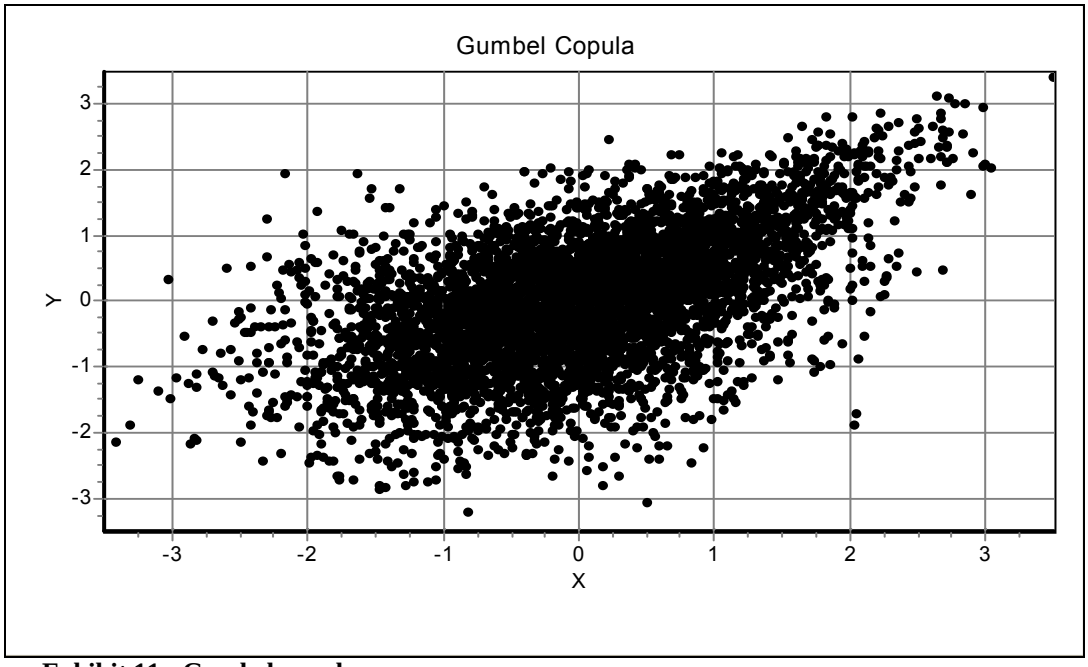

Exhibit 11 : Gumbel copula

In this case the probability of X and Y both exceeding 1 is 7.5%.

These examples illustrate the important point of tail dependency. Many financial and insurance entities (eg London and New York stock market returns, catastrophe claims in insurance) seem to have tail dependence. There is a formal definition of dependence, formal tests of tail dependence and formal measures of the strength of tail dependency. See the papers and books cited earlier for details.

4.2 What is a Copula?

The previous section has illustrated the value of a copula in terms of a multivariate probability function which:

- takes as inputs the marginal probabilities (which frees us up to make quite a wide range of assumptions concerning the marginal distributions of each variable); and
- allows flexibility to model varying tail dependence.

For those who want a formal definition, from Embrechts et al (2001), Definition 2.4:

> "An *n*-dimension copula is a function *C* with domain $[0, 1]^n$ such that
>
> - *C* is grounded and *n*-increasing.
> - *C* has margins C_k, $k = 1,2,\ldots,n$, which satisfy $C_k(u) = u$ for all u in $[0, 1]$."

This definition defines C as a multivariate (of dimension *n*) probability function which maps from each of *n* marginal probabilities (domain on $[0, 1]^n$) to return a valid probability.

For the further curious, see the references.

4.3 Application of Copulas

If the reader considers here that copulas are too theoretical and of no practical use, we illustrate here that almost every actuary signing off on technical provisions according to APRA Standards GP 210 has used a copula to estimate the reserve.

Consider the 'typical' approach to finding the 75th percentile of technical provisions for a portfolio of claims, as illustrated for example in Collings and White (2001).

Step 1: find central estimates of reserves for each line of business;

Step 2: make assumptions about the coefficient of variation, and about the distribution of the provisions (this defines the marginals);

Step 3: make an assumption of the 'correlation' between lines (our Gaussian copula, *C*); and

Step 4: combine the marginals using simulation or analytic approximation to derive the required percentile.

5 Structural Modelling

According to Sklar's Theorem[13], there is *always* a copula that can be used to map from the marginal distributions to the multivariate distribution, allowing evaluation of probabilities. It would seem then that a thorough knowledge of copulas is all that is needed to address the dependency issues confronted by actuaries. Unfortunately this is not so.

The main problem is that while Sklar tells us there is a copula to solve our problem:

- We may not be able to find it; and
- If we find it, it may be so analytically intractable that it is of no use.

Finally we would like to know *why* there is dependency between variables and why there is tail dependency[14].

Consider two long tail lines of business – Compulsory Third Party and Workers Compensation. Both have a significant exposure to medical cost inflation. Medical cost inflation would seem to be a *cause* of dependence between the two lines.

One approach to modelling the dependency structure of claims from these two lines, for example, is:

- Model medical cost inflation;
- Model the *effect* of medical cost inflation on claims;
- Model any residual (independent) random impact on the two factors.

This is referred to as a structural approach to modelling dependencies. The structural drivers need not be economic variables – they could be accident year, number of days of rain in an accident period or any other variable that impacts on claim frequency or severity.

[13] See Nelsen (1999) for particulars.

[14] Consider the (possible) tail dependency between fire claims and life insurance claims from poison. When all lighting was done by gas lamps, if the lamps went out briefly the rooms in a house would be filled by gas – poisoning the occupants. Should a match be lit to examine the cause of death then potentially there would be a very large fire indeed. Strong tail dependency indeed, but probably not as relevant today as was in the 1800's.

Structural models are very well established in finance (BARRA is an example of a company that produces sophisticated structural models of stock risk).

In actuarial circles, asset model such as the Wilkie model and Global CAP:Link[15] are structural risk models. In general insurance circles, examples are given from Taylor (1988) and Barnet and Zehnwirth (2000) (where the dependent variables are related to accident year and development year). For a partially non-parametric approach, see again the presentation given by Greg Taylor in August 2005.

6 Conclusions

This paper tries to provide a bit of insight into some of the problems associated with dependency structures, and helps to explain the desirable properties that a good dependency measure should have.

Whilst linear correlation is perhaps the most commonly used measures of dependency it has certain drawbacks. It will fail to show perfect dependence whenever the relationship between two variables is not a linear relationship and also does not always maintain the same value for a monotonic increasing or decreasing transformation of one of the variables.

There are a number of other possible dependency measures, such as Kendall's Tau and Spearman's Rho which may be more useful especially when the data is available.

We hope this paper will help actuaries to more fully consider the issues associated with dependency among risks and give some thought to this difficult topic.

[15] See Mulvey and Thorlacius (1998)

7 References

Ashe F R (1986), "An Essay at Measuring the Variance of Estimates of Outstanding Claim Payments", Astin Bulletin, Vol 16

Barnett G & Zehnwirth B (2000), "Best Estimates for Reserves", Proceedings of the Casualty Actuarial Society Vol LXXXV11, pp 245-306. See http://www.casact.org/pubs/proceed/proceed00/00245.pdf.

Bateup & Reed (2001), "Research and Data Analysis Relevant to the Development of Standards and Guidelines on Liability Valuations for General Insurance", XIII'th General Insurance Seminar 2001

Collings & White (2001), "APRA Risk Margin Analysis", XIII'th General Insurance Seminar 2001

Embrechts P, McNeil A & Straumann D (1999), "Correlation and Dependence in Risk Management: Property and Pitfalls", See http://www.risklab.ch/Papers.html

Embrechts P, Lindskog F & McNeil A (2001), "Modeling Dependence with Copulas and Applications to Risk Management", See http://www.risklab.ch/Papers.html

Faivre F (2003), "Copula: A New Vision for Economic Capital and Application to a Four Line of Business Company", ASTIN Conference. See www.astin2003.de/img/papers/faivre.pdf

Isaacs D (2003), "Capital Adequacy and Dependence", XIV'th General Insurance Seminar 2003

Johnson, Kotz & Balakrishnan (1995), "Continuous Univariate Distributions Volume 2", Second Edition, John Wiley & Sons Inc.

Lindskog F (2000), "Linear Correlation Estimation", See http://www.risklab.ch/Papers.html

Mulvey J & Thorlacius A (1998), "The Towers Perrin Global Capital Market Scenario Generation System: CAP:Link" in World Wide Asset and Liability Modeling by Ziemba W & Mulvey J, Cambridge University Press.

Nelson R (1999), "An Introduction to Copulas", Springer, New York

Priest C (2003), "Correlations: What They Mean – And More Importantly What They Don't Mean", IAAust Biennial Convention 2003, See http://www.actuaries.asn.au/PublicSite/events/Events2005/events_frameset.htm

Statsoft, www.statsoftinc.com

Taylor G (1998), "Regression Models in Claims Analysis 1: Theory", Proceedings of the Casualty Actuarial Society Vol 74, pp 354-383

Taylor G (2005), "Bootstrapping of Loss Reserves", New Developments in Quantitative Modeling of Insurance and Financial Risk Research Conference, 5 August 2005

Appendix: Sampling Error on Correlation?

Another aspect of correlation which needs consideration is the sampling error associated with any estimate of correlation. The correlation coefficient is bounded on [-1, 1] and hence we see immediately the normal distribution (or any unbounded distribution) will not give realistic distributions. The generally accepted method is to apply the Fisher transform:

$$Z = arctanh(R) = \frac{1}{2}\ln\left(\frac{(1+R)}{(1-R)}\right)$$

and assume this is normally distributed with appropriate mean and variance (consult any quality text for the values). The table below shows the 10% - 90% confidence levels of correlation coefficient for various sample size. We assume that the true value is 0.0.

Sample Size (n)	*90% Confidence Levels (assuming rho = 0)*
10	+/- 0.63
25	+/- 0.40
50	+/- 0.28
100	+/- 0.20
250	+/- 0.12
500	+/- 0.09

Exhibit 12 : Confidence Intervals on Correlation

Considering that correlations are bounded on [-1, 1], it is not feasible to draw many conclusions with small sample sizes. For example, with 25 observations (drawn from a normal distribution) an observed correlation of 0.4 is just barely statistically significant at the 90% level. The bounds would, of course, be wider if the underlying distributions were not drawn from normal distributions!

One observation that can be drawn from this is that frequently it is wise to obtain prior estimates of dependence to supplement data unless the sample size is indeed larger (> 250 observations).