Institute of Actuaries of Australia

Assessment of Diversification Benefit in Insurance Portfolios.

Prepared by Bartosz Piwcewicz

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

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

Abstract

This paper focuses on the practical approach to diversification and its impact on the value of insurance liabilities at various confidence levels.

In an insurance portfolio, the diversification benefit is intended to allow for the likelihood that not all lines of business (risks) will develop adversely or favourably at the same time. Whilst there are a number of factors that can influence the development of all lines of business, the impact is reduced at portfolio level.

There are a growing number of papers addressing the issue of dependence in a finance and insurance framework. However, their focus is on theoretical discussion with only marginal insight into practical applications. This paper is intended to fill this gap. In particular, it discusses computing empirical correlations and applying various dependence structures (both parametrical and empirical) in the multidimensional case. It also addresses various pitfalls and practical issues inherent in applications of this theory.

The discussion is supported by empirical examples and suggestions on how different stages of the assessment can be done in practice using a simulation approach.

Keywords: Diversification; dependence; correlation; elliptical copulas; empirical copula; Iman-Conover method;

1. Introduction

One of the more important issues in the insurance framework is aggregation of individual risks and understanding the associated joint distribution of outcomes. This is especially important when assessing insurance liabilities or capital requirements at various confidence levels. In Australia in particular the regulatory environment requires assessment of insurance liabilities at a 75% confidence level.

The complexity arises since not all risks (lines of business) in a portfolio will develop adversely or favourably at the same time. Whilst there are a number of factors that can influence the development of all lines of business, the impact is often reduced at portfolio level. The insurance and finance literature refer to this phenomenon as a diversification benefit. The purpose of this paper is to discuss practical methodologies for assessment of this diversification benefit.

In practice many different approaches have been developed to deal with the assessment of diversity. However, there is no golden method. In general, the most important factor is to design a sound process, which takes into account all of the stages required in developing a robust model, whilst acknowledging any data or resource limitations.

A core component of this paper is a suggested process to assess the diversification benefit for a portfolio of general insurance liabilities[1]. Importantly the main focus of the paper is on derivation of diversification benefits given underlying distributions, not on assessment of underlying distributions for individual liability classes. The paper will address the following main topics:

- How to define a diversification benefit and what methods are available for its calculation;
- How to derive correlation assumptions between classes of business;
- What dependence structure is adequate; and
- How to assess the overall level of diversification.

There is potentially a very wide range of considerations that can be taken into account eg. aggregate limits, non-lognormal distributional assumptions for individual classes, limited data etc. The more these are explicitly allowed for, the greater the complexity of the required model. In practice, an actuary has to decide on the overall level of diversification by taking into account each of the main assumptions and also by factoring in any additional allowances. At the end, considerable judgement may be needed to understand the overall impact of all relevant factors and determine an appropriate level of diversity.

The paper consists of the following main sections:

- A definition of diversification benefit and a suggested process for its assessment (chapter 2);
- A review of four modelling approaches to the calculation of the diversification benefit (chapter 3);

[1] Although this paper usually refers to claims liabilities, the methods can be generalised and also applied to premium liabilities, and hence the total insurance liabilities.

- A description of the various diversity modelling stages - selection of correlation assumptions (chapter 4), selection of a dependence structure (chapter 5) and assessment of implied diversity (chapter 6); and
- An appendix section, which shows details on algorithms used in practical modelling and some detailed mathematical workings.

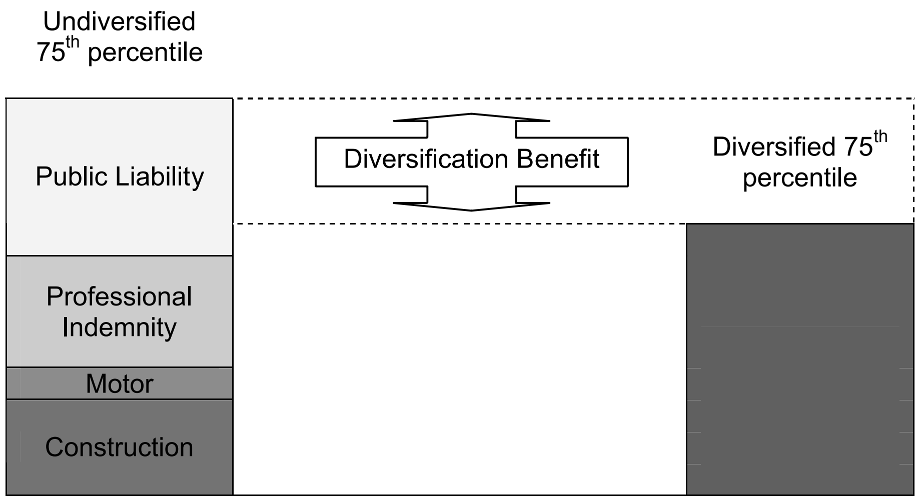
The reader should note that different chapters have different levels of complexity. Whilst I have tried to minimise the use of mathematical jargon throughout the text or deferred it to the Appendices, some sections include a considerable number of equations and formulas. As a guide, the reader who has not used mathematics for a while should focus on chapters 2, 6 and 7. Chapters 3 to 5 are more technical, however there are some sections (eg. 3.3) that provide useful high-level and non-technical discussions on the practical issues involved in the modelling stages.

2. General approach to diversification benefit

2.1 Definition

Diversification benefit can be defined either in relative or in absolute terms. For the purposes of this paper it reflects the difference between an undiversified value and a diversified value of insurance liabilities expressed as a percentage of the undiversified value.

The diagram below illustrates the above definition by considering the 75th percentile of insurance liabilities for a given portfolio. The left hand side shows the overall size of the liabilities if we did not allow for any diversification between lines of business. The right hand side shows the result after allowing for inter-dependencies between these classes.



In order to obtain an undiversified number in a simulation approach, one could generate a single set of random uniform variates and use them for each of the classes to calculate simulated liabilities[2]. The overall 75th percentile for the total portfolio would be then exactly equal to a sum of the 75th percentiles for individual classes.[3]

In practice the undiversified calculation sets an upper bound on the portfolio's liability, as it implies that given individual distributional and variability assumptions, the maximum loss at a given confidence level is equal to the sum of individual class losses at that confidence level.

On the other hand a diversification benefit allows for inter-relations within the portfolio and, in particular, it implies that not all lines develop in the same direction at the same time.

---

[2] By application of an inverse transform $u_i \rightarrow F_i^{-1}(u_i)$, where $F_i\ (x_i)$ are cumulative distribution functions ("CDF") for individual classes.

[3] In practice, this approach is equivalent to assuming a perfect dependence (comonotonicity) between classes.

It is important to note that a portfolio's diversification is only visible in the context of the total variability of the portfolio, which can be represented by a standard deviation, a variance or various percentiles. We will not be able to observe any impact of diversification benefit if we are calculating the mean of a portfolio.

2.2 General approach

A diversification benefit is a measure of how much diversity is allowed for in a particular portfolio. It is useful for benchmarking different portfolios with a similar mix of business or for comparing a basis from valuation to valuation. In order to derive a diversification benefit we must calculate both a diversified and an undiversified distribution.

It is useful to consider a portfolio with several classes of business, for which we want to estimate an insurance liability risk margin[4].

Theoretically the most appropriate way to deal with this task is to use the available data to estimate a joint distribution describing the outstanding losses for the portfolio. Then we can calculate the standard deviation and the 75th percentile of the distribution and hence calculate the required risk margin. However, as is common in insurance, there is generally not enough data to sufficiently and reliably estimate such a distribution.

A common approach to overcome the lack of sufficient data is to utilise the available data to estimate individual class distributions and pair-wise linear correlations, then to use this information in a simulation process to generate dependent vectors. There are however some flaws in this approach. The most important one is that the linear correlations may not be consistent with the individual distributions, so the overall portfolio distribution is not unique. Whilst we can try to overcome this problem by using rank correlations (eg. Spearman's rho) instead, the problem of uniqueness re-appears. In practice, actuaries often "forget" about this uniqueness problem. The selected approach depends on various aspects such as available software, skills and required accuracy of the calculation.

Many alternative methods are available. Some of them are "quick and dirty" approximations, which are not necessarily mathematically correct, but are often used in practice and described by practitioners. Such methods generally rely on special features of the Normal distribution. Other methods are more complex and usually require more advanced skills and resources. The main advantage of them is mathematical soundness and in some cases flexibility. An actuary usually has to decide which method is best for a given task.

It is also important to comment on two approaches, which are probably the most commonly used in Australia[5]. Both of them were published in 2001 aiming to present an end-to-end process for risk margin estimation. The primary focus of these papers was on

[4] In Australia, this refers to the standards imposed by APRA (the local prudential regulator), which state that the central estimate (mean) plus risk margin of the general insurance liabilities must secure the 75th percentile of the underlying distribution (subject to the risk margin being greater than or equal to half of the standard deviation). In this paper, the risk margin refers to the 75th percentile increment above the mean.

[5] Collings and White [3] and Bateup and Reed [1]

the analysis of the variability inherent within individual classes. However, they also outlined how actuaries should assess diversification benefit at a total portfolio level.

In summary, each of the papers assumes a very similar approach to diversification, which is generally equivalent to the "Lognormal approximation" described in section 3.2.1[6]. In particular, the following steps are outlined:

- Estimate variability assumptions for individual classes.
- Estimate correlation assumptions for the total portfolio (Collings and White [3] use empirical data and judgement, while Bateup and Reed [1] use judgement and industry survey).
- Given individual-class variances derived in the first stage, calculate the overall portfolio variance and use it to derive the diversified risk margin[7].

Neither of the papers differentiates between various types of correlation measures, nor separates dependence from correlation (see section 4.1 for the discussion). Overall these approaches are quick and easy to use[8], but actuaries should be aware of their deficiencies.

2.3 Step-by-step approach

When dealing with a problem such as assessment of diversification benefit, it is important to have a robust process. The approach should both correctly identify and allow for all of the crucial factors affecting the analysis, and should also have a feedback loop, which serves as an additional assessment tool for reasonableness.

Given the above, the following process for selection of a diversification benefit for a portfolio of general insurance liabilities is recommended:

1. Estimate or judgementally select distributional and variability assumptions for individual lines of business (marginal distributions $F_1, \ldots, F_n$). This may involve various stochastic methods like bootstrapping, stochastic chain ladder, Markov Monte Carlo method etc., which investigate the underlying distributions.[9]
2. Estimate or assume a matrix of pair-wise linear/rank correlations. This point is a subject of the next chapter, where I discuss the appropriateness of different measures in the context of a general insurance portfolio.
3. Estimate a joint distribution for outstanding losses of the portfolio. This step is basically the heart of the process and not surprisingly causes most problems and errors.

---

[6] Bateup and Reed [1] mention a simulation process in section 6.5.1 of their paper, which could suggest that they used a procedure similar to the one described above with a judgementally selected correlation matrix.

[7] Bateup and Reed [1] explicitly mention use of a lognormal distribution to derive the 75th percentile.

[8] In particular, Bateup and Reed [1] present a "rule of thumb" formula, which as they suggest, can be used to calculate diversification discount for various portfolios. More discussion on that is included in chapter 6 of this paper.

[9] This point is outside the scope of this paper. There are a number of papers, which deal with this topic. The reader may refer to Bateup and Reed [1], Collings and White [3] for descriptions of methods such as bootstrapping, stochastic chain ladder or the Mack method.

4. Assess the joint distribution against the underlying data and assumptions used. This step should also involve some sensitivity and scenario testing, to check if the results are consistent with the assumptions and observed experience.
5. Based on the selected joint distribution, calculate the required diversified standard deviation and the required percentiles (in our case we would calculate the 75^{th} percentile). The difference between the undiversified and the diversified percentiles gives us a diversification benefit.
6. If required, adjust the result for any judgemental loadings to allow for results of the scenario/sensitivity analysis or any other factors, which we believe are not properly modelled by the statistical methods selected.
7. Allocate the diversified figure down to the individual classes.

The above overview gives a high-level algorithm, which can be transferred into real life using a number of techniques. The main difference between approaches is usually the approach to step 3. Other steps often use very similar techniques and generally can be performed regardless of the selected method.

The focus of this paper is on steps 2 to 7. Due to its importance I have a particular emphasis on step 3.

3. Selection of Joint Distributions

I will start the overview of available methods with a description of copulas, which are a crucial technique in estimation of appropriate joint distributions. Later I will describe another mathematically sound technique and two common approximations. At the end of the chapter, I will compare all of the outlined methods and point out some flaws or possible extensions. It is important to bear in mind that the presented list is not exhaustive and could be quite easily extended.

3.1 The Copula Theory.

In terms of the developments in an insurance and finance framework, this method is arguably the most recent one. It has generally only appeared within wider actuarial literature over the last 5 to 10 years. In terms of theoretical attractiveness and practical tractability, it is probably the most appealing method at the moment.

A copula is simply a mathematical function that is used to join different distributions together. More mathematically a copula can be defined as a multivariate uniform distribution that is used to describe dependence[10]. Precisely speaking, for given continuous marginal distributions $F_1, \ldots, F_n$ and an n-copula C, a unique multivariate distribution G expressed by

$$G(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n))$$

exists, with marginal distributions $F_1, \ldots, F_n$. Additionally, if we have a proper rank correlation matrix it is possible to find a copula, which can be parameterised so that it reproduces this matrix.

Copulas do not put any restrictions on the underlying marginal distributions and also maintain their integrity. They allow for a variety of dependence structures[11], and some copulas can be easily employed in simulation algorithms. There are a great number of copula families, each offering some distinct features.

There have been many papers written on copulas, which give a good overview of the underlying theory. Many of these are listed in the Bibliography. In particular copulas in relation to bivariate distributions have received a lot of attention both in terms of the analysis of properties and also practical applications. However, there are still only a limited number of publications on the multivariate case.

In this paper, I will focus on a family of the elliptical copulas (in particular, Normal and t-Student copulas). There are three main reasons for this choice:

- These copulas can be applied in multivariate cases without any loss of dimensions.

[10] This fact is very important from a practical modelling perspective. In particular, a simulation approach usually requires generating uniform variates (see Appendix 9.1). These are generally random, and so “independent” (see “6.8 Important Caveats” section). If one wants to generate dependent samples, then a copula must be used, as it allows generating dependent uniform variates.

[11] Dependence structure relates to a specific function, which fully describes the relationship between events or between random variables. In this context, each copula implies a particular dependence structure.

- There exist quite simple simulation algorithms, which can be easily programmed in Visual Basic for Applications ("VBA").
- The family allows quite easy and flexible modelling of tail dependence[12] (eg. a Normal copula provides zero tail dependence, while a t-Student copula allows varying the strength of tail dependence).

The reader should also be aware of other copula families. In particular, Archimedean copulas have gained wide interest due to their diversity and simplicity in bivariate-case applications. However, the main challenge with this family is its extension to multivariate cases.

In particular, there are two main restrictions imposed by the construction of Archimedean copulas. Firstly, an Archimedean copula can only be parameterised by up to $n-1$ parameters (instead of $n(n-1)/2$), which represent a series of $n-1$ suitable bivariate relationships[13]. Additionally, there is also the lack of relatively easy-to-use algorithms that can be applied in practical modelling.[14]

Alternatively, practitioners could focus on parameterisation of Archimedean copulas[15] and then use some specialist software package to utilise them. However, the software can be expensive.

The theory and application of the elliptical copulas is described in-depth in chapter 5 of this paper. At this point, I only present a general algorithm for the application of any family of copulas.

The general approach is as follows:

1. Estimate or assume marginal distributions $F_1, \ldots, F_n$,
2. Estimate or assume a matrix of pair-wise rank correlations,
3. Select an n-copula C consistent with the selected rank correlation matrix (see section 5.2 for more detail),
4. Simulate a random vector of correlated uniforms $(U_1, \ldots, U_n)$ (which has the joint distribution C).
5. Calculate a correlated sample using marginal distributions by applying an inverse transform $u_i \rightarrow F_i^{-1}(u_i)$ for $i = 1, \ldots, n$. The vector $(F_1^{-1}(U_1), \ldots, F_n^{-1}(U_n))$ then has the joint distribution H with marginal distributions $F_1, \ldots, F_n$ and the selected rank correlation matrix.

---

[12] In the context of this paper, tail dependence means that extremely high/low values from one line of business imply an increased chance of a similar extraordinary effect on other lines. Refer to Embrechts, Lindskog, McNeil [7] for a mathematical definition.

[13] Note that for some portfolios this restriction may not be important, as the correlation assumptions naturally reduce to $n-1$ parameters (ie. there are only $n-1$ distinct correlation coefficients in the correlation matrix).

[14] General algorithms do exist, however practitioners need to be aware that their application requires more advanced skills and understanding of multivariate Archimedean copulas. For more discussion refer to Whelan [16].

[15] In particular, Isaacs [10] describes an approach to parameterise a Gumbel copula, which can be then used within a commercial modelling package.

I also comment on the above algorithm in the context of the Iman and Conover technique, which is described below. There are several similarities between that method and a Normal copula approach.

3.2 Other possible approaches

This section includes an overview of three alternative methods: two common approximations and one exact method called the Iman and Conover technique.[16]

3.2.1 Lognormal approximation

This is a common method in practice, as it can be done in a spreadsheet without any simulation process. The method relies on some properties of lognormal and normal distributions (abusing the latter). Sometimes, only a normal distribution is used throughout this method. This makes it more mathematically correct, but is not consistent with the generally accepted view that claim loss distributions are skewed.

The method can be outlined as follows:

1. Assume that outstanding losses for the total portfolio have a lognormal distribution,
2. Estimate or calculate variances / standard deviations for individual classes σ_i for $i=1, \ldots, n$,
3. Estimate or assume pair-wise linear correlations between individual classes r_{ij} ,
4. Calculate the portfolio variance (and so the standard deviation) using the following formula $\sigma_P^2 = \sum_{i=1}^{n} \sigma_i^2 + 2\sum_{i \neq j}^{n} \sigma_{ij}^2$, where σ_{ij}^2 is a covariance,
5. Parameterise the portfolio distribution, so $P \sim LN(\mu_P, \sigma_P)$, where $\mu_P = \sum_{i=1}^{n} \mu_i$,
6. Calculate the required 75th percentile for the total portfolio[17].

3.2.2 Normal approximation.

This method uses a simple observation made about the probability density function ("pdf") of a multivariate Normal distribution. The mathematical formula with all necessary details is shown in Appendix 9.5.

The general idea behind this approximation is to use and simplify a quadratic form, which appears in the formula for the multivariate pdf. The main difference between this approximation and the previous one is that instead of assuming a lognormal distribution for the total portfolio, we need to assume some sort of distribution (usually lognormal) for each of the individual classes.

The approximation uses the quadratic form of $\mathbf{z}^T\mathbf{R}^{-1}\mathbf{z}$, which is then simplified to the following formula $(\mathbf{x}-\mu)^\mathbf{T}\mathbf{R}(\mathbf{x}-\mu)$. The correctness of this formula is questionable. However,

---

[16] Comparison of these methods and the copula theory is presented in section 3.3

[17] If we are using a spreadsheet such as Excel we have a pre-defined function for such calculations, which finds percentiles for a lognormal distribution.

as an approximation the formula often behaves well[18] and avoids complicated matrix inversion. The (**x-μ**) is a column vector of increments between a given percentiles (the 75th in our example) and means (the best estimate) for individual classes. The square root of the quadratic form $(\mathbf{x}-\boldsymbol{\mu})^{\mathbf{T}}\mathbf{R}(\mathbf{x}-\boldsymbol{\mu})$ is a diversified increment for a given portfolio.

The application of this method can be outlined as follows:

1. Calculate differences between the 75th percentile and the mean for each of the individual classes and record them in a column vector **(x-μ)**,
2. Estimate or assume pair-wise linear correlations between individual classes r_{ij} and record them in a matrix **R**,
3. Calculate the required 75th increment above the mean[19] for the total portfolio using a square root of the formula $(\mathbf{x}-\boldsymbol{\mu})^{\mathbf{T}}\mathbf{R}(\mathbf{x}-\boldsymbol{\mu})$[20].

Again, as in the case of the lognormal approximation, all of the above can be done in a spreadsheet without any simulation.

3.2.3 The Iman and Conover Technique[21]

The main advantage of this method is that it is a mathematically proven "distribution-free" approach, which means that any type of distributions may be correlated. Besides, the procedure is also relatively easy to apply. The only complication may come with some matrix transformations such as the Cholesky decomposition or an inversion of matrices. Otherwise, there are no complex calculations.

The method can also be used under a sampling approach. In particular, we can combine it with other available correlation methods (discussed at the end of this chapter).

Additionally, the resulting samples for each distribution reflect the input distribution function from which they were drawn.

As outlined in the original work by the authors, the procedure relies on the theoretical concept that:

- If we have a random matrix **A**, in which columns have a correlation matrix **I** (the identity matrix), and a so-called “target” (input) correlation matrix **B**,
- Then there exists a transformation matrix **C** such that the columns of $\mathbf{AC^T}$ (where $\mathbf{C^T}$ is the transpose of **C**) have a correlation matrix **B**.
- Since **B** is positive definite[22] and symmetric, there exists the matrix **C**, the Cholesky decomposition of **B,** such that $\mathbf{B} = \mathbf{CC^T}$.

---

[18] Ie. the error increases with the number of lines included in the portfolio and also with the level of skewness of marginal distributions.

[19] The increment refers to the diversified 75th percentile less the portfolio mean.

[20] In Excel, this can be done using a combination of three functions:
SQRT(MMULT(TRANSPOSE(**x-μ**),MMULT(**R**, **x-μ**)))

[21] This was the first method (introduced by Iman and Conover [9]), which could be applied to arbitrary marginal distributions. Following that, there were some other similar methods presented eg. Clemen and Reilly [2].

[22] Refer to section 4.3.5 for details.

The algorithm for this method ("IC algorithm"), which can be used in practical applications, is included in Appendix 9.2. The reader can also refer to the original work by Iman and Conover for more detailed description of this technique. The IC algorithm can be employed in Monte Carlo simulations using either a spreadsheet or more cleanly as a set of procedures in VBA. There are also commercial simulation packages that rely on the IC method in a slightly modified version.

It is worth noting that the IC algorithm is quite similar to a simulation procedure for a Normal copula (i.e. if we use normal or van der Waerden scores). In particular, the step 3 (refer to the IC algorithm in Appendix 9.2) could be replaced by simulated uncorrelated normal deviates. However, the method still would have some distinct features such as:

- The Iman and Conover method works with existing samples produced from marginal distributions, while a Normal copula generates samples using the inverse transform.
- The IC technique approximates percentiles from the marginal distributions, but a Normal copula calculates them exactly.

Given the above, the IC method can be considered as a non-parametrical approach, which combines pseudo empirical data (generated from marginal distributions) into a single joint distribution without any explicit parameterisation. In particular, The IC algorithm implicitly assumes a Normal-copula-like dependence structure. However, the reader should be also aware that it is possible to extend the IC method to allow for other dependence structures. This can be done by using different scores in the IC algorithm (see Mildenhall [14]). The main requirement is that the selected scores need to be suitably normalized (eg. have mean zero and standard deviation one).

Summarising, it is useful to show how the IC algorithm performs against the Normal copula (used here as a benchmark). The table shows differences between these two approaches in a bivariate case (both of the variables have a Pearson distribution with α=13.1 and β=100,000, Spearman's correlation is set to 0.2).

Statistics	IC method	Normal copula	Difference
Mean	16,524	16,450	0.5%
Std. Dev.	3,564	3,678	-3.1%
CV	21.6%	22.4%	-3.5%
50th Perc.	16,006	15,995	0.1%
75th Perc.	18,512	18,550	-0.2%
99.5th Perc.	29,008	28,690	1.1%
99.95th Perc.	35,353	35,591	-0.7%

The results are subject to a random error as only 1,000 simulations were used to produce the above table. However, we can observe that the IC method gives a reasonably good approximation of the Normal copula output. The main concern is in regard to higher/lower percentiles and the standard deviation. These statistics are highly influenced by tail values, so a larger sample needs to be generated in order to achieve convergence.

3.3 Comparison and pitfalls

In the table below I have listed each of the described methods and their main properties (both advantages and disadvantages).

Methods	Advantages	Deficiencies
Lognormal approximation	• Fast and easy to apply, no specialist software required (only a spreadsheet). • Does not require marginal distributions to be specified. • For close to symmetrical distributions produces reasonably correct results.	• As marginal distributions become more skewed, the error becomes larger. • Does not provide any easy method to allocate the diversification benefit back to the individual classes. • Cannot be used to derive a joint distribution.
Normal approximation	• Fast and easy to apply, no specialist software required (only a spreadsheet). • Does not require marginal distributions to be specified (unless they are necessary to calculate individual percentiles). • For close to symmetrical distributions produces reasonably correct results.	• As marginal distributions become more skewed, the error becomes larger. • Diversification benefit is assumed constant for all percentiles. • Does not provide any easy method to allocate the diversification benefit back to the individual classes. • Cannot be used to derive a joint distribution.
The copula theory	• Maintains integrity of marginal distributions. • Maintains input rank/linear correlation matrix reasonably well. • Distribution-free approach. • Can be employed in simulation procedures. • Allows for various dependence structures. • Allows modelling of tail dependence. • Can be combined with other methods. • Flexible in application. • Generates an exact joint distribution.	• Requires advanced skills. • Might be quite time-consuming (in particular in development of models). • Not all available copula families can be easily used in multivariate cases. • Not all available copula families can be easily used in simulations. • Given the available data there is often no obvious single correct copula. • Sometimes it is hard to assess if the selected copula properly represents the dependence structure of the joint distribution (this may lead to large errors in particular at high percentiles).
The Iman and Conover Technique	• Maintains integrity of marginal distributions. • Maintains input rank/linear correlation matrix reasonably well. • Distribution-free approach. • Can be employed in simulation procedures. • Can be combined with other methods. • Reasonably flexible in application. • Provides a close approximation of a joint distribution. • It is possible to extend this method to other dependence structures by using non-normal scores.	• Requires advanced skills, however it is easier to implement than copulas. • The quality of an approximation of the joint distribution depends on the number of observations (this is particularly evident in calculation of higher percentiles for more than two dimensions). • In order to use the method correctly, the entire output sample from the IC method must be taken. • The number of data points in the output is equal to the number of data points in the input.

The above table highlights the main areas where a practitioner has to be careful when selecting a particular approach. Usually, the main pitfall is that users overlook or dismiss the shortcomings of the above methods. It is particularly important that the user understands the difference between dependence and correlation. There are a number of myths, some of which are discussed further in section 4.1, which can lead to misunderstanding of the methods and results.

Finally, even if we use the best method available, there are some key assumptions that need to be properly and carefully specified. If these are mis-estimated then the basic rule of GIGO (Garbage-In-Garbage-Out) applies.

Closing this chapter, I consider particular problems, where merging the copula theory and the Iman and Conover Technique may be the most efficient solution:

- Consider a portfolio with many individual valuation classes that were created to maintain homogeneity within each class. The number of classes makes it very hard to select a single correlation matrix that is positive definite (unless it was derived empirically to ensure that this property holds). In this situation one solution is to group similar classes (based on correlation assumptions) in sub-portfolios, calculate correlated results for these sub-portfolios and then perform a second-stage correlation between the sub-portfolios to derive the overall portfolio distribution. In particular, the first step could be performed using copulas and the second one using the Iman and Conover Technique. Alternatively, we could also use a second layer of copulas in the second stage, but this would require us to estimate joint distributions for the sub-portfolios.
- A portfolio consisting of a number of classes, for which dependence structures vary between different pair-wise relationships (ie. some catastrophe classes require strong upper tail dependence, while others do not require any tail dependence). Again this problem could be dealt with using a two-stage diversification process, which would group some of the classes together and allow for these special features.

The main issue with a two-stage diversification process is that some correlation can be lost. This is especially true if no tail dependence is selected in any of the stages, which could compensate for such a loss. I have set out a simple example in the Appendices, which illustrates this issue. Practitioners have to be aware of this feature and test their results, so they understand any differences between the input assumptions and final numbers.

4. Selection of correlation assumptions

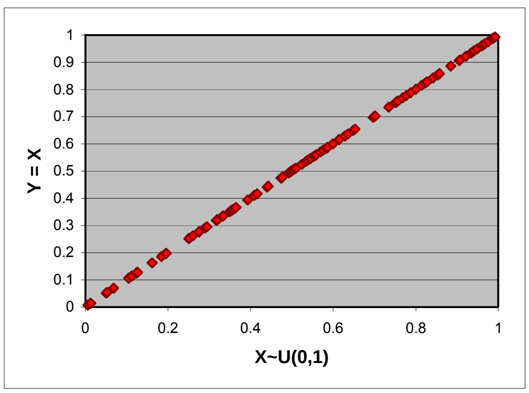
This chapter is dedicated to the second step in the general process of diversification benefit assessment. In particular, it gives details on the use of the Spearman's rank correlations for a portfolio of general insurance liabilities.

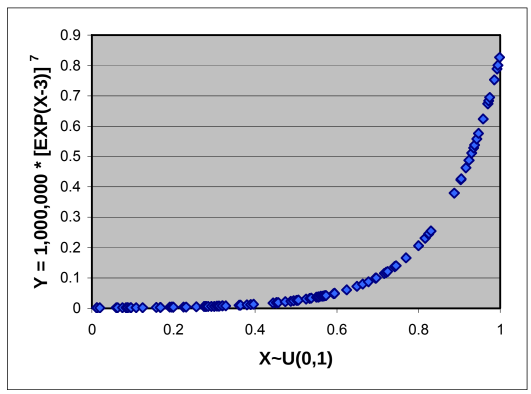
However, before moving onto the main part of this section I have documented some of the potential pitfalls.

4.1 Pitfalls and misinterpretations

The most common pitfall encountered by practitioners is that when they think about correlations, they usually have in mind linear correlations. There are two main reasons for this situation. Firstly, the linear correlation coefficient is very easy to calculate and secondly it is a natural measure of dependence for elliptical distributions (such as a multivariate normal distribution).

The two diagrams below can illustrate the problem of using linear correlation. The diagram on the left hand side shows that X and Y are perfectly linearly correlated (in fact X = Y, so the linear correlation is 100%). On the right hand side, the diagram also shows a perfect (100%) correlation, but a non-linear one. The linear correlation coefficient is only around 80%, which understates the strength of the relationship between X and Y.





Apart from the above, there are many myths about correlations and dependence[23] that affect the overall understanding of inter-relations between variables and events. I have listed several that I believe are the most common.

[23] Correlation can be simply defined as a tendency of two paired variables to move in the same direction. The dependence is more about the relationship between these two variables. In this sense, the dependence is infinite-dimensional, whilst the correlation coefficient is a scalar, which only captures a single dimension and not the potential complexity of the dependence function. This fact is also illustrated on the above right-hand-side diagram. Although the Spearman rank correlation coefficient gives us the correct perfect correlation, we would not be able to reproduce a relationship between X and Y ie. dependence (as on that graph) using the Spearman 100% coefficient or any other correlation measure.

Common misinterpretations are:

1. If X and Y are independent and Y and Z are independent, then X and Z are also independent.
   - ⇒ This is false as independence is not transitive.

2. If variables X and Y are uncorrelated then they are independent.
   - ⇒ This is generally false. However there are some specific exceptions, for which nil correlation does imply independence[24].

3. If X and Y are perfectly dependent and X and Z are perfectly dependent, then Y and Z are perfectly dependent.
   - ⇒ This observation is only false for dependence and does not directly relate to correlations. However, it highlights a difference between dependence and correlations. As is well known, there is a constraint of positive-definiteness that a correlation matrix must satisfy. This condition, however, does not generalize to the case of dependence.

4. Sensitivity testing of correlation coefficients is equivalent to sensitivity analysis of the dependence risk.
   - ⇒ In order to fully specify dependence between variables we need to use a copula (dependence function). Correlations capture only linear dependence. If we base our sensitivity analysis only on correlations we will not see the whole universe of outcomes.

5. No observations or other evidence available about the dependence between variables implies that they are independent.
   - ⇒ This issue is self-explanatory (no information about dependence should not automatically imply independence). Practitioners should be generally prudent in their approach to unknown risks, as in practice independence is usually the most optimistic scenario.

Examples of the first three misinterpretations are shown in Appendix 9.9.

4.2 Judgement vs. Empirical estimation

In general insurance, a key problem when assessing any modelling assumptions is often a data limitation, either quality or quantity. This issue is especially difficult if, in any part of the analysis, we rely on a statistical approach. Actuaries usually deal with this by applying professional judgement, which may or may not suitably compensate for the lack of data.

This problem is particularly apparent in estimation of correlation assumptions. As a result actuaries have developed some rules of thumb to select correlations, which “feel”

[24] In particular, if X and Y have Bernoulli distributions then it can be proven that nil correlation implies independence. In this context, it is also worth mentioning that normality of the marginal distributions eg. $X,Y \sim N(0,1)$ does not guarantee independence if they are uncorrelated. They also have to be jointly normally distributed.

intuitively correct for a given portfolio. Such an approach usually involves specifying a correlation structure between different classes using categories such as high, medium and low. The initial outcome could produce a correlation matrix shown below.

	Marine & Aviation	Property	Liability
Marine & Aviation	1	Medium	Medium-to-Low
Property	Medium	1	Medium
Liability	Medium-to-Low	Medium	1

The next step involves changing the words into numbers whilst maintaining relativities between classes. In doing so we have to make sure that the matrix is symmetrical and positive definite.

As this approach is entirely judgemental, we generally would not worry about specifying if the matrix includes rank correlations or linear correlations. The definition however would have to be consistent with the next steps of our analysis.

Alternatively, we could use actual data to derive a correlation matrix (ie. an empirical approach). In this situation the only questions are in regard to the benefit of using empirical correlations and the measure to be used.

Naturally, the first of the concerns is the most important. If we do not see any added value in using empirical correlations, we will not calculate them. However, there can be some significant benefits from such an approach. These include:

- Obtaining a truer picture of the correlations between individual classes in a portfolio.
- Consistency and internal integrity of the resulting correlation matrix.
- The possibility of seeing correlations from different perspectives by varying factors and criteria.
- Particularly when used over time, the analysis becomes an additional tool in understanding any material changes within the portfolio (including big incurred movements, commutation of parts of the business or writing new business).

It is important to consider these issues when setting up a framework to assess diversification for a given portfolio.

However, every approach has its difficulties. In estimating empirical correlations there may be lack of (or poor) data for some classes, which distorts the overall picture. Additionally there may be some other outside information, which is not captured by the analysis. In general, actuaries should understand their data and its limitations and apply judgemental adjustments when necessary.

4.3 Estimation of empirical correlations

There are a number of papers that satisfactorily describe and compare various correlation measures. I have included some of them in the Bibliography section (eg. Embrechts, McNeil and Straumann [6]). I believe that they give a complete picture of the topic, so I limit my description of these measures to a very high-level outline of three key measures with the main focus on practical use and deficiencies.

4.3.1 Pearson's correlation

This measure is commonly known as a linear correlation or a product moment correlation. For two variables X, Y the linear correlation coefficient equals

$$r(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)}\sqrt{Var(Y)}} .$$

For a given set of data, a linear correlation coefficient can be simply calculated using Excel formulas (based on sample versions of the covariance and variance).

4.3.2 Kendall's tau

This is the first of the rank correlation measures. Let (X_1, Y_1) and (X_2, Y_2) be independent and identically distributed random vectors, each with a joint distribution F. Then Kendall's tau for two variables X and Y equals

$$\tau(X,Y) = P\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - P\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right]$$

The first term is the probability of concordance while the second one is the probability of discordance.

In practice, we could use one of the available sample versions of this measure such as

$$\tau(X,Y) = \binom{n}{2}^{-1} \sum_{i<j} \mathrm{sgn}(X_i - X_j)\,\mathrm{sgn}(Y_i - Y_j)$$ [25]

The "*sgn*" function is a sign function, which returns +1 for positive and -1 for negative values.

4.3.3 Spearman's rho

This nonparametric index is sometimes knows as the grade correlation. It measures the monotone association between the variables. It is calculated using rankings of values, not actual values themselves. A value's "rank" is determined by its position within the min-max range of possible values for the variable.

If x and y are observations of random variables X and Y with marginal distributions F_X and F_Y then the "grades" (equivalent of ranks for the sample) are given by $F_X(x)$ and $F_Y(y)$ respectively. Spearman's rho (correlation coefficient) for these two variables is then equal to

$$\rho(X,Y) = \frac{Cov\left(F_X(X), F_Y(Y)\right)}{\sqrt{Var\left(F_X(X)\right)}\sqrt{Var\left(F_Y(Y)\right)}} .$$

---

[25] The logic behind this formula is as follows. Consider two samples, x and y, each of size n. The total number of possible pairings of x with y observations is $n(n-1)/2$. Now consider ordering the pairs by the x values and then by the y values. If $x_3 > y_3$ when ordered on both x and y then the third pair is concordant (ordered in the same way), otherwise the third pair is discordant (ordered differently). The numerator is the difference between the number of concordant and discordant pairs.

4.3.4 Comparison of the measures

There are several differences between the three approaches described. On one side, we can put the linear correlation and on the other side both Kendall's tau and Spearman's rho, as these measures have very similar characteristics.

Although the Pearson's correlation is very popular and easy to use, there are some important advantages of the other two indices. Firstly, they always exist and can take any values in the interval [-1,1]. Secondly, they are independent of the underlying marginal distributions. The latter difference is particularly important in general insurance, where we usually deal with skewed distributions from various families.

There are other more technical advantages of these two measures (for example the fact that they can be expressed in terms of copulas, which is not the case for the linear correlation). More details on these can be found in Embrechts, McNeil and Straumann [6].

The reader should note that there exist proven relationships, which connect the described measures. These are very useful, especially when we want to parameterise various copulas using the same set of assumptions. However, as Kurowicka and Cooke [12] point out, not all correlation matrices can be freely transformed. In particular, for some cases a positive-definite (or a positive semi-definite) rank correlation matrix may give us a non-positive-definite (or a non-positive-semi-definite) linear correlation matrix. This should be considered when dealing with these transformations in practice and an appropriate correction algorithm should be applied[26] if necessary.

The transformation formulae are as follows:

$r(X,Y) = \sin\left(\frac{\pi}{2}\tau(X,Y)\right)$ and

$r(X,Y) = 2\sin\left(\frac{\pi}{6}\rho(X,Y)\right)$.

4.3.5 Correlation matrix consistency

There are mathematical constraints imposed on correlation coefficients. For example, one variable cannot be strongly positively correlated with two other variables if they are strongly negatively correlated with each other. These constraints relate to the fact that a correlation matrix must always be a positive semi-definite matrix. If this criterion is not met then a simulation procedure will produce incorrect results.

[26] See Appendix 9.7 for an example of such algorithm.

Additionally, some algorithms/methods require positive definiteness (instead of positive semi-definiteness). This condition is stronger and sometimes harder to obtain[27].

4.4 Rank correlations in practice

This section sets out some steps, which could be followed in practice to estimate a correlation matrix for a general insurance portfolio.

For this example we will consider an insurance portfolio consisting of several classes. The available information includes monthly incurred movements by contract, and also various contract details such as underwriting year, line of business / modelling class, region, broker (or other distribution channel information), underwriter, etc.

We want to apply a technique that is not influenced by varying distributions of the individual classes.[28] Hence we use Spearman's rho, which uses rankings of values.

The calculation can be entirely performed in Excel, however the reader will need to develop a VBA procedure to correctly rank observations (incurred movements) in the given sample.

A suitable algorithm is:

1. Aggregate the data, taking care to suitably limit the impact of random "noise" (eg. use development half-years instead of monthly).
2. Select a number of key criteria, which are the most appropriate for the analysis (eg. we can analyse the information by development periods, by region, by currency, by underwriter, by broker, by distribution channel etc.).
3. Group data according to the selected criteria (so incurred claim movements have two dimensions: by class and by a given criterion).
4. Rank the data and calculate Spearman's rhos.

For some criteria it may be necessary to additionally split information into sub-groups in order to obtain more homogenous samples. The aggregation can be then done using appropriate weights (eg. premium information).

The final output of the above analysis is a number of rank correlation matrices, which present a wide picture of the portfolio. These can be analysed and (if required) adjusted manually, so the most appropriate assumptions are selected. The last step, depending on the size of adjustments, would have to include an overall check of positive definiteness.[29]

---

[27] In mathematical terms, the strictly positive definite condition requires eigenvalues of a matrix to be positive; the positive semi-definite condition requires them to be non-negative. The reader should refer to the following website for more details on these important properties http://en.wikipedia.org/wiki/Positive-definite_matrix

[28] This problem would be especially visible, if in our data we had a number of criteria (development periods or other groupings), for which there were nil or very large incurred movements.

[29] A simple approach of adjusting a correlation matrix to remove this lack of consistency is outlined in Appendix 9.7. The algorithm corrects an invalid correlation matrix and creates a semi-positive definite one instead.

5. Selection of a dependence structure

Generally, specifying a correlation coefficient is not sufficient to fully determine the dependence between two (or more) variables. Some approaches that seem to do this, implicitly assume some copula family (eg. the Iman and Conover Technique uses a Normal copula), which takes the correlation and specifies a particular dependence function. It is important that actuaries and other practitioners, who model dependencies, are aware of such implied structures and preferable that they have the ability to choose the best dependence structure for a particular situation.

Copulas were introduced in chapter 3 by giving a high level generic definition, which I believe is sufficient for the purpose of this discussion. There are a number of publications that give a detailed definition of copulas and their different families that offer a more theoretical background on this topic. Some of these are listed in the Bibliography.

In this chapter, I focus on a family of elliptical copulas and also describe an approach that can be used in assessment of the choice of dependence structure.

5.1 Characteristics of elliptical copulas

5.1.1 Overview

This family of copulas is based on elliptical multivariate distributions. These elliptical copulas are uniquely determined by their correlation matrix and knowledge of their type.[30]

There are two copulas in this family the Gaussian (Normal) copula and the t-Student copula. In terms of the notation, these copulas have the following forms:

- Normal copula (a copula of the n-dimensional normal distribution Φ_L^n with linear correlation matrix **L**)

  $$C_L^{Norm}(u_1,...,u_n) = \Phi_L^n\left(\Phi^{-1}(u_1),...,\Phi^{-1}(u_n)\right),$$

  where Φ^{-1} denotes the inverse of the distribution function of the univariate standard normal distribution.

- t-Student copula (a copula of the n-dimensional t-distribution with ν degrees of freedom $\Theta_{\nu,L}^n$ with linear correlation matrix **L**)

  $$C_{\nu,L}^t(u_1,...,u_n) = \Theta_{\nu,L}^n\left(t_\nu^{-1}(u_1),...,t_\nu^{-1}(u_n)\right),$$

  where t^{-1} denotes the inverse of the distribution function of the univariate standard t-Student distribution with ν degrees of freedom.

[30] The proof of this statement comes from the necessary conditions for the uniqueness of an elliptical distribution (i.e. its mean, covariance matrix and knowledge of its type). For more detailed discussion refer to Embrechts, Lindskog, McNeil [7].

5.1.2 Properties

As in the previous sections, the discussion in this chapter is in the context of practical approach to diversification in general insurance. There are several reasons why elliptical copulas fit well in this framework.

Firstly, they are fast and easy to implement in practice. I have included in Appendices (9.3 and 9.4) details on simulation algorithms, which can be successfully programmed in VBA or other programming languages. In practice this is a very important feature, as not all families of copulas can be easily programmed - especially for the multivariate case (more than two variables).

These copulas are easily parameterised by a linear correlation matrix, which can be obtained from a rank correlation matrix using the transform formulae introduced in section 4.3.4.

They are easily extended to the multivariate case. We can use them to model any number of modelling classes in a portfolio without any loss of parameterisation (note that for n classes, there are $n(n-1)/2$ correlation coefficients). This is not true, for example, for Archimedean copulas, for which the number of parameters is reduced to $n-1$. This added flexibility makes elliptical copulas especially useful in large, multi-class portfolios, where more than $n-1$ parameters are required.

The flexibility of this family of copulas is further enhanced by the fact that they can produce dependence structures with tail dependence. This is true for t-Student copulas, for which we can control this property for a given correlation matrix by the choice of the number of degrees of freedom. In particular, the lower the number of degrees of freedom, the stronger the tail dependence is. On the other hand, as the number of degrees of freedom tends to infinity, a t-Student copula becomes more like a Normal copula with no tail dependence at all.

There are also several other properties, which are not necessarily advantages, but are very important in practical applications. These are:

- Only a Normal copula parameterised by nil correlations can produce independent outcomes (this cannot be achieved with a t-Student copula).
- Elliptical copulas only exist for linear correlation coefficients taking values in (-1,1). This constraint means that it is impossible to perform simulations using a perfect-correlation assumption (all of the correlation coefficients are equal to 1)[31].
- Elliptical copulas, by their construction, are symmetrical. This means that in case of t-Student copulas, they have both upper and lower tail dependence.[32]

[31] Elliptical copulas require the correlation matrix to be strictly positive definite. Otherwise their construction is not possible. One could check that a perfect positive correlation matrix does not satisfy this condition (ie the Cholesky decomposition does not exist). In practice, this problem can be overcome if we simulate a single sample of random numbers (uniforms) and then use them for all modelling classes.

[32] Refer to chapter 3 for a definition and to chapter 6 for an example, which presents the impact of tail dependence on a portfolio.

5.2 Copula selection criteria

So far, in the discussion on copulas I have only looked at so-called parametrical copulas. They are based on prescribed forms and usually have well-studied properties. However, the question of which of them is best for a particular situation is still open.

In practice, there are several approaches available to assess goodness-of-fit of the selected parametrical copula. For example, Durrleman, Nikeghbali and Roncalli [5] give a good overview of such methods. In this paper, I describe one of them, which relies on comparison of parametrical copulas with an empirical copula.

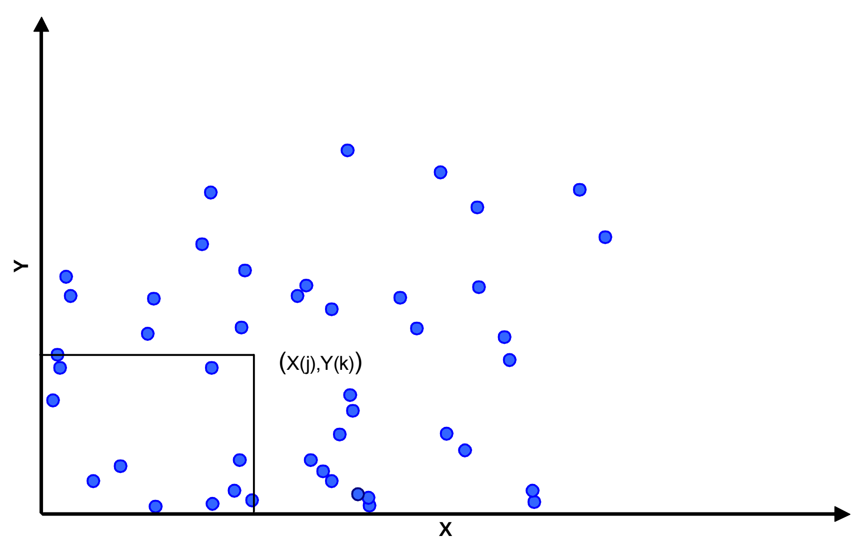
5.2.1 Empirical copula

An empirical copula (or dependence function) reflects the dependence between variables based on the actual data. It can be interpreted as the dependence equivalent of an empirical distribution function.[33]

In particular, for a bivariate sample $(x_i, y_i),\ i=1,\ \ldots,\ n$ an empirical copula can be derived using the following formula:

$$C^{Emp}(x,y) = C(\frac{j}{n},\frac{k}{n}) = \frac{\#\left\{(x,y) : x \le x_{(j)}, y \le y_{(k)}\right\}}{n},$$

where # denotes the cardinality of a set (ie. a measure of the number of elements of the set[34]), $x_{(j)}$ and $y_{(k)}$ (for $1 \le j$ and $k \le n$) denote order statistics from the sample. The numerator is just the number of data points, which lie to the left of and below the point $(x_{(j)},\ y_{(k)})$, where $x_{(j)}$ is the j[th] smallest value of the x sample and $y_{(k)}$ is the k[th] smallest value of the y sample. This can be illustrated by a graph shown below.



[33] In fact, if we transformed the marginal distributions into standard uniforms the empirical copula would be just the empirical joint distribution for the given data set.
[34] For more detail see http://en.wikipedia.org/wiki/Cardinality

The above formula allows us to calculate an empirical copula for a given sample. In practice, we would do it according to the following process:

1. Sort the observations (ie. incurred claim movements) for each of the individual classes separately. This produces vectors of the order statistics.
2. For each of the order statistics calculate the number of original data points (prior to sorting) that are simultaneously equal or smaller in each of the dimensions.
3. Normalize the calculated counts of data points by dividing them by the total number of observations.

The best way to visualize the product of the above calculation is by considering a sample, which consists of two dimensions with n observations in each of them. In this situation, the empirical copula would be calculated for each of the nxn points creating a surface in a 3-dimensional space.

The above method is easy and fast to use. If we had enough data points in the space of outcomes we would be able to create a complete empirical dependence structure. Rather than using parametrical copulas in a simulation process, we could then use an empirical copula with its values read from a pre-determined table of normalised counts of data points. Such an approach would allow us to reconstruct the overall statistical pattern of the dependence in the joint distribution.

However, there are three major practical problems with the above approach:

- In order to create a complete empirical dependence structure, we would have to have visibility of all possible outcomes, which is simply unrealistic (particularly in the tail of distributions).
- If we use the empirical copula in assessment of parametrical copulas (as described below), we face a problem of different universes (lattices[35]).
- The more classes we have in our portfolio, the more data points are required.
- Besides, even if we had such a full dataset, there is still an issue with future dependence vs. past dependence.

In regards to the last point, actuaries typically assume that past patterns will prevail and serve as a good approximation of the future (subject to their professional judgement). One could argue that this is also the case for dependencies.

Although the above arguments are valid and need to be kept in mind, there are some ways of overcoming these difficulties (eg. by aggregating some of the classes, ranking data etc.).

[35] A lattice is a partially ordered set, in which all nonempty finite subsets have both a supremum (the least upper bound) and an infimum (the greatest lower bound). More in-depth definition can be found at http://en.wikipedia.org/wiki/Lattice_(order)

5.2.2 Copula selection

There are a number of methods, which can be used to help select the best parametrical copula (eg. a t-Student copula etc.) for a given portfolio. Durrleman, Nikeghbali and Roncalli [5] give a good overview of the available techniques for selecting copulas, splitting them into two main groupings: parametric and non-parametric methods. Suitable techniques depend on the available software, available data and preferences of the user. In case of the first group (as classified by the authors), the methods are statistical and are based on estimation of copula parameters. The second group of methods is based on various empirical measures calculated directly from the data.

In terms of the efficiency of the available methods, it is hard to decide which of them is best for general insurance applications. All of them are potentially prone to material errors as available data is usually scarce.

The method, which I suggest for the purpose of this paper, is based on the empirical copula (and would appear under the non-parametric group). In general, we want to select the parametrical copula, which is closest to the empirical copula.

For a given portfolio, the approach can be summarized in three steps as follows:

I. Derive the empirical copula:
   1. For each of the classes, aggregate incurred movements by origin years and developments (if there are many modelling classes, consider aggregation to a higher level).
   2. Apply the empirical copula algorithm (described in the previous section).
   3. Rank the actual data, so it is not dependent on the absolute values and pair it with the empirical copula values.

II. Derive parametrical copulas[36]:
   1. Using the simulation algorithm for a given copula and selected/estimated assumptions (i.e. marginal distributions, correlation matrix) generate correlated observations for each of the classes.
   2. Apply the empirical copula algorithm (described in the previous section).
   3. Rank the simulated outputs, so they are not dependent on the absolute values and pair them with the derived pseudo-empirical copula values.

III. Calculate the distance between the empirical copula and parametrical copulas:
   1. Calculate sum of squared differences between the empirical copula values and corresponding parametrical outcomes for all matching rank vectors.
   2. Calculate the root of the sum of errors for each of the tested parametrical copulas and select the one, which gives the smallest value.

Mathematically, the above problem can be specified as selecting a copula $C_{(T)}^{Param}$ that minimizes $\left\| C_{(T)}^{Param} - C^{Emp}) \right\|_{L^2}$ on the discrete $\mathbf{L^2}$ norm. In this context, following Deheuvels

[36] This step is particularly important for elliptical copulas, as they do not have a closed multivariate form. For other families like Archimedean copulas, we could simply use their multivariate representation.

[4] let $X = \{(x_1^t, ..., x_N^t)\}_{t=1}^T$ be a sample of the random vector $(X_1, ..., X_N)$ and the empirical copula is given by

$$C^{Emp}\left(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\right) = \frac{1}{T}\sum_{t=1}^{T} \mathbf{1}_{\left[x_1^t \le x_1^{(t_1)}, ..., x_n^t \le x_n^{(t_n)}\right]}$$

where $x_k^{(t)}$ is the order statistics. Then we want to minimise

$$\left\| C_{(T)}^{Param} - C^{Emp} \right\|_{L^2} = \left(\sum_{t_1=1}^{T} ... \sum_{t_n=1}^{T} ... \sum_{t_N=1}^{T} \left[C_{(T)}^{Param}\left(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\right) - C^{Emp}\left(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\right) \right]^2 \right)^{\frac{1}{2}} \to \min$$

on the given lattice $\pounds = \left\{ \left(\frac{t_1}{T}, ..., \frac{t_n}{T}, ..., \frac{t_N}{T}\right) : 1 \le n \le N, t_n = 0, ..., T \right\}$.

6. Analysis of a diversification benefit

This chapter covers the last four steps of the diversification selection process. In particular, I look at implications of various assumptions/considerations on overall diversity levels and also review important aspects of the analysis.

6.1 Capital adequacy applications

Throughout this paper, I have generally taken the risk margin as the example of where diversity analysis is used. This was done to satisfy interests of actuaries who are primary concerned with the 75% probability of sufficiency. However, it is increasingly necessary to assess insurance liabilities at other (generally higher) levels of sufficiency (eg. 99.5% confidence level for run-off portfolios or the calculation of Risk Based Capital).

In this context, the methods and points presented in this paper create a foundation for the analysis of diversification benefit at various confidence levels.

This chapter focuses on the important considerations that must be taken into account at the final stages of the diversification selection process. In particular, I have included graphical/numerical examples, which give an insight into changes in diversity at different percentiles.

6.2 Diversification at different confidence levels

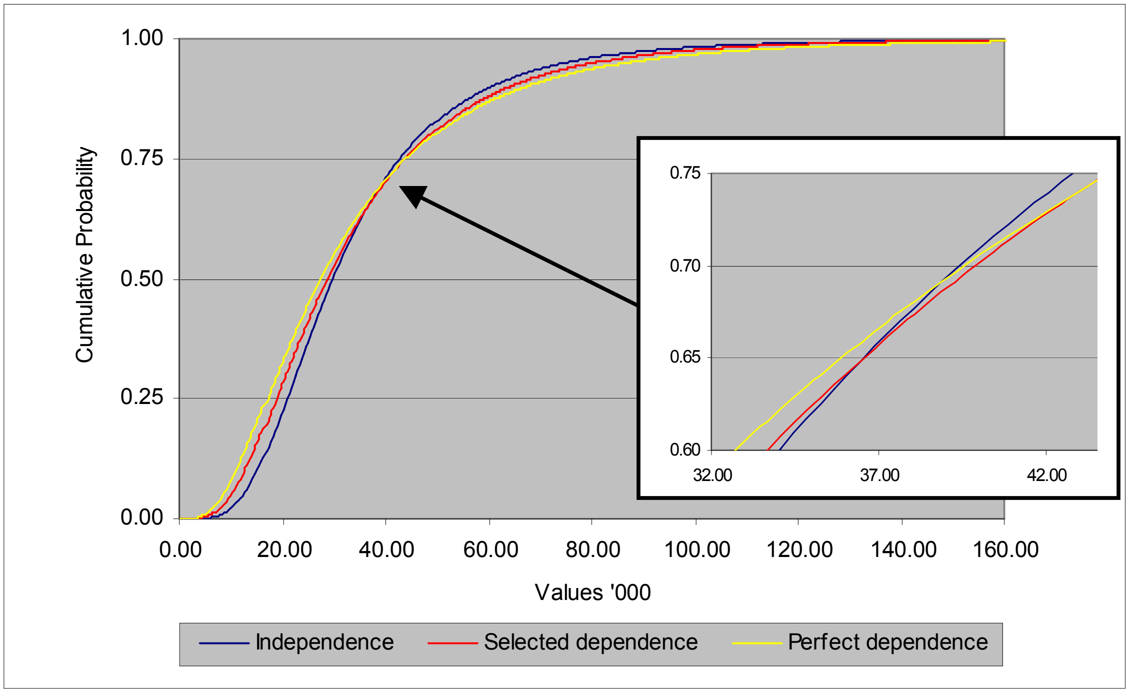
At the beginning of this paper, I defined diversification benefit as the difference between the undiversified figure and the diversified figure (expressed as a percentage of the undiversified value).

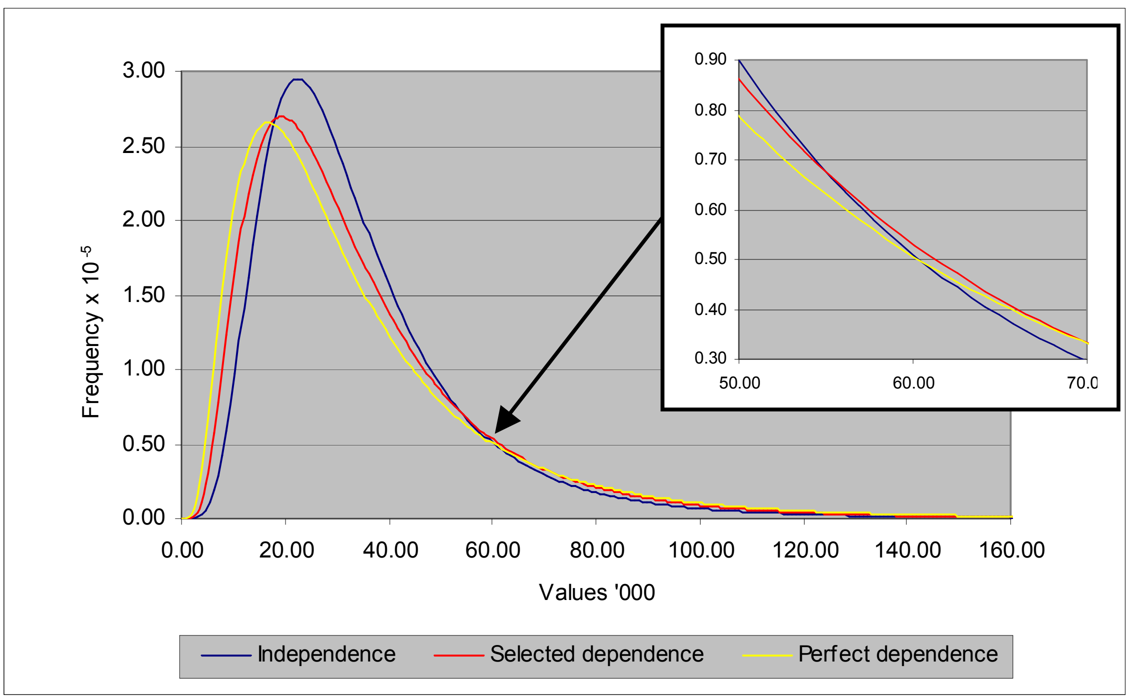
Apart from the technique described above, another useful measure is how the diversified figure relates to two boundaries. A lower bound is specified by the independence case (assuming a Normal copula with nil correlation) and an upper bound by the perfect dependence case.

In the context of these two measures, it is particularly interesting to understand how the diversification benefit changes over various confidence levels.

Practitioners quite often assume that the diversification benefit stays constant regardless of the confidence level. However, this is incorrect. The lower bound, the selected dependence and the upper bound have different distributions that intersect or overlap each other at various confidence levels. This can be best represented by the two graphs below that show joint CDF's and joint PDF's for a portfolio consisting of two classes (each with a lognormal distribution[37]) produced using a Normal copula.

[37] The classes have means of 25,000 and 10,000 with coefficients of variation of 80% each. The linear correlation assumed is 50%.





In particular, we can see in more detail how diversification benefit changes between the 60th percentile and the 99.9th percentile in the following table.

Percentiles	Independence	Selected dependence	Perfect dependence	Range-through	Diversification benefit
60th Perc.	34,141	33,648	32,710	34%	(3%)
70th Perc.	39,638	39,837	39,519	(167%)	(1%)
75th Perc.	43,092	43,834	43,923	89%	0%
80th Perc.	47,156	48,642	49,524	63%	2%
85th Perc.	52,485	54,675	56,587	53%	3%
90th Perc.	60,162	63,789	67,100	52%	5%
95th Perc.	74,596	80,732	86,938	50%	7%
97.5th Perc.	90,208	98,939	108,356	48%	9%
99.5th Perc.	131,852	146,765	168,355	41%	13%
99.9th Perc.	194,263	209,960	257,277	25%	18%

Quite often, practitioners are not aware of this fact and assume the same diversification benefit for all percentiles. This may lead to either overestimation or underestimation of the final result. It all depends on the actual position and shape of the joint probability distributions for each of the cases.

Additionally, the above example also highlights how dangerous it is to apply a fixed benchmark diversification benefit to various portfolios.

This method was applied by Bateup and Reed [1]. The authors derived a simple ("rule of thumb") formula:

Diversification Discount = 59% x (1-C)

where C was a coefficient of concentration (defined as net outstanding claims liability for largest line of business divided by total net outstanding claims liability).

The authors argued that this "rule of thumb" could be applied to an average industry portfolio. The example above highlights the danger of applying such a formula blindly. In the example the actual diversification benefit is effectively 0%, which compares to 17% (as suggested by the formula). The difference is apparent and practitioners should be aware. The diversification benefit depends on a number of factors and one can imagine situations where the difference is even larger.

6.3 Impact of individual class assumptions

There are a number of assumptions, which accompany the selected marginal distributions. Some of them can be allowed for directly in the individual class models (so they are incorporated in the shape of the marginal distributions), others need to be explicitly modelled at the portfolio level.

The following assumptions/items should be considered, while assessing diversification benefit:

- Aggregate limits. These refer to either analytical (derived using some analysis) or actual (based on the contract/policy information[38]) total exposures. Especially in regards to the latter case, it is sometimes practical to judgementally adjust these limits, so they are not unrealistically high (some covers may have no limits, which would have a dramatic impact on the top percentiles).
- Discounting. In the context of Australian regulations, this is generally not an issue at the 75th level of sufficiency. However, one could imagine that if a portfolio experienced a 99.5th scenario then the discounting profile could change quite considerably and affect various percentiles to a different extent.
- Exchange rates. This issue is particularly important for multi-currency portfolios, for which a movement in exchange rates can create a second order effect on diversity as the relative sizes of sub-portfolios shift.
- Reinsurance/retrocession arrangements. These usually are allowed on the individual class level. However, there might be some aggregate programs (eg. catastrophe arrangements), which cover a part or the total portfolio. In these cases, it is important to allow for them in the diversification analysis, as they may have a significant impact.
- Non-reinsurance recoveries (eg. subrogation, salvage). These generally relate to direct insurance and usually are allowed on the individual class level. Depending on materiality, they may have an impact on the relative sizes of various classes, and so on the overall diversity.
- Reinstatement premium[39] and other contract-specific upside items. Especially in case of reinsurance portfolios, it is possible to calculate a reinstatement premium as on offsetting item for insurance liabilities. However, this approach also has to take into account diversification. Otherwise, we could end up with inconsistent approach and too much upside.

The above list is not exhaustive. All portfolios have their own different characteristics and there will be circumstances, where other features also need to be taken into account.

Generally, each of the suggested assumptions has a direct impact on the diversified and the undiversified figure. Practitioners should be particularly careful when the relative sizes of individual classes are affected, as this situation usually leads to a shift in the overall diversification benefit.

6.4 Impact of changes in dependence structure

There are two major factors that influence the diversification benefit from the modelling perspective: correlations between classes and a dependency structure. The impact of these two assumptions usually exceeds all other allowances. Practitioners who assess risk should be looking at sensitivity and scenario testing of not only the correlation assumptions, but also the dependence structure.

---

[38] In regard to inwards reinsurance classes, the total exposures would be given by contract limits. In terms of direct insurance lines of business, we would consider sums insured or indemnity limits. Both of these could further adjusted in order to allow for individual-class outwards reinsurance protections. This would depend on our approach to individual-class assumptions (distributional and variability based either on gross or on net of reinsurance basis).

[39] Applies to inwards reinsurance portfolios.

In the table below, I included results for a portfolio consisting of ten classes. The figures were produced using five different dependence structures and the same correlation matrix. The upper part of the table shows absolute values of various percentiles, while the bottom part presents values of the top percentiles (75th and higher) relative to Normal copula as a base.

Statistics	**t-Student df = 3**	**t-Student df = 10**	**t-Student df = 15**	**t-Student df = 25**	**Normal**
Mean	318.8	318.8	318.8	318.8	318.8
St. Dev.	163.1	158.5	158.3	157.7	156.5
5th Perc.	141.3	138.3	137.9	137.5	137.4
10th Perc.	167.3	163.3	162.6	162.2	161.4
25th Perc.	215.9	212.6	212.1	211.6	211.0
50th Perc.	282.7	284.2	284.3	284.5	284.9
75th Perc.	376.0	383.6	385.2	386.2	387.8
95th Perc.	615.0	614.8	617.4	618.6	619.9
99.5th Perc.	1,090.2	1,009.8	999.5	985.9	961.9
99.95th Perc.	1,612.0	1,433.7	1,395.4	1,349.8	1,302.8
99.97th Perc.	1,709.5	1,519.8	1,477.5	1,456.2	1,373.8
99.99th Perc.	1,886.9	1,673.0	1,633.1	1,581.4	1,491.1
75th Perc.	96.97%	98.92%	99.33%	99.61%	100.00%
95th Perc.	99.20%	99.18%	99.59%	99.78%	100.00%
99.5th Perc.	113.35%	104.99%	103.92%	102.50%	100.00%
99.95th Perc.	123.74%	110.05%	107.11%	103.61%	100.00%
99.97th Perc.	124.44%	110.63%	107.55%	106.00%	100.00%
99.99th Perc.	126.54%	112.20%	109.52%	106.05%	100.00%

Each of the columns presents different dependence structure and so a different impact of tail dependence.[40] The far left figures include most of the tail dependence, while the output from a Normal copula does not include any tail dependence.

It is important to understand that the tail dependence is not a linear function. For example if we compare any two columns, we can see how the differences either widen or narrow as we move between percentiles. As expected, the biggest gap can be observed for the most extreme probability. The reader will also notice that the joint distributions intersect just above the 95th percentile.

The table clearly suggests that practitioners cannot ignore selection of the dependence structure. A simple sensitivity test, as shown above, gives a great insight into the financial impact of the selected copula. The tail dependence might be less important in the context of the risk margin. However, its significance increases if one considers capital adequacy (eg. 99.5th percentile and higher). In particular, the reader can see that the difference between copula outputs (comparing to the Normal copula) is at most -3% at the 75th percentile. At the 99.5th, this gap increases to between +2.5% and +13.4%.

[40] Refer to chapter 3 for definitions.

The presented numbers can also be assessed in the context of perfect dependence and perfect independence cases. This can give further insight into the overall range of possible outcomes.

6.5 Incremental analysis

In the analysis of diversification, it is also important to understand how much each of the classes contributes to the overall diversified figure. If we only look at the marginal distributions, we may quite easily lose sight of the bigger picture. This is especially true if there are many classes involved and the marginal distributions have different shapes.

Each of the classes contributes to the overall diversified number in two ways:

- Through their individual variability (if we add an independent risk to a portfolio then the overall standard deviation increases due to this risk's variance contribution).
- Through dependence (if we add a risk, which is correlated with other classes within the portfolio, then the overall standard deviation increases not only due to the variance, but also due to the covariance contributions)

One way to investigate the above issue is to perform a so-called "incremental analysis", where we exclude each of the risks (one at the time) and see the size of the impact on the portfolio diversification.

As in the case of diversification benefit, the diversified contributions for a particular class differ between various percentiles. This is mainly due to the shape of its distribution. However, other assumptions like tail dependence, correlation, aggregate limits and etc. also may have a significant influence. Given that, a simple linear interpolation between confidence levels usually does not give the correct approximation.

This analysis can be especially useful if we want to commute or simply remove (for a separate treatment) various parts of the portfolio. Additionally, we may also use this approach as an overall reasonableness check to see if any of the classes do not stand out or do not look sensible.

6.6 Selection criteria

In the final stage of the analysis an actuary has to decide on the total diversified figure. The decision is highly uncertain and, even if we have done all of the steps of the analysis, we still may have some doubts about the selection. This problem becomes even more complex if we are dealing with higher than 75% probabilities of sufficiency.

This is why it is very important to perform sensitivity and scenario testing. This can be done both at individual-class and at portfolio level. The analysis may not give us the required single answer, but it certainly helps to see possible upsides or downsides of our approach. On the other hand, this may serve as a base for any outside model adjustments (for a modelling and parameter error[41], for specific risks not captured by the approach, etc.)

[41] It is theoretically possible to allow for a parameter error in the modelling, albeit not discussed here.

In terms of diversity, we have to analyse the main factors driving correlations and the dependence structure. This can be done through sensitivity and scenario testing. The empirical copula (or an equivalent method) may also serve as an overall check, since it looks at the overall dependence pattern of the actual data and not only at correlations.

6.7 Allocation of diversity

It is usually required by reporting standards to allocate diversified results down to the class level[42]. In practice, there are several approaches available to deal with this task and a separate paper can be written to analyse their advantages and disadvantages. Whilst such a discussion is not the main focus of this paper, due to its sufficient importance I included a brief consideration of diversity allocation in this section.

Some of the most common methods to allocate a diversified figure (eg. the risk margin or a particular diversified increment above the mean[43]) are:

- Allocation in proportion to individual means (best estimates). This approach is probably the fastest, but it may also be the least correct of all. In particular, it does not allow for both the dependencies between classes and individual class variability.
- Allocation based on the incremental analysis (as outlined above). In some situations (eg. commutations), this approach might be the most appropriate. However, it is time-consuming.
- Allocation based on undiversified increments above the mean. This method is better than allocation by means (as it allows for inherent risks within individual classes) and is usually very fast to apply. However, it does not allow for any dependencies between classes.
- Allocation based on the Myers-Read method[44]. This method became very popular after it was published. However, there is a growing criticism of its assumptions.
- Various allocations based on option valuation formulas (eg. the insolvency exchange option[45]). These methods are reasonably new and prove to be much more adequate than other older methods like the Myers-Read method.

One could derive many other methods based on various risk measures. However, the reader must be aware of any deficiencies and select an approach that is reasonable for a particular application.

In this paper, I describe one of my preferred methods (in its two variations) that can be used in practical applications and is not included in the above list. This approach uses a particular type of the risk measures called co-measures[46] and was initially introduced in

[42] For example in Australia, Prudential Standards and Guidance Notes (GGN210.1) require actuaries to report insurance liabilities for each class of business specified by APRA. This includes allocation of the risk margin.

[43] The diversified increment above the mean is just the difference between a percentile of the joint distribution and the mean of that distribution (sum of the marginal means).

[44] Refer to Mildenhall [13] to see advantage and deficiencies of this method in practical applications.

[45] Refer to Sherris [15] for more details on this methods

[46] The definition of co-measures is outlined in Appendix 9.8, for more detailed reading refer to Kreps [11].

the context of capital allocation. The co-measures have two main advantages. They encompass positive properties of the above methods (eg. they are sensitive to the probability of default like the option valuation formulas). Additionally they are also easy to apply in simulations.

Let's assume that we have a portfolio X consisting of $X_1, \ldots, X_n$ classes (each of the X's relates to best estimates of outstanding losses[47]). Then we can define two statistics:

- A tail Value-at-Risk ("VaR") such as $TVaR_P(X) = E[X / X > P]$ and
- A co-measure-based tail VaR such as $Co - TVaR_P(X_i) = E[X_i / X > P]$

In the above notation, P relates to a given percentile for the total portfolio (for the risk margin assessment, this is equivalent to the 75th percentile). The method generally works fine and can be easily used in simulations. However, there is one downside. If we want to allocate the portfolio risk margin then it will allocate part of it to individual losses, which have no variability (the estimate of losses is constant).

The above problem is overcome by excess tail VaR. For the portfolio X, we have:

- An excess tail VaR such as $XTVaR_P(X) = E[X - m / X > P]$ and
- A excess co-measure-based tail VaR such as $Co - XTVaR_P(X_i) = E[X_i - m_i / X > P]$

In the above, m and m_i refer to means (best estimates of outstanding losses).

In most of the cases the first approach works fine, however if we deal with a portfolio, which does have some fixed losses then it is necessary to use the second method[48].

6.8 Important Caveats

There are a number of caveats, which need to be kept in mind by the reader while using the approaches outlined in this paper. In most cases, I have highlighted any possible flaws or problems throughout the text. However, there are other ones, which I present in this section.

Most of the issues, which are related to the presented methods, lie at the heart of the projection process itself. Others are due to the limitation imposed either by software or other sources.

I believe that the following caveats must be kept in mind:

---

[47] Note that in a simulation approach, these would refer to simulated losses. In particular, if we had 50,000 simulations this would produce 50,000 vectors of individual-class losses.

[48] We could extend this approach even further by adjusting probabilities, so that larger losses receive increasingly larger weights. In this case, the co-measure could be called a weighted excess $TVaR$ ("$WXTVaR$").

- Any technique used to model real-life processes is only an approximation. As such, the methods cannot replicate future events with certainty.
- Actual data is generally limited, which makes any estimation difficult and possibly incorrect. Judgement needs to be used to assess any estimates.
- The past experience may not repeat in the future. This is true for most (if not all) of the assumptions including correlations, the dependence structure and assumed distributions for individual classes. The time-wise stability of all of the above must be considered in detail. In particular, practitioners should be mindful when using non-parametrical methods (eg. simulating from an empirical copula, bootstrapping etc.). These approaches usually apply past patterns to the future automatically without any flexibility.
- The past experience is very likely to represent only part of the total distribution of outcomes. In this sense, if we are estimating liabilities at extreme probabilities of sufficiency, we may not have any actual observations at this level.
- The results obtained for one portfolio (eg. correlation matrix, dependence structure etc.) are usually not appropriate to other portfolios. Practitioners need to keep this in mind when using any pre-determined assumptions.
- Various simulation procedures rely on randomly generated numbers. In the context of any simulation algorithm presented in this paper, random numbers are actually pseudo-random numbers.
- Any simulation approach is prone to a random error, especially if we are interested in calculating extreme confidence levels. In general, results do converge with increased number of simulations, however in order to obtain a stable 99.5th percentile we need a much larger output sample than for a 75th percentile.

7. Summary and Conclusions

The main purpose of this paper is to suggest a general approach to diversification, which can be readily adopted in practice. In particular, practitioners could use an Excel spreadsheet supported by VBA to implement all of the steps of the approach.

The described framework is quite general and although there are a number of methods suggested for each stage, the reader should be aware of their limitations and consider alternative approaches if available. In practice there is usually no single correct formula, which applies to all circumstances.

The paper also focuses on a number of important considerations, which should be taken into account while assessing diversity. In particular, I have included quite detailed discussion of assumptions on both individual-class (eg. marginal distributions, aggregate limits, exchange rates etc.) and portfolio level (eg. dependence structure, correlations etc.), which have a direct impact on the diversification benefit.

Additionally, the paper also looks at a number of common myths and misinterpretation commonly committed by practitioners. All of those should be well understood and avoided in practical modelling.

It is also important to note that although this paper discusses diversification in the context of general insurance, some or all of the concepts can easily be extended to other areas of insurance, finance or, even, engineering.

Practitioners should bear in mind that existing methods are not entirely studied nor widely understood. This is especially true in regards to the assessment of dependence structures. This paper describes use of the elliptical copulas, which are valued for their properties. However, there are many other families available and one could use them instead.

There is also a big gap in copula assessment methods. The described empirical copula approach is one of many and, as pointed out, it has its limitations. The main problems when dealing with a copula selection in general insurance are multidimensionality and scarcity of data. In general, none of the existing methods properly addresses these issues.

In contrast, the profession and academics have done quite a lot of study on loss distributions. As a result, many actuaries quite often assume lognormal distribution for outstanding claim losses in their analysis of various sufficiency levels. This assumption is almost treated as a standard/base from which one can deviate in case of existing evidence.

In my view, the same study should be done in terms of dependence. There might be sufficient evidence to justify a particular copula or a family of copulas that is the most appropriate for general insurance applications. If this cannot be achieved then at least more reliable methods of copula assessment should be derived.

8. Acknowledgments

This paper was motivated by my work experience gained over the past two years at Cobalt Solutions (formerly Cobalt Run-Off Services), where I had an opportunity to develop a risk margin and a 99.5% modelling framework.

In particular, I would like to express my gratitude and appreciation to the peer-reviewers of this paper:

- Henry Onions - for his encouragement, valuable peer-review feedback and general support in writing this paper.
- Emiliano Valdez - for his valuable comments and suggestions.

I also would like to thank my colleagues from the Actuarial Team at Cobalt Solutions for their encouragement and valuable feedback.

This paper is the work of its author and does not represent the views of the author's employer. Whilst, as the author, I have made every effort to provide accurate and current information I do not warrant that the information contained herein is in every respect accurate and complete. I expressly disclaim any responsibility for any errors or omissions or for any reliance placed on this paper.

If you have any comments on or queries about the paper, please feel free to email me at my email address:

bartosz_piwcewicz@cobaltsolutions.net

9. Appendices

9.1 Overview of a general simulation approach

This section may be less interesting to practitioners who are familiar with simulations. It might be however useful for others who are new to this topic and want to apply some of the ideas presented in this paper.

A general approach to generate random numbers from a given distribution $F\ (x_i)$ can be outlined as follows:

1. Using some random-number generator produce a set of uniformly distributed variates ("uniforms") $u_1, \ldots, u_n \sim U(0,1)$.
   - This can be done using an Excel function RAND()[49] or some other algorithms.
   - The uniforms $u_1, \ldots, u_n$ are generally uncorrelated and independent (note that some algorithms may create dependent samples).
2. Using inverse transform $u_i \to F^{-1}(u_i)$ obtain a sample from the given distribution.
   - For some distributions, this can be done directly using their CDF eg. Pareto, Exponential etc.
   - For other distributions, a numerical calculation needs to be done by integration of the density function.
   - Some distributions like Normal or Lognormal are standard functions in Excel.

In some cases instead of generating random uniforms one needs to generate random variates, which are standard normally distributed ("normals") ie. $\sim N(0,1)$. This is particularly true for application of elliptical copula algorithms (as shown below). The reader should be aware that there have been a number of algorithms developed which allow generating random normals.

If the reader wants to simulate *n* dependent samples, then a copula (or some equivalent method) must be used in the simulation algorithm. The obtained dependent uniforms can be then utilised in the second step of the above process.

9.2 The Iman and Conover algorithm[50]

The algorithm for the Iman and Conover technique can be outlined as follows:

1. Using marginal distributions generate N values (draws) for each of the classes and store them in a matrix **X** (record values for each class in columns).
2. Estimate or assume pair-wise rank correlations between individual classes ρ_{ij} and record them in a matrix **T** (the input rank correlation matrix for a transformation of the columns of **X**). Note that **T** needs to be positive definite and symmetric.

[49] The reader should be aware of documented drawbacks of this function.

[50] This algorithm is consistent with the original work presented by the authors. An application of this method can also be found in Hart, Hayes and Babcock[8]

3. Generate a matrix **R** of the same size as **X** that contains van der Waerden scores ($\Phi^{-1}(i / N+1)$, where Φ^{-1} is the inverse of the standard normal distribution function, N is the number of draws, and $i = 1, ..., N$).
4. Randomly shuffle van der Waerden scores within each of the columns, so that **R** has a correlation matrix **I** (identity matrix).
5. Find the Cholesky decomposition for **T** such that $\mathbf{T} = \mathbf{PP}^{\mathbf{T}}$ (where **P** is a lower triangular matrix).
6. Create a transformed score matrix $\mathbf{R^*} = \mathbf{RP}^{\mathbf{T}}$ (where columns of **R*** have a rank correlation matrix **M**, which approximates to the input rank correlation matrix **T**).
7. Sort elements of **X** so they are arranged in the same ranking as in **R*** (the columns of the transformed **X** matrix will also have a rank correlation matrix equal to **M**).
8. Matrix **M** only approximates matrix **T**. This is due to the random simulation error that affects the correlation matrix of **R** (ie. it is not an identity matrix). The authors of the method suggested applying a variance reduction procedure in order to minimize deviations between **M** and **T**. According to this procedure, we need to find a matrix **M***, which is closer to **T** than **M**. The algorithm is as follows:
   i. Let **J** represent the actual correlation matrix for the columns of **R**.[51]
   ii. Let **T** represent the input correlation matrix.
   iii. Find for **T** a matrix **S** such that $\mathbf{T} = \mathbf{SJS}^{\mathbf{T}}$.
   iv. Find Cholesky decomposition for both matrices **T** and **J** (both are positive definite and symmetric matrices) such that $\mathbf{J} = \mathbf{UU}^{\mathbf{T}}$ and $\mathbf{T} = \mathbf{VV}^{\mathbf{T}}$. So $\mathbf{VV}^{\mathbf{T}} = \mathbf{T} = \mathbf{SJS}^{\mathbf{T}} = \mathbf{SUU}^{\mathbf{T}}\mathbf{S}^{\mathbf{T}}$. This implies that $\mathbf{S} = \mathbf{VU}^{-1}$ (where $\mathbf{U}^{-1}$ is the inverse of **U**).
   v. Calculate the transformation $\mathbf{RS}^{\mathbf{T}}$ and note that the columns of $\mathbf{RS}^{\mathbf{T}}$ have a rank correlation matrix that is equal to **M***. The matrix **M*** provides a better estimate of matrix **T** than **M** does.
9. Sort the original values from the marginal distributions (the columns of **X**) so they match ranks with the data in the columns of $\mathbf{RS}^{\mathbf{T}}$ (the rank correlation matrix of the sorted original values is equal to **M*** and approaches the input rank correlation matrix **T**).
10. The sum of each row represents the total loss for the portfolio. The diversified percentile can be calculated based on the row totals.

9.3 Normal copula algorithm

Let's assume that we have a portfolio X consisting of $X_1, ..., X_n$ classes with marginal distributions $F_1, ..., F_n$ and a linear correlation matrix **R**.

Then the following algorithm can be applied to produce correlated outcomes:
1. Generate n independent variates $z_1, ..., z_n$ from $N(0,1)$.
2. Find the Cholesky decomposition **A** of **R**.
3. Set **x**=**Az**.
4. Set $u_i = \Phi(x_i)$ for $i=1, ..., n$.

[51] Note that in theory **J** would be equal to **I**. However it is not the case due to the random error.

5. $u_1, \ldots, u_n$ is just a vector of correlated uniforms from a Normal copula.
6. Use an inverse transform $u_i \rightarrow F_i^{-1}(u_i)$ for $i = 1, \ldots, n$ to derive simulated losses for each of the classes (if we run 50,000 iterations of this algorithm then we obtain 50,000 vectors $(F_1^{-1}(U_1), \ldots, F_n^{-1}(U_n))$ of simulated losses).

9.4 t-Student copula algorithm

Let's assume that we have a portfolio X consisting of $X_1, \ldots, X_n$ classes with marginal distributions $F_1, \ldots, F_n$ and a linear correlation matrix **R**.

Then the following algorithm can be applied to produce correlated outcomes:
1. Generate n independent variates $z_1, \ldots, z_n$ from $N(0,1)$.
2. Generate a variate s from χ_v^2 independent of $z_1, \ldots, z_n$.
3. Find the Cholesky decomposition **A** of **R**.
4. Set **y**=**Az**.
5. Set **x**=$\frac{\sqrt{v}}{\sqrt{s}}$**y**.
6. Set $u_i = t_v(x_i)$ for $i=1, \ldots, n$.
7. $u_1, \ldots, u_n$ is just a vector of correlated uniforms from a t-Student copula with v degrees of freedom.
8. Use an inverse transform $u_i \rightarrow F_i^{-1}(u_i)$ for $i = 1, \ldots, n$ to derive simulated losses for each of the classes (if we run 50,000 iterations of this algorithm then we obtain 50,000 vectors $(F_1^{-1}(U_1), \ldots, F_n^{-1}(U_n))$ of simulated losses).

9.5 Multivariate Normal distribution

If **x** is a n-vector and **x**~N(μ, Σ), then the joint pdf is

$$f(\mathbf{x}) = (2\pi)^{-n/2} |\mathbf{\Sigma}|^{-1/2} \exp\left[-\frac{1}{2} (\mathbf{x}-\boldsymbol{\mu})^T \mathbf{\Sigma}^{-1} (\mathbf{x}-\boldsymbol{\mu}) \right]$$

Now let

$$\mathbf{R}_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}$$ (**R** is a correlation matrix)

$$f(\mathbf{x}) = (2\pi)^{-n/2} (\sigma_1 \sigma_2 \ldots \sigma_n)^{-1} |\mathbf{R}|^{-1/2} \exp\left[-\frac{1}{2} \mathbf{z}^T \mathbf{R}^{-1} \mathbf{z} \right]$$

where $z_i = \frac{(x_i - \mu)}{\sigma_i}$

9.6 Two-stage correlation example

Assuming A,B and C have normal distributions.

st dev (A)	20,000
st dev (B)	8,000
st dev (C)	15,000

corr(A,B)	0.7
corr(A,C)	0.7
corr(B,C)	0.2

Linear correlation matrix

	A	B	C
A	1	0.7	0.7
B	0.7	1	0.2
C	0.7	0.2	1

Cov(A,B)	112,000,000	=stdev(A)*stdev(B)*corr(A,B)
Cov(A,C)	210,000,000	=stdev(A)*stdev(C)*corr(A,C)
Cov(B,C)	24,000,000	=stdev(B)*stdev(C)*corr(B,C)

Variance of the product of the 1st stage correlations

Variance of the product of the 2nd stage correlations

Cov(A,B,C)	1,381,000,000	=var(A) + var(B) + var(C) + 2*(cov(A,B) + cov(A,C) + cov(B,C))
CombVar(B,C)	337,000,000	=var(B) + var(C) + 2*cov(B,C)
Cov(A,(B,C))	1,251,011,673	=var(A) + CombVar(B,C) + 2*stdev(A)*sqrt(CombVar(B,C))*corr(A,B) [or '*corr(A,C)']

sqrt(cov(A,B,C))	37,162	<--- standard deviation for the one-stage correlation approach
sqrt(cov(A,(B,C))	35,370	<--- standard deviation for the two-stage correlation approach

Note:
The difference becomes bigger if:
corr(B,C) becomes lower
corr(A,C) and corr(A,B) become bigger

9.7 Positive semi-definiteness correction formula

In general, an inconsistent correlation matrix can be uncovered by application of the Cholesky decomposition[52]. A matrix is not positive semi-definite if one of the eigenvalues drops below zero.

The following algorithm can be used to adjust an inconsistent matrix:

1. Find the smallest eigenvalue λ_0 (note that λ_0 is negative and the Cholesky decomposition algorithm falls over at this point).
2. Adjust the eigenvalues so that the smallest eigenvalue equals zero by adding the product of - λ_0 and the identity matrix **I** to the correlation matrix **R**. So the modified correlation matrix **R'** = **R** - λ_0**I**.
3. Divide **R'** by $(1 - \lambda_0)$ so the resulting matirx **R''** $= 1/(1 - \lambda_0)$**R'** (this normalizes the overall matrix and brings the diagonal values back to 1)

[52] The Cholesky decomposition requires the correlation matrix to be strictly positive definite. The method presented here creates a positive semi-definite matrix. In order to use it in the Cholesky decomposition a further adjustment needs to be made to obtain a positive definite matrix (eg. by changing particular entries in the matrix).

4. This new **R''** matrix is positive semi-definite (i.e at least one of the eigenvectors is zero and others are greater than zero), and can be used in further calculations.

This method adjusts the whole correlation matrix and can be used iteratively until the problem is removed. Alternatively, one could adjust a single coefficient of correlation (keeping the overall symmetry of the matrix); which causes the problem. The selected approach depends on the user.

9.8 Definition of co-measures

A co-measure can be defined for any risk measure that is expressed in terms of a conditional expected value. In particular, if we have a risk measure $R(X)$ for a risk X with mean m such that:

$R(X) = E\left[(X - am)f(x)/condition\right]$, for some value *a* and function f.

Then given that X is the sum of n X_i's (each with mean m_i), the co-measure for X_i is:

$co\text{ - }R(X_i) = E\left[(X_i - am_i)f(x)/condition\right]$, where the condition is the same as for X.

Using the fact that expectations are additive, the sum of individual $co\text{-}R$'s of the n X_i's is $R(X)$.

9.9 Examples of correlation and dependence myths[53]

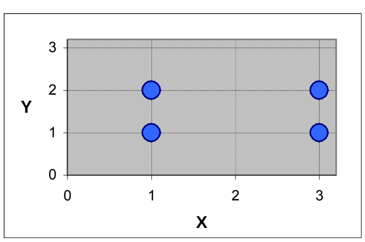
The below examples refer to section "4.1 Pitfalls and misinterpretations".

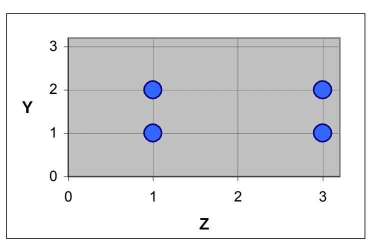
Example of point 1.

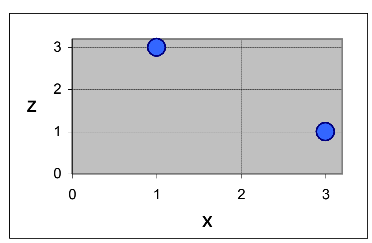
Let's consider discrete distributions (X, Y, Z) such as:

X	Y	Z	Probability
1	1	3	0.25
1	2	3	0.25
3	2	1	0.25
3	1	1	0.25

The relationships between X and Y, Y and Z and X and Z can be presented as follows:





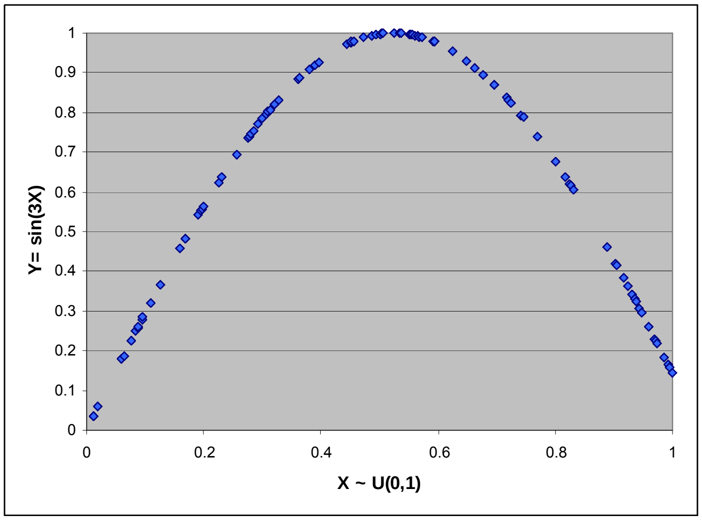


[53] A useful list of examples (including the ones presented in this section) can be found at http://www.ramas.com/wttreprints/Myths.pdf

The above diagrams show, that although X and Y and Y and Z are independent – X and Z are perfectly negatively dependent.

Example of point 2.

Let's consider variables X and Y. The relationship between them is as shown on the diagram. There is a clear dependence between the variables, although the correlation coefficients are nil.

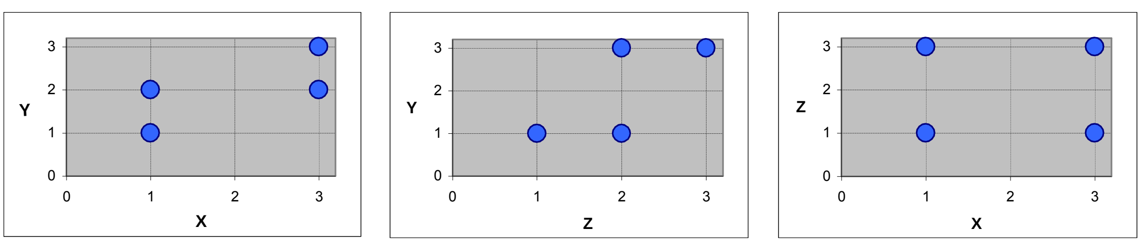


Example of point 3.

As in Example of point 1, let's consider discrete distributions (X, Y, Z) such as:

X	Y	Z	Probability
1	1	1	0.25
1	2	3	0.25
3	2	1	0.25
3	3	3	0.25

The relationships between X and Y, Y and Z and X and Z can be presented as follows:



The above diagrams show, that although X and Y and Y and Z are perfectly dependent – X and Z are independent.

10. Bibliography

1. Bateup, R. and I. Reed (2001): "Research and Data Analysis Relevant to the Developments of Standards and Guidelines on Liability Valuation for General Insurance". The Institute of Actuaries Australia.
2. Clemen, R. and T. Reilly (1999): "Correlations and copulas for decision and risk analysis. Management Science 45: pages 208-224.
3. Collings, S. and G. White (2001): "APRA Risk Margin Analysis". Paper presented to the Institute of Actuaries of Australia, XIII General Insurance Seminar 2001.
4. Deheuvels, P. (1979): "La fonction de dépendence empirique et ses propriétés: un test non paramétrique d'indépendence". Bulletin de la Classe des Sciences de l'Academie Royale de Belgique (ser. 5) 65: pages 274-292.
5. Durrleman, V., A. Nikeghbali and T. Roncalli (2000): "Which copula is the right one?" Groupe de Recherche Opérationnelle, Crédit Lyonnais. Available from http://defaultrisk.com/pp_corr_68.htm
6. Embrechts, P., A. J. McNeil and D. Straumann (1999): "Correlation and Dependence in Risk Management: Properties and Pitfalls". Available from http://www.ccfz.ch/files/Embrechtspitfalls.pdf
7. Embrechts, P., F. Lindskog, and A. J. McNeil (2001): "Modelling Dependence with Copulas and Applications to Risk Management". Department of Mathematics ETHZ Zurich. Available from http://defaultrisk.com/pp_corr_19.htm
8. Hart, C. E., D. J. Hayes and B. A. Babcock (2003): " Insuring Eggs in Baskets". Working Paper 03-WP 339, July 2003. Center for Agricultural and Rural Development, Iowa State University. Available from http://www.econ.iastate.edu/research/webpapers/paper_10669.pdf
9. Iman, R. L. and W. J. Conover (1982): "A distribution-free approach to inducing rank correlation among input variables". Communications in Statistics B11: pages 311-334.
10. Isaacs, D. (2003): "Capital Adequacy and Dependence". Paper presented to the Institute of Actuaries of Australia, XIV General Insurance Seminar 2003.
11. Kreps R.: "Riskiness Leverage Models". Available from http://www.casact.org/library/05pcas/kreps.pdf
12. Kurowicka, D. and R. M. Cooke: "Conditional, partial and rank correlation for the elliptical copula; dependence modelling in uncertainty analysis". Available from http://ssor.twi.tudelft.nl/~risk/br/uncertainty.html
13. Mildenhall, S. J. (2003): "A Note on the Myers and Read Capital Allocation Formula". Available from: http://www.casact.org/pubs/forum/03fforum/03ff419.pdf
14. Mildenhall, S. J. (2005): " Correlation and Aggregate Loss Distributions with an Emphasis On The Iman-Conover Method". CAS Working Party on Correlation Available from: http://www.mynl.com/wp/ic.pdf
15. Sherris, M. (2004): "Solvency, Capital Allocation and Fair Rate of Return in Insurance". Actuarial Studies, Faculty of Commerce and Economics UNSW, Sydney, AUSTRALIA. Available from: http://www.actuary.unsw.edu.au/files/sherriscapall.pdf
16. Whelan, N. (2004): "Sampling from Archimedean Copulas". Quantitative Finance Volume 4, Issue 3 (June 2004), pages 339-352.