ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/237205902

Dependence Modelling via the Copula Method

Technical Report · January 2006

CITATION
1

READS
1,162

3 authors:

Helen Arnold
1 PUBLICATION **1** CITATION

SEE PROFILE

Pavel V. Shevchenko
Macquarie University
156 PUBLICATIONS **1,271** CITATIONS

SEE PROFILE

Xiaolin Luo
The Commonwealth Scientific and Industrial Research Organisation
22 PUBLICATIONS **163** CITATIONS

SEE PROFILE

All content following this page was uploaded by Pavel V. Shevchenko on 07 July 2014.

The user has requested enhancement of the downloaded file.

Dependence Modelling via the Copula Method

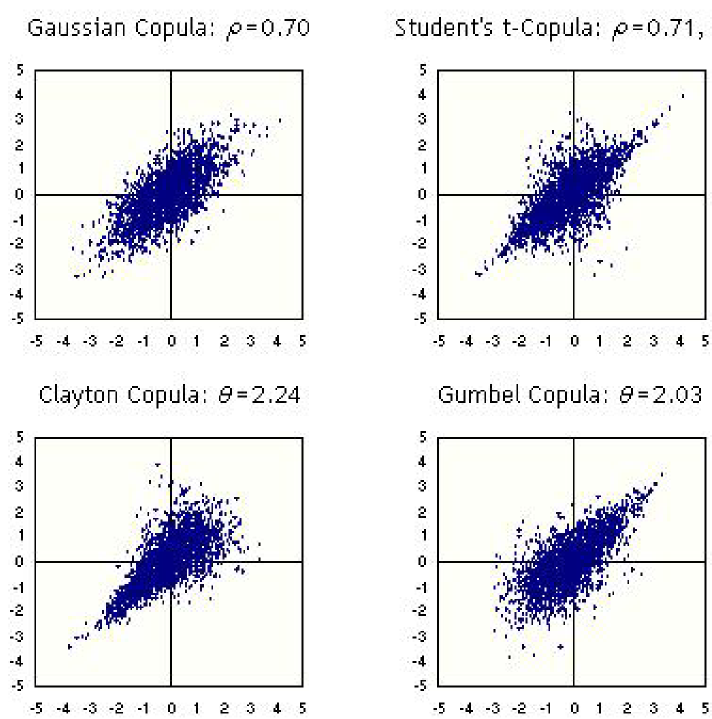
prepared by Helen Arnold

Vacation student project

Vacation student: Helen Arnold
Supervisor: Pavel Shevchenko
Co-supervisor: Xiaolin Luo

Quantitative Risk Management Group
CSIRO Mathematical and Information Sciences
Macquarie University Campus, Locked Bag 17, North Ryde NSW 1670

CMIS Report No. CMIS 06/15
February 2006



The datasets have a linear correlation of approximately 0.7 and standard normal marginal distributions but different dependence structures (copulas)

Important Notice

© Copyright Commonwealth Scientific and Industrial Research Organisation ('CSIRO') Australia 2005

All rights are reserved and no part of this publication covered by copyright may be reproduced or copied in any form or by any means except with the written permission of CSIRO.

The results and analyses contained in this Report are based on a number of technical, circumstantial or otherwise specified assumptions and parameters. The user must make its own assessment of the suitability for its use of the information or material contained in or generated from the Report. To the extent permitted by law, CSIRO excludes all liability to any party for expenses, losses, damages and costs arising directly or indirectly from using this Report.

Use of this Report

The use of this Report is subject to the terms on which it was prepared by CSIRO. In particular, the Report may only be used for the following purposes.

- this Report may be copied for distribution within the Client's organisation;
- the information in this Report may be used by the entity for which it was prepared ("the Client"), or by the Client's contractors and agents, for the Client's internal business operations (but not licensing to third parties);
- extracts of the Report distributed for these purposes must clearly note that the extract is part of a larger Report prepared by CSIRO for the Client.

The Report must not be used as a means of endorsement without the prior written consent of CSIRO.

The name, trade mark or logo of CSIRO must not be used without the prior written consent of CSIRO.

Table of Contents

Project Brief

Project Type: Vacation Student Project

Project Title: Dependence Modelling via Copula Method

Project Place:
Quantitative Risk management Group, CSIRO Mathematical and Information Sciences, North Ryde

Project major outcome:
Development of the procedure, algorithm and stand-alone application that allow to fit copula parameters and to chose the best fit copula for given bi-variate time series.

Project Timetable (10 weeks starting at the end of Nov 06 to Feb 06):

- 2 weeks – learning/reading background literature on copulas and goodness-of-fit techniques
- 2 weeks – development of the algorithm and application interface for fitting and goodness-of-fit testing in the case of a standard copula (Gaussian copula)
- 2 weeks – adding extra copulas (t-copula, Clayton, Gumbel) to the algorithm and application.
- 1 week – fitting copulas using bi-variate time series (e.g. A$/USD and Euro/USD)
- 2 weeks – documentation of the results as a technical report
- 1 week – preparation of the end-of-project presentation

Project deliverables:

- stand-alone application with MS Excel interface that allows to fit different copulas to a given bi-variate time series and chose the best fit copula.
- documentation of the procedures and results as a technical report.

Brief description: Accurate modelling of dependence structures via copulas is a topic of current research and recent use in risk management and other areas. Specifying individual (marginal) distributions and linear or rank correlations between the dependent variables is not enough to construct a multivariate joint distribution. In addition, the form of the copula should be specified. The copula method is a tool for constructing non-Gaussian multivariate distributions and understanding the relationships among multivariate data. While the functional forms of many copulas are well known, the fitting procedures and goodness-of-fit tests for copulas are not well-researched topics. The aim of this project is to develop an algorithm and application that will fit different copulas to bivariate time-series and allow to chose the best fit copula.

1. Introduction

Copula *(Lt) : a connexion; a link*

-- Oxford English Dictionary

The phrase ‘copula’ was first used in 1959 by Abe Sklar sixty-seven years ago, but traces of copula theory can be found in Hoeffding’s work during the 1940’s. This theory languished for three decades in the obscurity of theoretical statistics before re-emerging as a key analytical tool in the global scene of financial economics, with particular usefulness in modelling the dependence structure between two sets of random variables. Prior to the very recent explosion of copula theory and application in the financial world, the only models available to represent this dependence structure were the classical multivariate models (such as the ever-present Gaussian multivariate model). These models entailed rigid assumptions on the marginal and joint behaviours of the variables, and were almost useless for modelling the dependence between real financial data.

Copula theory provides a method of modelling the dependence structure between two sets of observations without becoming inextricably tangled in these assumptions. Simply expressed, copulas separate the marginal behaviour of variables from the dependence structure through the use of distribution functions. As the empirical marginal distribution functions can be used instead of their explicit analogues, it is not even necessary to know the exact distribution of the variables being modelled.

Copulas are extremely versatile, and can be used as an analytical tool in a broad range of financial situations such as risk estimation, credit modelling, pricing derivatives, and portfolio management, to name but a few. Although the majority of the vast literature dedicated to copula application lies in the financial sector, the applications of copula theory are not confined to the financial world. Any situation involving more than one random variable can be modelled and analysed using copula theory – although as is usually the case with statistics, the more variables there are present in the model, the more complicated and time-consuming the analysis becomes.

In this paper, we attempt to analyse the dependence structure present between foreign exchange spot rates for the US dollar quoted against the following: the Australian Dollar (denoted $A), the British Pound Sterling (£GB), Swiss Franc (₣Sz), the Euro (€), the Japanese Yen (¥), and the Canadian Dollar ($C). We focus only on the bivariate case due to project time constraints and a difficulty defining the multivariate goodness-of-fit statistics and their distributions.

We first provide some definitions and fundamental framework, then introduce four copulas (Gaussian, Student's t, Clayton, and Gumbel) that will be the key focus in this study. We developed procedures to simulate each copula, estimate the parameter underlying each copula (dependence structure) for a given data set using an approach similar to that of Genest et.al (1995) and Dias et.al (2004), and apply several goodness-of-fit tests to test the appropriateness of these copulas. These procedures were implemented into stand-alone application with MS Excel interface using VBA. Finally, this methodology is applied to the weekly log-returns of the foreign exchange rates mentioned above.

2. Definitions

A copula function is a method of modelling the dependence structure between two sets of paired observations by linking the joint cumulative distribution function (cdf) of a n-dimensional random vector to the marginal cdfs of its components[1].

Technically, a copula is a mapping $C:(0,1)^n \to (0,1)$ from the unit 'hypercube' to the unit interval, with marginals that are uniformly distributed on the interval (0,1). The fundamental theorem underpinning all copula-based analysis is known as Sklar's Theorem

2.1 Sklar's Theorem

Let $F:\Re^n \mapsto (0,1)$ be a joint distribution function with margins X_1, X_2, … , X_n. Then there exists a copula $C:(0,1)^n \to (0,1)$ such that for all $\mathbf{x} \in \Re^n$, $\mathbf{u} \in (0,1)^n$

$$F(\mathbf{x}) = C\{F_1(x_1),...,F_n(x_n)\} = C(\mathbf{u})$$

Conversely, if $C:(0,1)^n \to (0,1)$ is a copula and F_1, …, F_n are distribution functions, then there exists a joint distribution function F with margins F_1, …, F_n such that for all $\mathbf{x} \in \Re^n$, $\mathbf{u} \in (0,1)^n$

$$F(\mathbf{x}) = F\{F_1^{-1}(u_1),...,F_n^{-1}(u_n)\} = C(\mathbf{u})$$

Furthermore, if F, F_1, F_2,,… F_n are continuous, then the copula C is unique.

2.2 Bivariate Copulas

This study considers only the bivariate (two-dimensional) case:

$$C(u,v) = F\{F_X^{-1}(x), F_Y^{-1}(y)\}$$

We consider continuous distributions where F is the joint cdf of the random vector $\mathbf{X} = (X,Y)$ and F_X and F_Y are the marginal cdf's of X and Y respectively. Note particularly that X and Y do not necessarily have the same distribution, and that the joint distribution may differ again; for example, it is quite possible to link a normally-distributed variable and an exponentially-distributed variable together through a bivariate gamma function. Bivariate copulas further satisfy three necessary and sufficent properties (Joe, 1995):

1. $\lim_{u\to 0} C(u,v) = \lim_{v\to 0} C(u,v) = 0$
2. $\lim_{u->1} C(u,v) = v$, $\lim_{v->1} C(u,v) = u$
3. $C(u_1,v_1) - C(u_1,v_2) - C(u_2,v_1) + C(u_2,v_2) \geq 0 \;\; \forall\, (u_1,v_1),(u_2,v_2)$ with $u_1 \leq u_2$, $v_1 \leq v_2$

[1] For some definitions and simple results on statistical distributions, see Appendix I.

2.3 Empirical CDF's and the Copula

If the exact distribution of the marginal distributions is unknown, the empirical cdf can be plugged into the formula instead of the inverse cdf. The empirical cdf of a random variable is defined as follows:

In the univariate case, the empirical cdf of a random variable X evaluated at a point x is defined to be the proportion of observations not greater than x. Mathematically, if the observed values of X are x_1 , x_2 , ... , x_n , then the cdf is given by

$$F(x) = P(X \leq x) = \tfrac{1}{n} \sum_{i=1}^{n} 1(x_i \leq x)$$

A similar expression exists for the bivariate case:

$$F(x, y) = P(X \leq x, Y \leq y) = \tfrac{1}{n} \sum_{i=1}^{n} 1(x_i \leq x, y_i \leq y)$$

The joint empirical cdf of a multivariate distribution can be similarly defined.

3. Copulas Used in this Project

3.1 Elliptic Copulas

The elliptic copulas are a class of symmetric copulas, so-called because the horizontal cross-sections of their bivariate pdf's take the shape of ellipses. Because of their symmetry, a simple linear transformation of variables will transform the elliptic cross-sections to circular ones, producing 'spherical' [marginally uncorrelated] distributions.

We consider two elliptic copulas in this study, both of which are based on extremely common statistical distributions.

3.1.1 Gaussian or Normal Copula

The Gaussian copula is an extension of the normal distribution, one of the most commonly-used (and misused) statistical distributions. In the univariate case, the standard normal distribution is also known as the 'bell curve', and has applications in fields as diverse as mechanical engineering, medicine, pharmaceutical manufacture, and psychology.

The Gaussian copula is given by

$$C_\rho(u,v) = \Phi_\rho\{\Phi^{-1}(u) + \Phi^{-1}(v)\}$$

in which Φ and Φ_ρ are the univariate and bivariate standard normal cdf's repesctively, and ρ is the coefficient of correlation between the random variables X and Y[2].

3.1.2 Student t-Copula with ν degrees of freedom

The Student's t-copula is similar to the Gaussian copula, but it has an extra parameter to control the tail dependence. Small values of ν correspond to greater amounts of probability being contained in the tails of the t-copula, increasing the probability of joint extreme events. As the value of ν increases, the cdf of the Student's t-copula approaches that of the Gaussian. For this reason, the highest value of ν that we consider in this study is $\nu = 20$.

The Student's t-copula with ν degrees of freedom takes the form

$$C_{\rho,\upsilon}(u,v) = t_{\rho,\upsilon}\left\{t_\upsilon^{-1}(u) + t_\upsilon^{-1}(v)\right\}$$

where t_υ and $t_{\rho,\upsilon}$ are respectively the univariate and bivariate standard Student's t cdf's with ν degrees of freedom, and ρ is the coefficient of correlation between the random variables X and Y.

[2] See Appendix I for more details.

3.2 Archimedean Copulas

The Gaussian copula and the Student's t-copula are both suitable for modelling the dependence structure present in symmetric data. However, some financial data such as equity returns (Ang and Chen, 2001, and LInkoln and , 2002) display an asymmetric distribution that cannot be suitably modelled by an elliptic copula. We therefore introduce two asymmetric 'extreme-value' (EV) copulas: the Clayton copula, a left-tailed EV copula, and the Gumbel copula, a right-tailed EV copula.

Both the Clayton and Gumbel copulas are members of the single-parameter Archimedean family of copulas. A copula belongs to the bivariate Archimedean family if it takes the form

$$C(u,v) = \Psi^{-1}\{\, \Psi(u) + \Psi(v) \,\}$$

where $\Psi : (0,1) \mapsto (0,\infty)$ is a **generator function** satisfying three conditions:

1. $\Psi(1) = 0$
2. $\Psi'(s) < 0 \;\; \forall\, s \in (0,1)$, i.e. Ψ is strictly decreasing
3. $\Psi''(1) > 0 \;\; \forall\, s \in (0,1)$, i.e. Ψ is convex

Many such generating functions exist, and thus there are many possible copulas to fit to any dataset. We focus only on two such copulas due to project time constraints.

3.2.1 Clayton Copula

The Clayton copula is a left-tailed EV copula, exhibiting greater dependence in the lower tail than in the upper tail. This copula relies upon a single parameter θ with support in $[0,\infty)$, and takes the form

$$C_\theta(u,v) = \left\{ u^{-\theta} + v^{-\theta} - 1 \right\}^{-1/\theta}$$

with corresponding generator function $\Psi(s) = s^{-\theta} - 1$ and its inverse, $\Psi^{-1}(s) = (s+1)^{-1/\theta}$. Simple differentiation yields a pdf of

$$c_\theta(u,v) = (\theta+1)(uv)^{-\theta-1}\left\{ u^{-\theta} + v^{-\theta} - 1 \right\}^{-2-1/\theta}.$$

The dependence between the observations increases as the value of θ increases, with $\theta \to 0^+$ implying independence and $\theta \to \infty$ implying perfect dependence.

The Clayton copula can also be extended to negative dependence structures, i.e. those with negative linear correlation.

3.2.2 Gumbel Copula

The Gumbel copula exhibits upper tail dependence, indicating that large positive joint extreme events are more likely to occur than large negative ones. It incorporates a single parameter θ with support in $[1,\infty)$, and is written as

$$C_\theta(u,v) = \exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{-1/\theta}\right\}$$

with an associated generator function of $\Psi(s) = (-\ln s)^\theta$ and its inverse $\Psi^{-1}(s) = \exp\left(-s^{1/\theta}\right)$. Setting $\bar{u} = -\ln u$ and $\bar{v} = -\ln v$, the pdf of the Gumbel copula can be expressed thus:

$$c_\theta(u,v) = C_\theta(u,v)\,(uv)^{-1}(\bar{u}\bar{v})^{\theta-1}\left(\bar{u}^\theta + \bar{v}^\theta\right)^{-2+1/\theta}\left\{\left(\bar{u}^\theta + \bar{v}^\theta\right)^{1/\theta} + \theta - 1\right\}.$$

As with the Clayton copula, the dependence between the observations increases with θ, implying independence as $\theta \to 1^+$ and perfect dependence as $\theta \to \infty$. Similarly, an extension to negative dependence is possible.

3.2.3 Other Archimedean Copulas

There are many other type of Archimedean copulas that have been proposed in the last few decades. Some of them have simple closed forms, and can be easily simulated, while others are more complex.

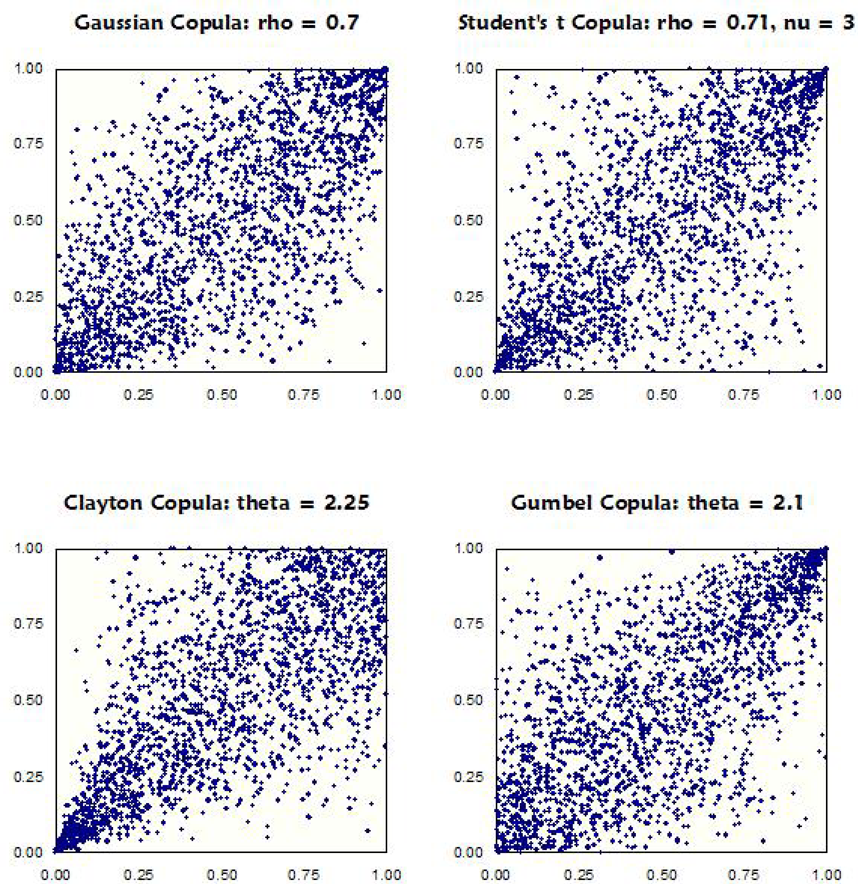
1. Independent Copula

   $C(u,v) = uv$; $\Psi(s) = \ln s$

2. Frank Copula

   $C(u,v) = -\dfrac{1}{\theta}\ln\left[1 - \dfrac{\left(1-e^{-\theta u}\right)\left(1-e^{-\theta v}\right)}{1-e^{-\theta}}\right]$; $\Psi(s) = -\dfrac{1}{\theta}\ln\left[1 - \left(1-e^{-\theta}\right)e^{-s}\right]$

The Frank Copula is symmetric, and has support for θ in $(-\infty,\infty)$.



All the datasets above have a linear correlation coefficient of 0.7 and uniform margins; however, the dependence structure displayed by each dataset is clearly different.

4. Simulation Toolkit

4.1 Elliptical Copulas

Given two independent random variables $U \sim \mathrm{U}(0,1)$ and $V \sim \mathrm{U}(0,1)$, we can set $X_1 = F^{-1}(U)$ and $X_2 = F^{-1}(V)$, where F(.) is the marginal cdf of the elliptical distribution used to construct copula (e.g. F(.) is the standard Normal distribution for Gaussian copula, F(.) is t-distribution with ν degrees of freedom for t copula with ν degrees of freedom). Since the copulas are elliptical, we can transform these variables by a rotation about the origin: This rotation factor is the linear correlation coefficient, ρ.

$$X_1 = X_1$$
$$X_2 = \rho X_1 + \sqrt{1-\rho^2}\, X_2$$

We then transform back into uniform variates by applying the CDF:

$$U_1 = F(X_1)$$
$$U_2 = F(X_2)$$

U_1 and U_2 are thus simulated from the elliptical copula with correlation coefficient ρ.

This procedure has been used for Gaussian and t copulas.

Another procedure can be used to simulate t-copula as follows:

1. Simulate independent U and V from U(0,1)

2. calculate $X_1 = F^{-1}(U)$ and $X_2 = F^{-1}(V)$ where F(.) is standard Normal distribnution

3. perform transformation: $X_1 = X_1; \quad X_2 = \rho X_1 + \sqrt{1-\rho^2}\, X_2$

4. find $X_i = \sqrt{\nu} X_i / \sqrt{S}, i = 1,2$ where S is simulated independently from the Chi-Square distribution with ν degrees of freedom

5. calculate $U_i = F_t(X_i)$ where $F_t(.)$ is t distribution with ν degrees of freedom.

4.2 Archimedean Copulas

Method 1

Given two independent random variables $U \sim \mathrm{U}(0,1)$ and $V \sim \mathrm{U}(0,1)$, we can set $U = C(S \mid V)$ independently of V. So, we can generate two random variables u and v that are uniformly distributed on (0,1), and set

$$s = C^{-1}(u \mid v)\ , \text{ where } C(u \mid v) = \frac{\partial C(u, v)}{\partial v}$$

Thus U and S are simulated from the desired copula. Often, the Archimedean copulas will have no closed form and this step will require numerical root finding until a required precision of convergence is reached.

Method 2

The distribution function of the Archimedean copula being simulated can be expressed as

$$K(t) = t - \frac{\Psi(t)}{\Psi'(t)}$$

where $\Psi(t)$, $t \in (0,1)$ is the generating function of the copula being simulated.

We can thus generate two random variables u and v that are uniformly distributed on (0,1), and setting $u = K(t)$, we can generate $q = K^{-1}(u)$ by numerical root finding. Then, we set

$$s = \Psi^{-1}[v\ \Psi(q)]$$

$$t = \Psi^{-1}[(1 - v)\ \Psi(q)]$$

The variables s and t are thus uniformly distributed on the interval (0,1) and have the required dependence structure.

5. Estimation of Parameters

5.1 Maximum Likelihood Estimation

Define the likelihood of a set of n observations of (X, Y) as

$$L(\theta) = \prod_{i=1}^{n} f_\theta(x_i, y_i)$$

where $f_\theta(x, y)$ is the pdf of the random vector $(X, Y)^T$ evaluated at the point (x, y). Parameter θ is estimated by maximising this likelihood with respect to θ.

Since the logarithmic transform is a monotonic increasing function, maximising the log-likelihood will provide the same optimal estimate for θ as maximising the original likelihood whilst reducing the complexity of calculations.

$$\ln L(\hat{\theta}) = \ln\left(\prod_{i=1}^{n} f_{\hat{\theta}}(x_i, y_i)\right) = \sum_{i=1}^{n} \ln\left(f_{\hat{\theta}}(x_i, y_i)\right)$$

We therefore take the natural logarithm of the likelihood and maximise it by an iterative process, extending the estimate of the parameter by one decimal place each time until an accuracy of twelve decimal places (the limit of the Excel spreadsheet) is reached.

5.2 Standard Error

The variance of the parameter estimate $\hat{\theta}$ can be calculated from the Fisher Information Matrix. Since all our models have only one parameter (in the Student's t-distribution, ν is regarded as fixed), the formula simplifies to:

$$I = -E\left[\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2} \ln f(X)\bigg|_{\theta=\hat{\theta}}\right] = \frac{1}{Var(\hat{\theta})}$$

where $f(X)$ is the pdf of the random variable X.

Calculation of the standard error was performed by an iterative routine:

1. For each pair of observations, calculate the component of variance contributed to the total Fisher Information by substituting the observed values of x_i and y_i, as well as the estimated value of $\hat{\theta}$ into the second derivative function.

2. Since the standard error is the square root of the variance, set

$$SE(\hat{\theta}) = \sqrt{\left|Var(\hat{\theta})\right|}\ .$$

6. Goodness-of-Fit Tests

6.1 Pearson's Chi-Squared

The chi-squared test relies upon an arbitrary binning of data to calculate the test statistic. In this paper, we have used equal probability binning in all procedures to minimise error. Furthermore, all data is transformed prior to binning into uniform variates that are uncorrelated under the null hypothesis – that is, our null hypothesis is $\overline{C}_\theta(u, v) = uv$.

The classical Pearson's χ^2-statistic is then calculated according to the following equation:

$$\chi^2 = \sum_{i=1}^{n} \frac{\left(N_i^{obs} - N_i^{\exp} \right)^2}{N_i^{\exp}}$$

where $N_i^{obs}, N_i^{\exp}$ are the number of observed and expected observations in the ith bin respectively. To calculate the p-value we simply note that the degrees of freedom are simply the number of bins that are not empty, and thus obtain the p-value from statistical tables.

While the chi-squared test is the most commonly-used statistical test for goodness-of-fit, it should be noted that the binning is entirely arbitrary and can produce different p-values depending upon which binning procedure is used. With this in mind, we make use of five different binning procedures for chi-squared testing in this paper.

1. Sectors

   Each pair of observations is allocated to one of eight bins depending upon the sign of $u_i - 0.5$ and the ratio $r_i = \dfrac{v_i - 0.5}{u_i - 0.5}$. This measures clustering around specific values of *r*.

2. Concentric Squares

   Each pair of observations is placed into one of ten bins centred at the point $(0.5, 0.5)$ as shown below. This measures clustering around the centre of the dataset.

3. Quadrants

   Taking the 'origin' to be the point $(0.5, 0.5)$, each pair of observations is allocated to one of four bins depending upon which quadrant that pair lies in.

4. 3x3 Squares

   Each pair of observations is allocated to one of nine bins as shown below.

5. 4x4 Squares

Each pair of observations is allocated to one of sixteen bins as shown below.

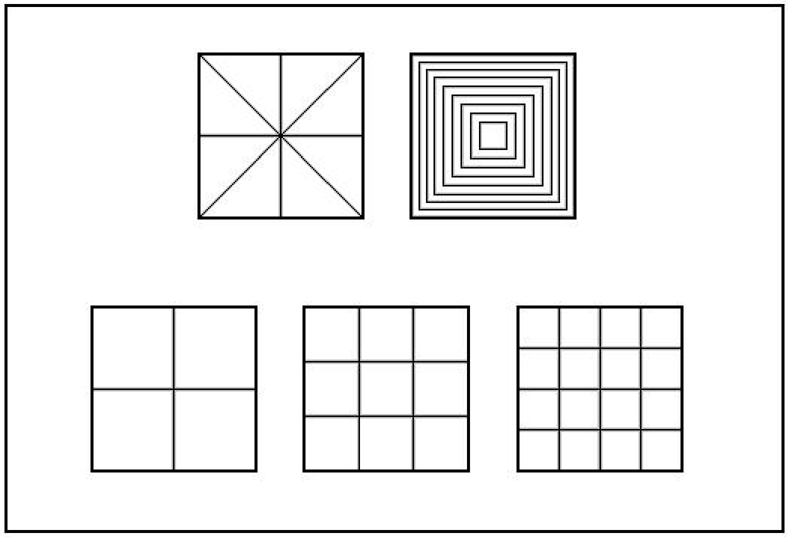


Figure 2: The binning procedures used in this study for chi-squared goodness-of-fit of copulas.

However, to balance out the limitations imposed upon the study by arbitrary binning, we also performed two tests for equality of continuous distributions, suitably modified for the discrete step-function exhibited by an empirical joint cdf.

6.2 Kolmogorov-Smirnov

The Kolmogorov-Smirnov test for equality of cumulative distribution functions is one of the most widely used tests for equality of distributions. Although the Kolmogorov-Smirnov test is less powerful than tests such as the Shapiro-Wilk's test and the Anderson-Darling test, it remains easy to implement and calculate the p-value associated with an observed value of the statistic as the distribution is well-known - unlike the Anderson-Darling test statistic, below.

Let $Y_{(i)}$ be the ordered data - then the Kolmogorov-Smirnov test statistic is defined as

$$D^* = \max_{1 \le i \le n} \left\{ \left| F(Y_{(i)}) - \tfrac{i-1}{n} \right|, \left| F(Y_{(i)}) - \tfrac{i}{n} \right| \right\}$$

The p-value associated with this statistic can be easily calculated by evaluating the sum

$$1\text{-}Q_{KS} = 1 - 2\sum_{i=1}^{\infty} \left(-1\right)^n e^{-2j^2\lambda^2} \quad \text{where} \quad \lambda = \left(\sqrt{n} + 0.11 + \frac{0.12}{\sqrt{n}}\right) D^*$$

until it converges to an acceptable degree of precision.

We apply the Kolmogorov-Smirnov test to both the fitted cdf and to the uniform variate obtained during transformation of the data. Only one marginal cdf needs to be tested, as the other marginal is simply the empirical cdf of the first original marginal, and therefore is already uniformly distributed.

6.3 Anderson-Darling

The Anderson-Darling test is widely used in the literature for testing normality, although tables of critical values have been developed for the exponential, Weibull, and log-normal distributions, as well as some others. However, no function similar to the Kolmogorov-Smirnov function has yet been found, and the p-values associated with a statistic are usually generated through the technique of Monte Carlo simulation.

The Anderson-Darling test is technically an integral, but since the empirical cdf is a step function we can express this integral as a finite discrete sum. The discrete form of the Anderson-Darling test statistic is

$$A^2 = -N - \sum_{i=1}^{N} \frac{2i-1}{N} \left\{ \ln F\left(Y_{(i)}\right) + \ln\left[1 - F\left(Y_{(n+1-i)}\right) \right] \right\}$$

Unfortunately, there was insufficient time allocated to this project to allow a Monte Carlo simulation of the critical values for the estimated value of the parameters in each copula. The Anderson-Darling test statistics are useful only as a rough estimate only, and cannot be used to determine whether a copula is significantly different from the empirical copula.

6.4 p-values

In general, p-values for all of the above tests should be calculated numerically via a Monte Carlo simulation procedure because the copula parameter is estimated from the data and thus the distribution of the test statistic does not have a closed form. Given the time limits of the project, the p-values for the Kolmogorov-Smirnov and chi-squared tests were calculated using the Kolmogorov-Smirnov and chi-squared distributions.

7. Testing Simulated Data

To verify simulation and testing procedures, we performed numerous simulation tests.

- 10,000 data points from each copula were simulated for a specified parameter value– for example, a Gaussian copula with correlation coefficient ρ=0.7 was simulated. The maximum likelihood procedures and goodness-of-fit tests were run to test each copula.

This was repeated for five different values of the parameter for each copula – for example, with the Gaussian copula, specified values of the correlation coefficitent were -0.6, -0.2, 0.3, 0.7, and 0.95. In each case, the testing routine rejected all copulas but the one that was simulated, verifying the estimation and testing procedures.

8. Application to Financial Data

In this paper, we attempt to use the copula method to analyse the dependence structure present between foreign exchange spot rates for the US dollar quoted against the following: the Australian Dollar (denoted $A), the British Pound Sterling (£GB), Swiss Franc (₣Sz), the Euro (€), the Japanese Yen (¥), and the Canadian Dollar ($C).

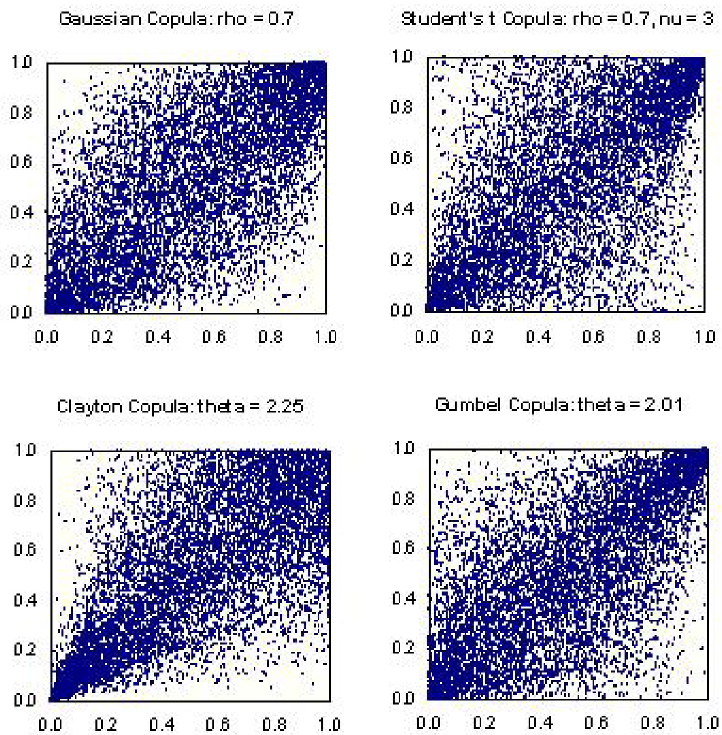
We obtained foreign exchange rate data from the New York stock exchange data provided by Yahoo! Finance, taking daily data from 1996 to 2005. This data was then reduced to weekly data to eliminate seasonality. The day chosen to take observations on was Wednesday, as fewer public holidays fell on Wednesdays than on any other day of the week in this time period. Whenever a public holiday did fall on a Wednesday, the data from the Thursday immediately after was substituted instead. This substitution occurred only four times out of 522 data observations.

The weekly data was then transformed into log-returns:

$$Y_i = \log\left(\frac{X_{i+1}}{X_i}\right)$$

We then selected a particular pair of data observations (there were 15 pairs in all) and for each pair, we performed the following procedure:

1. The log-returns were transformed to [univariate] empirical cdf's.
2. For each copula, the second vector of observations was transformed to a vector that should be uncorrelated with the first vector, provided the copula choice is correct. We made this transformation by setting
$$s = C(u \mid v) = \frac{\partial C(u,v)}{\partial v}$$
where u and v are the observed values of the empirical cdf's, and s is also uniformly distributed on the interval (0,1), but is uncorrelated with v.
3. We calculated the joint empirical cdf for a specified pair (x_i, y_i) of these modified data using the definition from Section 2.3:
$$F(x_i, y_i) = P(X < x_i, Y < y_i)$$
4. Each of the four copulas focused on in this study were fitted by maximum likelihood estimation and the goodness-of-fit testing routines were applied as described above to obtain a set of p-values for each test (or in the case of the Anderson-Darling test, a test statistic).



Disappointingly, the results were inconclusive. For most of the pairs of exchange rates, none of the copula models could be rejected at the 0.05 significance level. However, some preliminary conclusions and hypotheses could be drawn from the data based upon a comparison of the sets of p-values. In particular, the pairings involving the Canadian Dollar and the Japanese Yen showed complex behaviour that requires further investigation, perhaps due to their geographical isolation from the other currencies.

- The only significant rejection occurred for the pairing of the Swiss Franc and the Euro, with the extreme-value copulas performing very badly in comparison with the Gaussian and t-copulas.
- The dependence between the Australian Dollar and the Euro appears to be a t-copula with a low number (5-8) of degrees of freedom. The dependence between the Australian Dollar and the Swiss Franc appears similar.
- The relationship between the Australian Dollar and the Canadian Dollar may be a mixture model of the Gaussian and Gumbel copulas, as may the relationship between the British Pound and the Canadian Dollar.
- The British Pound appears to have a Gaussian dependence structure with the Euro, the Swiss Franc, and the Australian Dollar.

The p-values (and test statistic values for the Anderson-Darling test) are shown on the next few pages. The p-values that are significant at the 0.05 level are highlighted in red.

$A vs £GB, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.348	0.414	1.266	0.294	0.312	0.323	0.329	0.334	0.337	0.339	0.340	0.345	0.346
	std error	0.001	0.008		0.011	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.880	0.628	0.669	0.163	0.331	0.387	0.428	0.467	0.639	0.669	0.669	0.739	0.800
	Circles	0.880	0.534	0.985	0.932	0.920	0.936	0.936	0.946	0.903	0.874	0.905	0.791	0.846
	Quadrants	0.998	0.811	0.998	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	3x3 Squares	0.978	0.328	0.981	0.772	0.916	0.954	0.976	0.983	0.983	0.978	0.978	0.978	0.978
	4x4 Squares	0.697	0.161	0.941	0.505	0.473	0.528	0.496	0.589	0.632	0.589	0.646	0.646	0.646
Kolmogorov-Smirnov	**U2**	0.774	0.531	0.925	0.695	0.699	0.747	0.722	0.724	0.738	0.734	0.732	0.794	0.794
	Copula	0.954	0.623	0.893	0.893	0.956	0.960	0.960	0.960	0.960	0.959	0.959	0.959	0.958
Anderson-Darling	**U2 stat.**	0.395	0.304	0.328	0.372	0.384	0.394	0.400	0.404	0.408	0.411	0.412	0.409	0.406
	Copula stat.	124.7	132.9	129.6	140.4	136.3	133.7	131.9	130.7	129.8	129.1	128.6	127.0	126.3

Table 8.1 The p-values and test statistic values obtained for modelling the Australian Dollar ($A) against the British Pound (£GB)

$A vs FSz, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.333	0.408	1.267	0.315	0.329	0.337	0.341	0.343	0.344	0.345	0.345	0.345	0.344
	std error	0.001	0.008		0.011	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.699	0.178	0.229	0.421	0.800	0.883	0.912	0.912	0.907	0.907	0.907	0.856	0.807
	Circles	0.665	0.696	0.526	0.865	0.820	0.754	0.773	0.739	0.617	0.649	0.677	0.688	0.731
	Quadrants	0.354	0.203	0.345	0.863	0.789	0.638	0.638	0.638	0.638	0.638	0.638	0.505	0.439
	3x3 Squares	0.818	0.120	0.821	0.916	0.916	0.941	0.959	0.950	0.950	0.923	0.923	0.923	0.908
	4x4 Squares	0.797	0.317	0.636	0.542	0.627	0.646	0.688	0.751	0.751	0.751	0.776	0.742	0.711
Kolmogorov-Smirnov	**U2**	0.961	0.922	0.790	0.970	0.965	0.939	0.930	0.929	0.933	0.956	0.968	0.973	0.953
	Copula	0.983	0.699	0.891	0.992	0.995	0.995	0.992	0.989	0.987	0.985	0.984	0.981	0.981
Anderson-Darling	**U2 stat.**	0.298	0.212	0.374	0.214	0.219	0.236	0.250	0.258	0.265	0.269	0.271	0.278	0.281
	Copula stat.	124.9	131.4	127.9	136.5	132.7	130.4	128.9	127.9	127.2	126.7	126.3	125.3	125.0

Table 8.2 The p-values and test statistic values obtained for modelling the Australian Dollar ($A) against the Swiss Franc (FSz)

$A vs €, 1999-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.485	0.741	1.445	0.471	0.488	0.497	0.501	0.503	0.505	0.505	0.505	0.504	0.501
	std error	0.001	0.018		0.013	0.008	0.006	0.005	0.004	0.003	0.003	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.504	0.316	0.142	0.351	0.437	0.607	0.666	0.781	0.677	0.677	0.677	0.576	0.618
	Circles	0.725	0.921	0.747	0.943	0.954	0.981	0.875	0.911	0.839	0.810	0.810	0.879	0.834
	Quadrants	0.255	0.097	0.285	0.699	0.615	0.615	0.615	0.615	0.444	0.444	0.444	0.365	0.355
	3x3 Squares	0.936	0.257	0.819	0.999	0.995	0.987	0.971	0.971	0.906	0.930	0.909	0.895	0.920
	4x4 Squares	0.752	0.039	0.628	0.765	0.844	0.839	0.839	0.823	0.833	0.828	0.828	0.812	0.839
Kolmogorov-Smirnov	**U2**	0.695	0.622	0.410	0.856	0.790	0.809	0.906	0.903	0.843	0.800	0.806	0.799	0.804
	Copula	0.672	0.397	0.523	0.705	0.765	0.760	0.754	0.747	0.741	0.736	0.731	0.714	0.704
Anderson-Darling	**U2 stat.**	0.689	0.531	0.831	0.222	0.294	0.360	0.410	0.447	0.476	0.498	0.515	0.565	0.591
	Copula stat.	63.6	67.6	66.6	73.2	69.9	67.9	66.7	65.9	65.3	64.9	64.6	63.8	63.6

Table 8.3 The p-values and test statistic values obtained for modelling the Australian Dollar ($A) against the Euro (€)

$A vs ¥, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.288	0.328	1.207	0.248	0.262	0.270	0.275	0.279	0.281	0.282	0.284	0.286	0.287
	std error	0.001	0.006		0.011	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.936	0.915	0.837	0.830	0.924	0.899	0.883	0.891	0.915	0.915	0.915	0.904	0.922
	Circles	0.989	0.802	0.872	0.892	0.960	0.976	0.983	0.971	0.944	0.957	0.955	0.985	0.985
	Quadrants	0.752	0.885	0.723	0.752	0.730	0.666	0.666	0.687	0.687	0.687	0.687	0.687	0.687
	3x3 Squares	0.303	0.231	0.586	0.570	0.526	0.478	0.478	0.478	0.478	0.478	0.420	0.420	0.295
	4x4 Squares	0.746	0.262	0.416	0.565	0.608	0.598	0.551	0.738	0.738	0.733	0.733	0.733	0.733
Kolmogorov-Smirnov	**U2**	0.790	0.995	0.858	0.858	0.818	0.740	0.734	0.731	0.728	0.726	0.725	0.769	0.791
	Copula	0.810	0.689	0.791	0.812	0.812	0.812	0.812	0.812	0.812	0.812	0.812	0.812	0.811
Anderson-Darling	**U2 stat.**	0.286	0.144	0.375	0.262	0.280	0.285	0.291	0.294	0.295	0.296	0.295	0.293	0.292
	Copula stat.	137.7	144.6	142.6	150.8	147.3	145.1	143.7	142.6	141.9	141.3	140.9	139.6	139.1

Table 8.4 The p-values and test statistic values obtained for modelling the Australian Dollar ($A) against the Japanese Yen (¥)

$A vs $C, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.446	0.593	1.359	0.387	0.408	0.420	0.427	0.432	0.435	0.438	0.439	0.444	0.445
	std error	0.001	0.012		0.011	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.912	0.259	0.178	0.176	0.328	0.561	0.743	0.750	0.853	0.859	0.833	0.846	0.837
	Circles	0.507	0.903	0.954	0.977	0.816	0.688	0.573	0.519	0.393	0.380	0.370	0.370	0.500
	Quadrants	0.998	0.122	0.877	1.000	1.000	0.998	0.998	0.998	0.998	1.000	1.000	0.998	0.998
	3x3 Squares	0.959	0.021	0.945	0.970	0.975	0.971	0.964	0.975	0.973	0.973	0.973	0.973	0.967
	4x4 Squares	0.907	0.010	0.821	0.828	0.688	0.746	0.751	0.764	0.764	0.843	0.817	0.893	0.893
Kolmogorov-Smirnov	**U2**	0.963	0.427	0.891	0.985	0.961	0.963	0.978	0.968	0.973	0.967	0.969	0.966	0.951
	Copula	0.946	0.335	0.769	0.710	0.850	0.918	0.951	0.968	0.978	0.980	0.979	0.974	0.966
Anderson-Darling	**U2 stat.**	0.208	0.578	0.288	0.150	0.141	0.146	0.159	0.169	0.177	0.183	0.187	0.196	0.199
	Copula stat.	102.0	110.5	109.4	119.4	114.7	111.8	109.9	108.5	107.5	106.7	106.2	104.5	103.8

Table 8.5 The p-values and test statistic values obtained for modelling the Australian Dollar ($A) against the Canadian Dollar ($C)

£GB vs FSz, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.623	1.029	1.677	0.590	0.608	0.617	0.622	0.625	0.627	0.629	0.629	0.630	0.630
	std error	0.000	0.022		0.010	0.006	0.005	0.004	0.003	0.002	0.002	0.002	0.001	0.001
Chi-Squared Tests	**Spokes**	0.412	0.011	0.014	0.013	0.079	0.126	0.302	0.364	0.387	0.376	0.394	0.454	0.454
	Circles	0.948	0.692	0.908	0.936	0.895	0.944	0.959	0.944	0.960	0.948	0.950	0.965	0.964
	Quadrants	0.632	0.061	0.161	0.995	0.988	0.963	0.963	0.963	0.945	0.945	0.919	0.863	0.863
	3x3 Squares	0.979	0.010	0.735	0.983	0.991	0.992	0.992	0.995	0.995	0.995	0.995	0.991	0.991
	4x4 Squares	0.742	0.000	0.285	0.109	0.278	0.468	0.641	0.781	0.871	0.864	0.847	0.913	0.817
Kolmogorov-Smirnov	**U2**	0.357	0.488	0.096	0.713	0.540	0.520	0.535	0.509	0.428	0.419	0.414	0.394	0.346
	Copula	0.912	0.192	0.600	0.644	0.821	0.879	0.908	0.924	0.933	0.938	0.941	0.946	0.944
Anderson-Darling	**U2 stat.**	0.726	0.504	1.420	0.379	0.405	0.423	0.436	0.444	0.454	0.469	0.486	0.551	0.585
	Copula stat.	64.3	73.6	69.5	81.5	76.3	73.2	71.2	69.8	68.9	68.1	67.6	66.1	65.5

Table 8.6 The p-values and test statistic values obtained for modelling the British Pound (£GB) against the Swiss Franc (FSz))

£GB vs €, 1999-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.700	1.268	1.865	0.655	0.675	0.685	0.692	0.696	0.699	0.701	0.702	0.705	0.706
	std error	0.000	0.034		0.012	0.008	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	Spokes	0.853	0.000	0.189	0.235	0.351	0.470	0.645	0.645	0.645	0.645	0.745	0.761	0.761
	Circles	0.914	0.891	0.973	0.453	0.805	0.895	0.993	0.987	0.969	0.980	0.967	0.934	0.967
	Quadrants	0.945	0.068	0.777	0.978	0.974	0.974	0.974	0.974	0.974	0.974	0.945	0.945	0.945
	3x3 Squares	0.764	0.009	0.743	0.694	0.721	0.589	0.628	0.726	0.727	0.721	0.758	0.689	0.705
	4x4 Squares	0.992	0.001	0.740	0.771	0.817	0.926	0.953	0.995	0.995	0.991	0.988	0.994	0.994
Kolmogorov-Smirnov	U2	0.840	0.488	0.573	0.948	0.947	0.957	0.967	0.950	0.933	0.944	0.952	0.950	0.966
	Copula	0.910	0.127	0.889	0.729	0.891	0.921	0.921	0.921	0.921	0.862	0.863	0.871	0.878
Anderson-Darling	U2 stat.	0.399	0.700	1.098	0.295	0.313	0.315	0.308	0.302	0.302	0.305	0.308	0.312	0.317
	Copula stat.	35.1	42.4	38.8	48.3	44.4	42.1	40.6	39.5	38.8	32.1	32.3	32.9	33.3

Table 8.7 The p-values and test statistic values obtained for modelling the British Pound (£GB) against the Euro (€)

£GB vs ¥, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom									
					3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.278	0.338	1.188	0.247	0.258	0.264	0.268	0.270	0.272	0.273	0.274	0.277	0.277
	std error	0.001	0.006		0.012	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	Spokes	0.800	0.967	0.598	0.978	0.933	0.917	0.917	0.922	0.922	0.917	0.907	0.833	0.833
	Circles	0.979	0.995	0.927	0.880	0.895	0.886	0.908	0.903	0.942	0.938	0.973	0.962	0.981
	Quadrants	0.963	0.926	0.841	0.974	0.979	0.974	0.974	0.988	0.988	0.974	0.951	0.939	0.939
	3x3 Squares	1.000	0.916	0.903	1.000	1.000	1.000	0.998	0.999	0.999	0.999	0.999	0.999	0.999
	4x4 Squares	1.000	0.893	0.943	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	0.999
Kolmogorov-Smirnov	U2	0.994	0.957	0.877	0.988	0.992	0.991	0.994	0.994	0.995	0.996	0.988	0.990	0.989
	Copula	0.963	0.836	0.815	0.979	0.991	0.991	0.990	0.989	0.987	0.983	0.981	0.970	0.968
Anderson-Darling	U2 stat.	0.136	0.188	0.205	0.153	0.147	0.139	0.134	0.132	0.131	0.131	0.131	0.132	0.133
	Copula stat.	138.0	142.7	144.0	149.9	146.8	144.9	143.6	142.7	142.0	141.5	141.1	139.9	139.4

Table 8.8 The p-values and test statistic values obtained for modelling the British Pound (£GB) against the Japanese Yen (¥)

£GB vs $C, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.208	0.229	1.131	0.176	0.187	0.193	0.197	0.199	0.201	0.202	0.203	0.206	0.207
	std error	0.001	0.004		0.012	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	Spokes	0.990	0.617	0.999	0.999	1.000	1.000	1.000	1.000	1.000	0.999	0.996	0.990	0.990
	Circles	0.515	0.613	0.903	0.777	0.496	0.565	0.565	0.526	0.550	0.554	0.573	0.565	0.538
	Quadrants	0.974	0.687	0.963	1.000	1.000	1.000	0.998	0.988	0.988	0.988	0.988	0.974	0.974
	3x3 Squares	0.351	0.447	0.616	0.322	0.322	0.337	0.311	0.348	0.382	0.385	0.348	0.351	0.351
	4x4 Squares	0.982	0.692	0.995	0.983	0.985	0.986	0.980	0.976	0.976	0.976	0.976	0.956	0.978
Kolmogorov-Smirnov	U2	0.907	0.956	0.893	0.885	0.881	0.901	0.921	0.917	0.908	0.902	0.898	0.897	0.902
	Copula	0.900	0.914	0.848	0.916	0.916	0.916	0.915	0.915	0.914	0.913	0.913	0.910	0.908
Anderson-Darling	U2 stat.	0.202	0.188	0.168	0.214	0.207	0.205	0.204	0.204	0.205	0.205	0.206	0.206	0.206
	Copula stat.	155.5	160.3	161.0	165.9	163.3	161.6	160.5	159.7	159.1	158.6	158.2	157.2	156.7

Table 8.9 The p-values and test statistic values obtained for modelling the British Pound (£GB) against the Canadian Dollar ($C)

FSz vs €, 1999-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.942	4.582	4.091	0.933	0.937	0.939	0.941	0.941	0.942	0.942	0.942	0.942	0.943
	std error	0.000	0.151		0.004	0.002	0.002	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Chi-Squared Tests	Spokes	0.591	0.000	0.031	0.545	0.623	0.534	0.645	0.634	0.634	0.634	0.719	0.581	0.581
	Circles	0.398	0.002	0.131	0.192	0.388	0.256	0.473	0.495	0.432	0.500	0.432	0.538	0.489
	Quadrants	0.999	0.223	0.990	0.990	0.990	0.981	0.990	0.990	0.990	0.990	0.997	0.999	0.999
	3x3 Squares	0.913	0.000	0.005	0.243	0.243	0.419	0.519	0.611	0.711	0.694	0.764	0.870	0.834
	4x4 Squares	0.873	0.000	0.002	0.031	0.406	0.461	0.507	0.614	0.607	0.648	0.765	0.839	0.844
Kolmogorov-Smirnov	U2	0.305	0.045	0.487	0.986	0.971	0.914	0.914	0.899	0.823	0.784	0.729	0.487	0.424
	Copula	0.999	0.315	0.991	0.964	0.994	0.997	0.999	0.999	0.999	0.999	0.999	0.999	0.999
Anderson-Darling	U2 stat.	0.891	2.563	1.161	0.268	0.289	0.311	0.343	0.371	0.397	0.418	0.437	0.518	0.575
	Copula stat.	6.2	9.7	7.4	5.2	5.4	5.5	5.6	5.6	5.7	5.7	5.8	5.9	6.0

Table 8.10 The p-values and test statistic values obtained for modelling the Swiss Franc (FSz) against the Euro (€)

FSz vs ¥, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.360	0.431	1.291	0.323	0.340	0.350	0.356	0.359	0.362	0.364	0.365	0.367	0.367
	std error	0.001	0.008		0.011	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.993	0.580	0.750	0.441	0.710	0.856	0.907	0.938	0.950	0.964	0.973	0.973	0.989
	Circles	0.880	0.946	0.960	0.850	0.720	0.665	0.795	0.830	0.837	0.859	0.889	0.895	0.883
	Quadrants	0.995	0.826	0.974	0.979	0.988	0.995	0.995	0.995	0.998	1.000	1.000	1.000	0.998
	3x3 Squares	0.973	0.128	0.811	0.866	0.921	0.939	0.964	0.986	0.985	0.982	0.982	0.982	0.982
	4x4 Squares	0.751	0.142	0.627	0.288	0.433	0.451	0.473	0.500	0.519	0.598	0.669	0.688	0.751
Kolmogorov-Smirnov	**U2**	0.994	0.783	0.993	0.994	0.995	0.988	0.992	0.977	0.981	0.983	0.981	0.982	0.991
	Copula	0.983	0.523	0.880	0.957	0.989	0.993	0.993	0.992	0.992	0.991	0.990	0.979	0.972
Anderson-Darling	**U2 stat.**	0.179	0.273	0.232	0.186	0.184	0.182	0.185	0.182	0.179	0.176	0.175	0.173	0.172
	Copula stat.	120.1	128.5	123.9	134.2	130.0	127.4	125.8	124.6	123.7	123.1	122.6	121.3	120.8

Table 8.11 The p-values and test statistic values obtained for modelling the Swiss Franc (FSz) against the Japanese Yen (¥)

FSz vs $C, 1996-2005

		Normal	Clayton	Gumbel	Student's t degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	**rho / theta**	0.226	0.238	1.159	0.197	0.209	0.216	0.220	0.223	0.224	0.225	0.226	0.228	0.228
	std error	0.001	0.004		0.012	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	**Spokes**	0.491	0.174	0.661	0.280	0.344	0.302	0.358	0.358	0.358	0.358	0.358	0.358	0.358
	Circles	0.977	0.968	0.992	0.964	0.938	0.972	0.950	0.976	0.964	0.971	0.971	0.982	0.986
	Quadrants	0.926	0.477	0.752	0.926	0.951	0.939	0.919	0.919	0.919	0.919	0.919	0.919	0.919
	3x3 Squares	0.872	0.427	0.928	0.808	0.808	0.808	0.808	0.808	0.825	0.825	0.866	0.866	0.866
	4x4 Squares	0.982	0.482	0.987	0.968	0.975	0.978	0.980	0.980	0.980	0.980	0.980	0.980	0.980
Kolmogorov-Smirnov	**U2**	0.919	0.968	0.995	0.898	0.886	0.938	0.926	0.922	0.918	0.916	0.921	0.911	0.911
	Copula	0.810	0.696	0.687	0.832	0.851	0.848	0.846	0.843	0.841	0.838	0.836	0.829	0.825
Anderson-Darling	**U2 stat.**	0.187	0.171	0.131	0.213	0.205	0.205	0.206	0.206	0.205	0.204	0.204	0.200	0.197
	Copula stat.	151.2	157.8	154.8	161.6	158.7	156.9	155.6	154.8	154.2	153.7	153.4	152.4	152.0

Table 8.12 The p-values and test statistic values obtained for modelling the Swiss Franc (FSz) against the Canadian Dollar ($C)

€ vs ¥, 1999-2005

		Normal	Clayton	Gumbel	Student's t, degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.299	0.410	1.238	0.302	0.313	0.318	0.320	0.321	0.321	0.321	0.321	0.318	0.315
	std error	0.001	0.009		0.014	0.008	0.006	0.005	0.004	0.003	0.003	0.002	0.002	0.002
Chi-Squared Tests	Spokes	0.089	0.312	0.038	0.397	0.301	0.257	0.244	0.244	0.179	0.179	0.179	0.161	0.161
	Circles	0.365	0.914	0.479	0.533	0.468	0.489	0.370	0.379	0.338	0.309	0.281	0.252	0.356
	Quadrants	0.974	0.846	0.903	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999
	3x3 Squares	0.983	0.838	0.758	0.977	0.994	0.996	0.999	0.999	0.996	0.995	0.995	0.996	0.995
	4x4 Squares	0.992	0.668	0.839	0.993	0.994	0.990	0.986	0.978	0.978	0.978	0.978	0.981	0.985
Kolmogorov-Smirnov	U2	0.759	0.793	0.720	0.674	0.707	0.713	0.731	0.720	0.726	0.728	0.729	0.828	0.762
	Copula	0.984	0.956	0.852	0.992	0.998	0.996	0.994	0.994	0.993	0.993	0.993	0.990	0.989
Anderson-Darling	U2 stat.	0.301	0.378	0.310	0.271	0.269	0.274	0.281	0.284	0.285	0.285	0.285	0.286	0.289
	Copula stat.	92.3	93.8	93.8	98.0	95.8	94.4	93.6	93.1	92.7	92.4	92.3	91.9	91.8

Table 8.13 The p-values and test statistic values obtained for modelling the Euro (€) against the Japanese Yen (¥)

€ vs $C, 1999-2005

		Normal	Clayton	Gumbel	Student's t, degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.318	0.384	1.238	0.297	0.311	0.318	0.322	0.325	0.326	0.327	0.327	0.327	0.326
	std error	0.001	0.009		0.014	0.009	0.006	0.005	0.004	0.003	0.003	0.002	0.002	0.002
Chi-Squared Tests	Spokes	0.709	0.144	0.990	0.946	0.875	0.801	0.776	0.776	0.776	0.776	0.776	0.682	0.682
	Circles	0.484	0.866	0.940	0.313	0.309	0.384	0.479	0.577	0.561	0.511	0.549	0.511	0.453
	Quadrants	0.529	0.140	0.762	0.783	0.699	0.615	0.620	0.620	0.620	0.620	0.620	0.533	0.533
	3x3 Squares	0.838	0.373	0.448	0.939	0.906	0.805	0.805	0.805	0.689	0.672	0.672	0.617	0.795
	4x4 Squares	0.839	0.326	0.923	0.969	0.948	0.904	0.887	0.887	0.887	0.887	0.887	0.752	0.817
Kolmogorov-Smirnov	U2	0.594	0.790	0.396	0.552	0.485	0.470	0.432	0.427	0.437	0.448	0.461	0.439	0.483
	Copula	0.952	0.558	0.791	0.938	0.969	0.964	0.962	0.962	0.961	0.961	0.961	0.959	0.958
Anderson-Darling	U2 stat.	0.489	0.261	0.612	0.338	0.386	0.412	0.428	0.438	0.444	0.448	0.451	0.460	0.467
	Copula stat.	89.1	93.4	92.4	97.2	94.7	93.1	92.1	91.4	90.9	90.6	90.3	89.6	89.4

Table 8.14 The p-values and test statistic values obtained for modelling the Euro (€) against the Canadian Dollar ($C)

¥ vs $C, 1996-2005

		Normal	Clayton	Gumbel	Student's t, degrees of freedom 3	4	5	6	7	8	9	10	15	20
Parameter Estimates	rho / theta	0.214	0.220	1.131	0.151	0.167	0.177	0.184	0.189	0.192	0.195	0.197	0.204	0.206
	std error	0.001	0.004		0.012	0.007	0.005	0.004	0.003	0.003	0.002	0.002	0.002	0.001
Chi-Squared Tests	Spokes	0.650	0.255	0.676	0.297	0.252	0.428	0.328	0.397	0.397	0.397	0.390	0.558	0.628
	Circles	0.519	0.865	0.903	0.905	0.892	0.927	0.925	0.886	0.823	0.735	0.743	0.743	0.597
	Quadrants	0.834	0.926	0.939	0.919	0.945	0.939	0.877	0.877	0.877	0.877	0.877	0.877	0.885
	3x3 Squares	0.991	0.713	0.991	0.982	0.997	0.992	0.988	0.988	0.988	0.992	0.992	0.992	0.988
	4x4 Squares	0.877	0.801	0.995	0.935	0.963	0.884	0.867	0.867	0.890	0.890	0.832	0.881	0.893
Kolmogorov-Smirnov	U2	0.976	0.952	0.987	0.879	0.943	0.960	0.946	0.931	0.919	0.935	0.945	0.961	0.969
	Copula	0.900	0.933	0.772	0.949	0.922	0.898	0.879	0.863	0.963	0.956	0.951	0.933	0.924
Anderson-Darling	U2 stat.	0.248	0.147	0.136	0.191	0.205	0.217	0.224	0.230	0.233	0.236	0.238	0.241	0.242
	Copula stat.	156.4	163.1	162.7	170.6	167.4	165.3	163.9	162.8	162.0	161.3	160.8	159.3	158.5

Table 8.15 The p-values and test statistic values obtained for modelling the Japanese Yen (¥) against the Canadian Dollar ($C)

9. Conclusions and Further Extensions

The popularity of copulas looks certain to increase in future years; however, the study of copulas is far from complete. Research continues into topics such as extending bivariate copulas into the multivariate case and incorporating time series techniques into copula modelling to more accurately model the dependence structure observed in a dataset. Importantly, more research is needed in the field of multivariate goodness-of-fit statistics and their distributions to reduce the reliance upon Monte-Carlo simulations when considering goodness-of-fit of a particular model.

This report was intended to be a preliminary investigation into the dependence structure between foreign exchange rates, and has revealed several areas which could be vastly improved upon in further studies. The inconclusive results from this study may result from failing to consider fluctuations due to volatility into account when transforming the data. Volatility in financial data can vary weekly due to socio-economic and political events influencing the exchange markets, and techniques are available to detrend data by removing volatiliy (see Embrechts et. al, 2002).

Further, the dependence between the European currencies appears to be predominantly elliptic in nature, whilst the geographically isolated currencies appear to exhibit extreme-value dependence as well as elliptic dependence. Further research into estimating and fitting mixture models is therefore required.

- **Time-Dependent Copulas**

A major problem with the analysis of financial markets is that of time dependency and volatility. To properly analyse fluctuations and dependendencies in stock markets, etc., the seasonal, cyclical, and long-term trends need to be removed from the data, along with any other time-dependent behaviour. A straight analysis of data without removing these trends introduces errors into the analysis and thus no reliable conclusions can be made.

Current research is being conducted into the integration of statistical time series analysis with copula theory to provide a unified model of time-dependence. This is a relatively new field, and to date very few papers have been published. After more theoretical work is published, this field can be expected to become a major element in the world of financial economics.

Anoher key component of time series that needs to be incorporated into the model is the concept of volatility. Dias and Embrechts (2004) provide comprehensive details for fitting copulas to deseasonalised high-frequency data with the volatility removed.

- **Mixture Models**

. Another issue in analysing copulas is that the real copula may be a ‘mixture model’ – that is, composed of several different dependence structures. For example, a mixture model of the Gaussian and Clayton copulas with parameters ρ and θ respectively would be given by

$$C_{\alpha,\rho\theta}(u,v) = \alpha\, C_{N;\rho}(u,v) \ + \ (1-\alpha)\, C_{C;\theta}(u,v)\left\{ u^{-\theta} + v^{-\theta} - 1 \right\}^{-1/\theta}$$

where $C_{N;\rho}(u,v)$ and $C_{C;\theta}(u,v)$ denote the Gaussian and Clayton copulas respectively.

To further complicate matters, a mixture model may incorporate many sub-copulas, with many different parameters. Estimating these parameters becomes very difficult as there may also be interaction effects between the influences.

- **High-Frequency Data**

One method of reducing time-series error is to take high-frequency data: instead of daily or weekly observations, many observations are made per hour. Corrections must still be made for volatility, but the long-term seasonal and cyclical trends are not present.

- **Extension to Multivariate Models**

Whilst this study and many others have only focussed on bivariate copulas, the dependence structure between financial markets is clearly not as simple as the bivariate models suggest. In any financial market, there is a complex interaction of numerous variables to provide the fluctuations exhibited by the data. With this in mind, another key area for investigation is the application of multivariate statistical theory to copulas - in particular, the distributions of the multivariate goodness-of-fit statistics.

- **Binning-Free Goodness-of-Fit Tests**

Several authors have tried to escape the limitations imposed upon them by the arbitrary ‘binning’ of data for the classical Pearson’s chi-squared goodness-of-fit tests by developing an alternative test or statistic. One of particular note is the ‘energy’ approach proposed by Aslan et.al (2003): they propose a test based upon the observed cdf and the corresponding simulated cdf, much like a Monte Carlo approach to obtaining critical values.

10. Bibliography

Aas, K. (2004) "*Modelling the dependence structure of financial assets: A survey of four copulas."* Technical report, Norwegian Computing Centre. http://www.nr.no/files/samba/bff/SAMBA2204.pdf

Ang, A. & Chen, J. (2002) "*Asymmetric correlations of equity portfolios*." Journal of Financial Economics, vol. 63, pp 443–494

Aslan, B., & Zech, G. (2005) "*A new class of binning-free, multivariate goodness-of-fit tests: the energy tests."* Working paper, University of Siegen. http://arxiv.org/PS_cache/hep-ex/pdf/0203/0203010.pdf

Capéraà, P., Fougères, A.-L., & Genest, C. (1997) "*A non-parametric estimation procedure for bivariate extreme value copulas.*" Biometrika, vol. 84(3), pp 567-577.

Clayton, D.G. (1978 "A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incident." Biometrika, vol. 65, pp 141–151.

Cook, R.D. & Johnson, M.E. (1981) "*A family of distributions for modelling non-elliptically symmetric multivariate data."* Journal of the Royal Statistical Society, B 43, pp 210–218.

Demarta, S., & McNeil, A.J. (2005) "The t copula and related copulas." International Statistical Review, vol. 73(1), pp 111–130.

Embrechts, P., McNeil, A.J. & Straumann, D. (2002) "Correlation and dependence in risk management: properties and pitfalls." In Risk Management: value at risk and beyond (ed. M. A. H. Dempster), pp 176–223, Cambridge University Press, Cambridge.

Genest, G., Ghoudi, K., & Rivest, L.-P. (1995) "A semi-parametric estimation procedure of dependence paramters in multivariate families of distributions." Biometrika, vol. 82(3) pp 543-552.

Genest, G. & Rivest, L.-P. (1993) "Statistical inference procedures for bivariate Archimedean copulas." Journal of the American Statistical Association, vol. 88(423), pp 1034-1043

Genest, C., Quessy, J.-F., & Rémillard, B. (2006) "Goodness-of-fit Procedures for Copula Models Based on the Probability Integral Transformation." The Scandinavian Journal of Statistics, vol. 33, in press.

Genest, C., & MacKay, J. (1986) "The joy of copulas: Bivariate distributions with uniform marginals." Journal of the American Statistical Association, vol. 40, pp 280-285.

Joe, H. (1997) Multivariate Models and Dependence Concepts, Chapman & Hall, London, chs 2&5.

Longin, F., & Solnik, B. (2001) "Extreme correlation of international equity markets." Journal of Finance, vol.56(2), pp 649–676.

Mashal, R., & Zeevi, A. (2002) "Beyond correlation: extreme co-movements between financial assets." Working paper, Columbia Graduate School of Business.

Schweizer, B., & Wolff, E. (1981) "On non-parametric measures of dependence for random variables." Annals of Statistics, vol. 9, pp 879-885

Appendix I : Distribution Theory

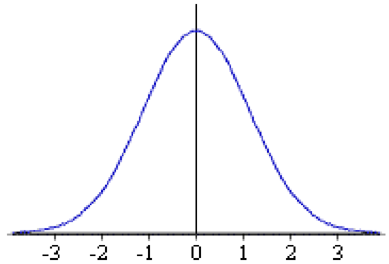


Fig 4.1 The 'bell curve', or standard normal pdf

1 Gaussian or Normal Distribution

Let Φ denote the standard normal cdf, i.e.

$$\Phi(x) = P(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left\{-\tfrac{1}{2}s^2\right\} ds \quad \text{where} \quad X \sim N(0,1)$$

$$\Phi_\rho(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\tfrac{1}{2}\frac{s^2 - 2\rho st + t^2}{1-\rho^2}\right\} dt\, ds\,,$$

where $(X,Y)^T \sim N_2\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$ and ρ is the coefficient of correlation between the random variables X and Y.

2 Student t-distribution with ν degrees of freedom

Let t_υ be the standard Student's t cdf with ν degrees of freedom, i.e.

$$t_\upsilon(x) = P(X \le x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \left\{1 + \frac{s^2}{\upsilon}\right\}^{-\frac{\upsilon+2}{2}} ds$$

$$t_{\rho,\upsilon}(x,y) = P(X \le x, Y \le y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{1}{2\pi\sqrt{1-\rho^2}} \left\{1 + \frac{s^2 - 2\rho st + t^2}{\upsilon(1-\rho^2)}\right\}^{-\frac{\upsilon+2}{2}} dtds$$

where ρ is the coefficient of correlation between the random variables X and Y.

Appendix II : Tail Dependence

The concept of tail dependence is crucial to the model selected for a particular set of financial data. Tail dependence measures the probability of joint extreme events, and can occur in the upper tails, the lower tails, or both.

Dependence in the upper tail is defined to be the probability of two extreme events ibn the upper tails occurring jointly. We can define this as

$$\lambda_{upper} = \lim_{u \to 1} \left[\Pr\left(Y \geq F_Y^{-1}(u) \mid X \geq F_X^{-1}(u) \right) \right]$$

For continuous random variables, this expression is equivalent to

$$\lambda_{upper} = \lim_{u \to 1} \left[\frac{1 - 2u + C(u,u)}{1-u} \right]$$

We can define lower tail dependence similarly:

$$\begin{aligned} \lambda_{lower} &= \lim_{u \to 0} \left[\Pr\left(Y \leq F_Y^{-1}(u) \mid X \leq F_X^{-1}(u) \right) \right] \\ &= \lim_{u \to 0} \left[\frac{C(u,u)}{u} \right] \end{aligned}$$

The tail dependency depends only upon the underlying copula, not the marginal distributions. Importantly, the Gaussian copula has asymptotically indepdendent tails, while the Student's t-distribution has asymptotically dependent tails. As mentioned before, the Gumbel copula exhibits upper tail dependence while the Clayton copula exhibits left-tail dependence.

Since the majority of financial data displays some tail dependency, the t-copula and the Archimedean copulas may better describe the actual behaviour of financial markets.

View publication stats