Multivariate Dependence Modeling Using Pair-Copulas

Doris Schirmacher
Ernesto Schirmacher[1]

Copyright 2008 by the Society of Actuaries.
All rights reserved by the Society of Actuaries. Permission is granted to make brief excerpts for a published review. Permission is also granted to make limited numbers of copies of items in this monograph for personal, internal, classroom or other instructional use, on condition that the foregoing copyright notice is used so as to give reasonable notice of the Society's copyright. This consent for free limited copying without prior consent of the Society does not extend to making copies for general distribution, for advertising or promotional purposes, for inclusion in new collective works or for resale.

[1] *Corresponding author:* Liberty Mutual Group, 175 Berkeley St., Boston, Mass. 02116. Ernesto.Schirmacher@LibertyMutual.com.

Abstract

In the copula literature there are many bivariate distribution families but very few higher dimensional ones. Moreover, most of these are difficult to work with. Some of the bivariate families can be extended to more dimensions but in general the construction of distribution functions with more than two variables is a difficult problem. We introduce a construction method that is straightforward to implement and can produce multivariate distribution functions of any dimension. In essence the method takes an arbitrary multivariate density function and decomposes it into a product of bivariate copulas and marginal density functions. Each of these bivariate copulas can be from any of the available families.

We also highlight the power of a graphical display known as a *chi-plot* to help us understand the dependence between pairs of variables. One illustration, based on changes in the exchange rate of three currencies, shows how we can specify the pair-copulas and estimate their parameters. In another illustration we simulate data that exhibits complex dependencies as would be found, for example, in enterprise risk management or dynamic financial analysis.

1. Introduction

Actuaries are routinely called upon to analyze complex financial security systems. The outcomes of these systems depend on many variables that have complex dependencies. Until recently, actuaries have had a rather limited set of tools to analyze, extract and make use of the information embedded in multivariate distributions. The best known tools have been the linear correlation coefficient and the scatterplot. Linear correlation or Pearson correlation is a global measure that attempts to summarize in a single number the dependence between two variables. We cannot expect the correlation coefficient to be able to adequately summarize complex dependencies into a single number, and it is well known [2, 12] that two datasets with very different dependence patterns can have the same correlation coefficient. To alleviate this situation the scatterplot has been used very effectively to display the entire dataset. Here one can fully appreciate any patterns, if they are strong enough. Unfortunately, our eyes would rather see a pattern where none exists. We will introduce a graphical display designed to alleviate this problem.

In recent years there has been increased attention in the combined management of all risk sources. It is no longer best practice to understand each risk source in isolation. Today we not only need to consider each risk source but more importantly how all risk sources relate to each other and their potential synergy to create catastrophic losses when several factors align properly. Thus we need to understand the joint distribution of all risk sources. Unfortunately, the number of tractable multivariate distributions of dimension three or higher, such as the multivariate normal and t-distributions, is rather limited. Moreover, for these two families the marginal distributions are also normal or t-distributed, respectively. This restriction has limited their useful application in practical situations. What is needed is a construction that would allow us to specify the marginal distributions independently from the dependence structure. This we can do with the theory of copulas. Early contributions to the application of copulas include [13, 15, 19, 27] and in recent years we have seen more activity closely linked with applications in finance [7, 26], insurance [5, 9, 25], enterprise risk management [7, 3] and other areas. Readers new to copula methods should consult [14].

In a recent contribution [26] the authors state:

> While a variety of bivariate copulas is available, when more than two variables are involved the practical choice comes down to normal vs. t-copula. The normal copula is essentially the t-copula with high degrees of freedom (df), so the choice is basically what df to use in that copula.

They go on to introduce two new copula families; the IT and MCC. While these two families are defined for any dimension their dependence properties are somewhat limited and they are not very tractable even in three dimensions.

In this paper we introduce the work of Aas et al. [1] to construct multivariate distributions of any dimension. Their method relies in using bivariate copulas and pasting them together appropriately to arrive at a distribution function. The construction is straightforward and easy to implement. In Appendix F we have reproduced the algorithms in [1] that carry out simulation and maximum likelihood estimation.

In Section 2, we start the study of dependence by showing that the marginal distributions affect our perception of the association between two variables. Therefore, if we want to understand how two variables are interrelated we must remove the marginal distributions and look at the rank transformed data [14, p. 349]. We also show that the widely used Pearson correlation coefficient is a poor global measure of dependence because it is not invariant under all strictly increasing transformations of the underlying variables. But the next two best known global measures of association; namely, Kendall's τ and Spearman's ρ_s, are invariant under all such transformations.

Section 3 introduces some elementary properties of copulas and Sklar's Theorem. This theorem is a crucial result. Basically it states that any multivariate distribution can be specified via two independent components:

1. marginal distribution functions, and
2. a copula function that provides the dependence structure.

We also state how both Kendall's τ and Spearman's ρ_s can be expressed in terms of the underlying copula.

Section 4 provides all the details on how to decompose a multivariate density function into pair-copulas. For a given density function there are many possible pair-copula decompositions. To help us organize them there is a graph construction known as a *regular vine* [4]. Regular vines are a rather large class of decompositions and so we will only work with two subsets known as *canonical* and *D-vines*.

In Section 5, we show how to obtain the maximum likelihood parameters for a canonical or D-vine. We also introduce a powerful graphical display, known as a chi-plot or χ-plot [10, 11], to help us assess the dependence between two variables.

Finally, Section 6 is devoted to a numerical example based, as in [26], on currency rate changes and in Section 7 we show some of the flexibility of the pair-copula construction by simulating a D-vine structure with various parameters. Appendices A–E provided some basic information on the following one-parameter families of copulas

1. Clayton
2. Frank
3. Galambos
4. Gumbel
5. Normal.

These copulas are just a small sample from all bivariate copulas. For more information on these and other copulas refer to [18, 20, 22].

Appendix F reproduces four algorithms taken from [1]. These algorithms perform simulation and maximum likelihood calculations for canonical and D-vine structures.

2. Understanding Dependence

The word "correlation" has been frequently used (or misused) as an over-arching term to describe all sorts of dependence between two random variables. We will use the word only in its technical sense of *linear correlation* or *Pearson's correlation*, denote it by ρ, and define it as

Definition 1. Let X and Y be two random variables with non-zero finite variances. The linear correlation coefficient for (X, Y) is

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)}\sqrt{\mathrm{Var}(Y)}}, \tag{1}$$

where Cov and Var are the covariance and variance operators, respectively.

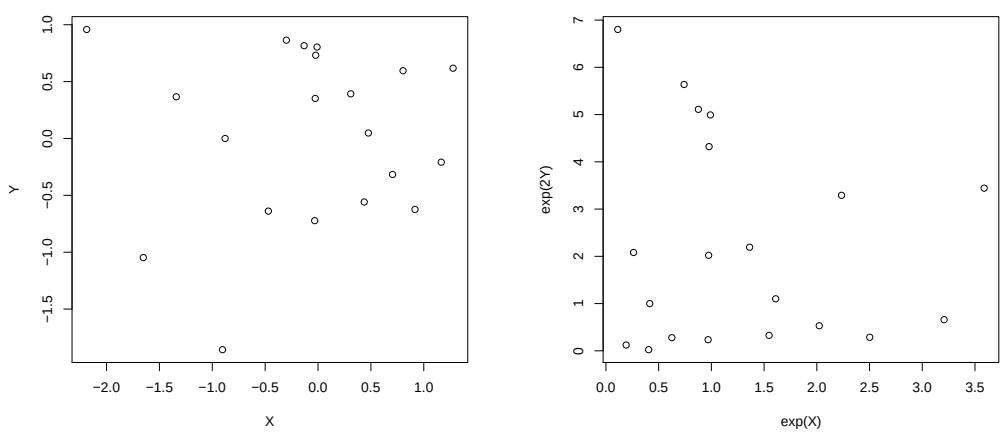
This quantity, ρ, is a measure of linear dependence; that is, if Y depends on X linearly (namely, $Y = aX + b$ with $a, b \in \mathbb{R}$ and $a \neq 0$), then the absolute value of $\rho(X, Y)$ is equal to 1. We also know that linear correlation is invariant only under strictly increasing *linear* transformations. In addition, linear correlation is the correct dependence measure to use for multivariate normal distributions; but for other joint distributions it can give very misleading impressions.

Suppose we have a set of observations $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ from an unknown bivariate distribution $H(x, y)$. We would like to identify the distribution H that characterizes their joint behavior. We could start our investigations by looking at the scatterplot of the data and try to discover some pattern that would point us to the correct choice of bivariate distribution. This approach has long been used with some successes and some

failures. The difficulty of effectively using the scatterplot stems from the fact that this display not only gives us information about the dependency between X and Y but also about their marginal distributions. While marginal distributions are vital for other analyses it distorts the information about their dependence.

For example, Figure 2.1 shows a conventional scatterplot on the left-hand side and a monotone transformation on the right-hand panel. The Pearson correlation coefficient on the left-hand side is approximately $\rho = 0.07$, but on the right-hand side is decidedly different at about $\rho = -0.15$.

Figure 2.1
Effect of Monotone Transformations



The left-hand panel shows 20 points (x_i, y_i) sampled at random from a bivariate standard normal distribution. The points on the right-hand panel are given by the monotone transformation $(x_i, y_i) \mapsto (\exp(x_i), \exp(2y_i))$.

Both scatterplots in Figure 2.1 definitely have a different qualitative feel to them. Nonetheless, in both displays the underlying bivariate dependence between the two variables has not changed. Only their marginal distributions are different. This shows that the Pearson correlation coefficient is a poor measure of the association between two variables. In particular, it is not invariant under strictly increasing transformations, and this is a major objection to its use as a global measure of dependence.

The fact that the dependence between two variables is invariant under increasing monotone transformations is based on two key results. The first one is a representation theorem due to Sklar [24, 23] that states that the joint distribution function $H(x, y)$ of any pair of continuous random variables (X, Y) may be written in the form

$$H(x, y) = C\Big(F(x), G(y)\Big), \qquad x, y \in \mathbb{R} \tag{2}$$

where $F(x)$ and $G(y)$ are the marginal distributions of X and Y, and C is a function mapping $[0,1] \times [0,1] \to [0,1]$ known as a *copula*.

The second result says that if the pair (W, Z) is a monotone increasing transform of the pair (X, Y), then the copula that characterizes the joint behavior of (W, Z) is exactly the same copula as for the pair (X, Y). That is, copulas are invariant under monotone increasing transformations [14, p. 348].

Since the copula that characterizes dependence is invariant under strictly monotone transformations, then a better global measure of dependence would also be invariant under such transformations. Both Kendall's τ and Spearman's ρ_s are invariant under strictly increasing transformations, and, as we will see in the next section, they can be expressed in terms of the associated copula.

Kendall's τ measures the amount of concordance present in a bivariate distribution. Suppose that (X, Y) and $(\widetilde{X}, \widetilde{Y})$ are two pairs of random variables from a joint distribution function. We say that these pairs are *concordant* if "large" values of one tend to be associated with "large" values of the other, and "small" values of one tend to be associated with "small" values of the other. The pairs are called *discordant* if "large" goes with "small" or vice versa. Algebraically we have concordant pairs if $(X - \widetilde{X})(Y - \widetilde{Y}) > 0$ and discordant pairs if we reverse the inequality. The formal definition is

Definition 2. Kendall's τ for the random variables X and Y is defined as

$$\tau(X, Y) = \text{Prob}((X - \widetilde{X})(Y - \widetilde{Y}) > 0) - \text{Prob}((X - \widetilde{X})(Y - \widetilde{Y}) < 0), \tag{3}$$

where $(\widetilde{X}, \widetilde{Y})$ is an independent copy of (X, Y).

Definition 3. Spearman's ρ_s for the random variables X and Y is equal to the linear correlation coefficient on the variables $F_1(X)$ and $F_2(Y)$ where F_1 and F_2 are the marginal distributions of X and Y, respectively.

3. Copulas

Sklar's Theorem is probably the most important theorem in the study of dependence and as in (2) it allows us to write any joint distribution in terms of a copula function and the marginal distributions. The theorem is valid not only in the bivariate case ($n = 2$), but also in all higher dimensions ($n > 2$). Thus copula functions play a crucial role in our understanding of dependence. Since the underlying copula is invariant under increasing monotone transformations, the study of dependence should be free of marginal effects.

That is, for dependence purposes, we should look at our data based on their ranks [14, p. 349].

Copulas are fundamental for understanding the dependence between random variables. With them we can separate the underlying dependence from the marginal distributions. This decomposition is analogous to the way the multivariate normal distribution is specified; namely, we need two components:

1. a vector of means and
2. a covariance matrix.

The vector of means provides the location for each of the marginal distributions and the covariance matrix tells us about the dependence between the variables. Moreover, these two components are independent of each other.

When we consider more general multivariate distributions we can have a similar decomposition as above. Basically, any multivariate distribution $F(x_1, \ldots, x_n)$ can be specified by providing two components:

1. the marginal distributions $F_1, F_2, \ldots, F_n$, and
2. a copula $C\colon [0,1]^n \to [0,1]$ that provides the dependence structure.

For example, suppose that $F_1, F_2, \ldots, F_n$ are given marginal distributions.[2] Let the copula function $C(x_1, \ldots, x_n)$ be equal to the product of its arguments; that is,

$$C(x_1, \ldots, x_n) = x_1 x_2 \cdots x_n.$$

Then the resulting distribution $F(x_1, \ldots, x_n)$ defined via

$$F(x_1, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$$

gives us a multivariate distribution F where the n random variables are independent of each other. The copula $C(x_1, \ldots, x_n) = x_1 \cdots x_n$ is called the *independence* copula and it is usually denoted by the symbol Π;

$$\Pi(x_1, \ldots, x_n) = x_1 x_2 \cdots x_n. \tag{4}$$

[2] There are no restrictions on the marginal distributions besides being continuous and even this can be relaxed. In particular, we don't have to choose the margins to be all from the same family of distributions as would happen if we used a multivariate normal distribution or a multivariate t-distribution.

So whenever we are dealing with a multivariate distribution F we should separate the marginal distributions F_i from the dependence structure that is given by some copula function C. Otherwise, we risk getting tangled up between whatever association there might be between the variables and their marginal distributions.

The fact that we can always decompose a multivariate distribution F in its marginal distributions F_i and a copula function C is known as Sklar's Theorem [23, p. 83].

Theorem 1. *Let F be an n-dimensional distribution function with marginal functions $F_1, F_2, \ldots, F_n$. Then there exists an n-dimensional copula C such that for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$,*

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{5}$$

If the functions $F_1, \ldots, F_n$ are all continuous, then C is unique; otherwise, C is uniquely determined on $\operatorname{Ran} F_1 \times \cdots \times \operatorname{Ran} F_n$.

Conversely, if C is an n-copula and $F_1, \ldots, F_n$ are distribution functions, then the function F defined via (5) is an n-dimensional distribution function with margins $F_1, \ldots, F_n$.

Another key result for understanding the dependence among variables is that copulas are invariant under strictly increasing transformations [23, p. 91].

Theorem 2. *For $n \geq 2$ let $X_1, X_2, \ldots, X_n$ be random variables with continuous distribution functions $F_1, F_2, \ldots, F_n$, joint distribution function F, and copula C. Let $f_1, f_2, \ldots, f_n$ be strictly increasing function from $\mathbb{R}$ to $\mathbb{R}$. Then $f_1(X_1), f_2(X_2), \ldots, f_n(X_n)$ are random variables with continuous distribution functions and copula C. Thus C is invariant under strictly increasing transformation of $X_1, X_2, \ldots, X_n$.*

The implication of this result is that *any property* of the joint distribution function that is invariant under strictly increasing transformations is also a *property* of their copula and it is independent of the marginal distributions. Thus the study of dependence among variables is really about the study of copulas.

We know from Figure 2.1 that the linear correlation coefficient is not invariant under strictly increasing transformations. But as stated informally in the previous section, both Kendall's τ and Spearman's ρ_s are invariant under such transformations. In fact, these two measures of dependence can be expressed in terms of the underlying copula as stated in the next two theorems. For the proofs see [22, p. 127, 135].

Theorem 3. *Let X and Y be continuous random variables with copula C. Then Kendall's τ is given by*

$$\tau(X,Y) = 4 \iint_{[0,1]^2} C(u,v)\, dC(u,v) - 1. \tag{6}$$

Theorem 4. *Let X and Y be continuous random variables with copula C. Then Spearman's ρ_s is given by*

$$\rho_s(X,Y) = 12 \iint_{[0,1]^2} uv\, dC(u,v) - 3. \tag{7}$$

4. The Pair-Copula Construction

The construction of distribution functions with $n > 2$ dimensions is recognized as a difficult problem. In [26] the authors say:

> The MMC[3] copula densities get increasingly difficult to calculate as the dimension increases. For this reason, some alternatives to MLE[4] were explored. One alternative is to maximize the product of the bivariate likelihood functions, which just requires the bivariate densities.

And later they add:

> It should be noted, however, that there are many local maxima for both likelihood functions, so we cannot be absolutely sure that these are the global maxima.

In this section we present the results of Aas et al. [1] and follow their exposition very closely. The basic idea behind the pair-copula construction is to decompose an arbitrary distribution function into simple bivariate building blocks and stitch them together appropriately. These bivariate blocks are two-dimensional copulas and we have a large selection to choose from [18, 22].

Before we present the general case let us illustrate the construction for two, three, and four dimensions. The method is recursive in nature. For the base case in two dimensions we can easily see that the density function $f(x_1, x_2)$ is given by

$$f(x_1, x_2) = c_{12}\Big(F_1(x_1), F_2(x_2)\Big) \cdot f_1(x_1) \cdot f_2(x_2). \tag{8}$$

[3] These are multivariate copulas with general dependence introduced in [18, p. 163].
[4] Maximum likelihood estimation.

This follows immediately by taking partial derivatives with respect to both arguments in $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$, where C is the copula associated with F via Sklar's Theorem.

Before we move on to the next case, note that from (8) we can determine what the conditional density of X_2 given X_1 is; that is,

$$f_{2|1}(x_2|x_1) = \frac{f(x_1, x_2)}{f_1(x_1)} = c_{12}\Big(F_1(x_1), F_2(x_2)\Big) \cdot f_2(x_2). \tag{9}$$

This formula in its general form

$$f_{j|i}(x_j|x_i) = \frac{f(x_i, x_j)}{f_i(x_i)} = c_{ij}\Big(F_i(x_i), F_j(x_j)\Big) \cdot f_j(x_j). \tag{10}$$

will come in very handy as we move up into higher dimensions.

Next let us build a three-dimensional density function. Any such function can always be written in the form

$$f(x_1, x_2, x_3) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|1,2}(x_3|x_1, x_2), \tag{11}$$

and this factorization is unique up to a relabeling of the variables. Note that the second term on the right-hand side $f_{2|1}(x_2|x_1)$ can be written in terms of a pair-copula and a marginal distribution using (9). As for the last term $f_{3|1,2}(x_3|x_1, x_2)$ we can pick one of the conditioning variables, say x_2, and use a form similar to (10) to arrive at

$$f_{3|12}(x_3|x_1, x_2) = c_{13|2}(F_{1|2}(x_1|x_2), F_{3|2}(x_3|x_2)) \cdot f_{3|2}(x_3|x_2). \tag{12}$$

This decomposition involves a pair-copula and the last term can then be further decomposed into another pair-copula, using (10) again, and a marginal distribution. This yields, for a three-dimensional density, the full decomposition

$$\begin{aligned} f(x_1, x_2, x_3) = f_1(x_1) \cdot & \\ & c_{12}\Big(F_1(x_1), F_2(x_2)\Big) \cdot f_2(x_2) \cdot \\ & c_{31|2}(F_{3|2}(x_3|x_2), F_{1|2}(x_1|x_2)) \cdot c_{23}\Big(F_2(x_2), F_3(x_3)\Big) \cdot f_3(x_3). \end{aligned} \tag{13}$$

For a four-dimensional density we start with

$$f(x_1, x_2, x_3, x_4) = f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|1,2}(x_3|x_1, x_2) \cdot f_{4|1,2,3}(x_4|x_1, x_2, x_3) \tag{14}$$

and use (10) repeatedly together with the previous results to rewrite it in terms of six pair-copulas and the four marginal densities $f_i(x_i)$ for $i = 1, 2, 3, 4$:

$$
\begin{aligned}
f(x_1, x_2, x_3, x_4) = f_1(x_1) \cdot & \\
& c_{12}\Big(F_1(x_1), F_2(x_2)\Big) \cdot f_2(x_2) \cdot \\
& c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) \cdot c_{13}\Big(F_1(x_1), F_3(x_3)\Big) \cdot f_3(x_3) \cdot \\
& c_{34|12}\Big(F_{3|12}(x_3|x_1, x_2), F_{4|12}(x_4|x_1, x_2)\Big) \cdot \\
& c_{24|1}\Big(F_{2|1}(x_2|x_1), F_{4|1}(x_4|x_1)\Big) \cdot \\
& c_{14}\Big(F_1(x_1), F_4(x_4)\Big) \cdot f_4(x_4)
\end{aligned}
\tag{15}
$$

Notice that in the construction many of the pair-copulas need to be evaluated at a conditional distribution of the form $F(x|\mathbf{v})$ where $\mathbf{v}$ denotes a vector of variables. The calculation of these conditional distributions is also recursive. Let $\mathbf{v}_{-j}$ denote the vector $\mathbf{v}$ but excluding the jth component v_j. For every j, Joe [17] has shown that

$$
F(x|\mathbf{v}) = \frac{\partial C_{x,v_j|\mathbf{v}_{-j}}\Big(F(x|\mathbf{v}_{-j}), F(v_j|\mathbf{v}_{-j})\Big)}{\partial F(v_j|\mathbf{v}_{-j})}, \tag{16}
$$

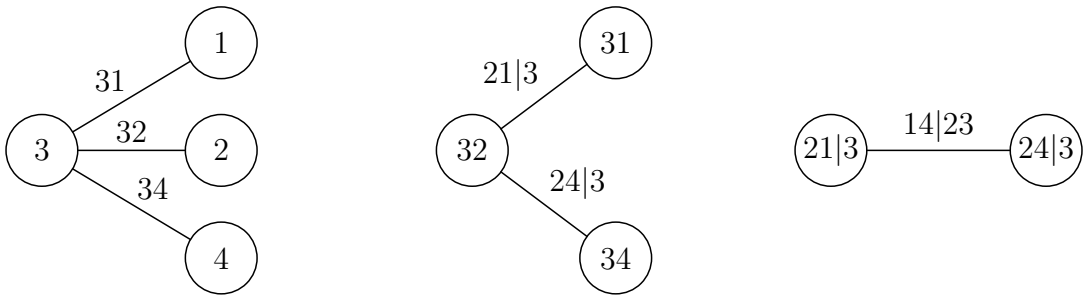
where $C_{x,v_j|\mathbf{v}_{-j}}$ is a bivariate copula function. For the special case where $\mathbf{v}$ has only one component we have

$$
F(x|v) = \frac{\partial C_{xv}\Big(F_x(x), F_v(v)\Big)}{\partial F_v(v)}. \tag{17}
$$

As we decompose a joint density function $f(x_1, \ldots, x_n)$ into a product of pair-copulas and the marginal densities $f_1, \ldots, f_n$ we need to make many choices in the conditioning variables. This leads to a large number of possible pair-copulas constructions. Bedford and Cooke [4] have introduced a graphical model, called a *regular vine*, to organize all possible decompositions. But regular vine decompositions are very general; therefore, we will only concentrate on two subsets called *D-vines* and *canonical vines*. Both models give us a specific way of decomposing a density function. These models can be specified as a nested set of trees. Figure 4.1 shows a canonical vine decomposition for a 4-dimensional density function and Figure 4.2 shows a D-vine.

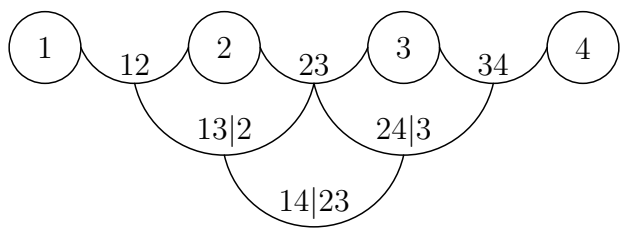
Similar constructions are possible for any number of variables. The intuition behind canonical vines is that one variable plays a key role in the dependency structure and so everyone is linked to it. For a D-vine things are more symmetric.

Figure 4.1
Canonical Vine Representation



Three trees representing the decomposition of a four-dimensional joint density function. The circled nodes on the left-most tree represent the four marginal density functions f_1, f_2, f_3, f_4. The remaining nodes on the other trees are not used in the representation. Each edge corresponds to a pair-copula function.

Figure 4.2
D-Vine Representation



Three trees representing the decomposition of a four-dimensional joint density function into pair-copulas and marginal densities. The circled nodes represent the four marginal density functions f_1, f_2, f_3, f_4. Each edge is labeled with the pair-copula of the variables that it represents. The edges in level i become nodes for level $i+1$. The edges for tree 1 are labeled as 12, 23 and 34. Tree 2 has edges labeled 13|2 and 24|3. Finally, tree 3 has one edge labeled 14|23.

In general, a canonical vine or D-vine decomposition of a joint density function with n variables involves $\binom{n}{2}$ pair-copulas. For the first tree in the representation we have $n-1$ edges. The second tree has $n-1$ nodes (corresponding to the edges in the previous tree) and $n-2$ edges. Continuing in this manner we see that the total number of edges across all trees in the representation is equal to

$$(n-1)+(n-2)+\cdots+2+1=\frac{(n-1)n}{2}=\binom{n}{2}.$$

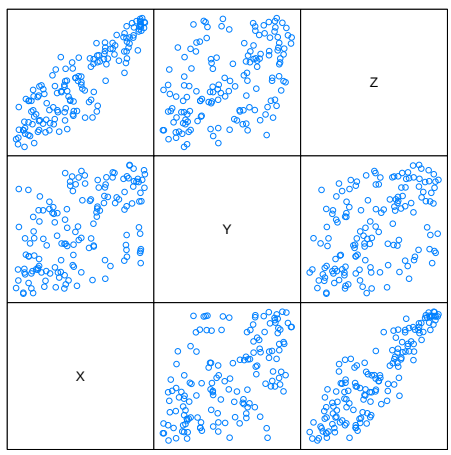
A four-dimensional joint density function requires the specification of six pair-copulas. For a five-dimensional joint density we have to specify ten pair-copulas and most of these need to be evaluated at conditional distribution functions. One way of reducing this complexity is by assuming conditional independence. Suppose we have a three-dimensional

problem where variable x_1 is linked to both x_2 and x_3 and so we would use a canonical vine representation.[5] If we assume that *conditional on* x_1 the variables x_2 and x_3 are independent, then the construction simplifies to

$$f(x_1, x_2, x_3) = f_1(x_1) f_2(x_2) f_3(x_3)\, c_{12}(F_1(x_1), F_2(x_2))\, c_{13}(F_1(x_1), F_3(x_3)) \tag{18}$$

because $c_{23|1}(F_{2|1}(x_2|x_1), F_{3|1}(x_3|x_1)) = 1$. Figure 4.3 shows the simulation of 150 points from the joint distribution in (18) (assuming uniform margins) where the copula c_{12} is from the Frank family with parameter 3, the c_{13} copula comes from the Galambos family with parameter 2, and $c_{23|1}$ is the independence copula.[6]

Figure 4.3
Simulated Canonical Vine



This display shows the pairwise scatterplots of 150 random points from a three-dimensional canonical vine with uniform marginals where the c_{12} copula is from the Frank family, the c_{13} copula comes from the Galambos family, and the conditional pair-copula $c_{23|1}$ is the independence copula. The parameters of the copulas c_{12} and c_{13} have been chosen so that the Kendall τ coefficient is approximately 0.3 and 0.65, respectively.

Simulation from canonical and D-vines is relatively straightforward and it is based on the following general sampling algorithm for n dependent uniform $[0, 1]$ variables. First, sample n independent uniform random numbers $w_i \in [0, 1]$ and now compute

[5] In the case of three variables every canonical vine is a D-vine and vice versa. In higher dimensions this is no longer the case.

[6] The parameters for the Frank and Galambos copulas have been chosen so that their Kendall τ coefficients are approximately equal to 0.3 and 0.65. For further parameter values see [18, Table 5.1].

$$
\begin{aligned}
x_1 &= w_1 \\
x_2 &= F^{-1}_{2|1}(w_2|x_1) \\
x_3 &= F^{-1}_{3|1,2}(w_3|x_1, x_2) \\
x_4 &= F^{-1}_{4|1,2,3}(w_4|x_1, x_2, x_3) \\
&\vdots \\
x_n &= F^{-1}_{n|1,2,3,\ldots,n-1}(w_n|x_1, \ldots, x_{n-1}).
\end{aligned}
$$

To implement this algorithm we will need to calculate conditional distributions of the form $F_{x|\mathbf{v}}(x|\mathbf{v})$ and their inverses where as before $\mathbf{v}$ is a vector of variables and $\mathbf{v}_{-j}$ denotes the same vector but excluding the jth component. The case where $\mathbf{v}$ has only one component this reduces to

$$
F(x|v) = \frac{\partial C_{xv}(F_x(x), F_v(v))}{\partial F_v(v)}. \tag{19}
$$

If we further assume that the marginal distributions $F_x(x)$ and $F_v(v)$ are uniform, then (19) reduces further to

$$
F(x|v) = \frac{\partial C_{xv}(x, v)}{\partial v}. \tag{20}
$$

This last construction will occur often in the simulation algorithm. So define the function $h(\cdot)$ via

$$
h(x, v; \Theta) = F(x|v) = \frac{\partial C_{xv}(x, v; \Theta)}{\partial v}, \tag{21}
$$

where the third argument Θ denotes the set of parameters for the copula associated with the joint distribution of x and v. Also the second argument of $h(\cdot)$ always denotes the conditioning variable. Moreover, let h^{-1} denote the inverse of h with respect to the first variable; that is, the inverse of the conditional distribution function.

Note that equation (16) can be used recursively, and at each stage the number of conditioning variables decreases by one. Eventually, we would arrive at the special case shown in (19). This will allow us to recursively compute the conditional distributions and so create a sample from the joint distribution function.

A appropriate choice of variable v_j to use in (16) will give us either the canonical vine or the D-vine. For the canonical vine we always choose v_j to be the last conditioning variable available

$$F(x_j|x_1,\ldots,x_{j-1}) = \frac{\partial C_{j,j-1|1,2,\ldots,j-2}(F(x_j|x_1,\ldots,x_{j-2}), F(x_{j-1}|x_1,\ldots,x_{j-2}))}{\partial F(x_{j-1}|x_1,\ldots,x_{j-2})}, \tag{22}$$

and for the D-vine we always choose the first conditioning variable

$$F(x_j|x_1,\ldots,x_{j-1}) = \frac{\partial C_{j,1|2,\ldots,j-1}(F(x_j|x_2,\ldots,x_{j-1}), F(x_1|x_2,\ldots,x_{j-1}))}{\partial F(x_1|x_2,\ldots,x_{j-1})}. \tag{23}$$

Algorithm 1, taken from [1], gives the pseudo-code for sampling from a canonical vine with uniform marginals. The use of uniform marginals is for simplicity only and the algorithm can easily be extended to other marginal distributions. We make heavy use of the h function defined in (21) and in the algorithm we set

$$v_{i,j} = F(x_i|x_1,\ldots,x_{j-1}).$$

The symbol $\Theta_{j,i}$ represents the set of parameters of the corresponding pair-copula density $c_{j,j+1|1,\ldots,j-1}$. Algorithm 2, also taken from [1], gives the pseudo-code to sample from a D-vine.

5. Estimating the Pair-Copula Decomposition

The last section showed how the canonical or D-vine constructions decompose an n-dimensional multivariate density function into two main components. The first one is the product of each of the marginal density functions. The second component is the product of the density functions of $n(n-1)/2$ bivariate copulas. To estimate the parameters of either construction we need to

1. decide which family to use for each pair-copula and
2. estimate all necessary parameters simultaneously.

5.1 Chi-plots to Determine Appropriate Pair-Copulas

To specify which bivariate copulas we want to use in the canonical or D-vine decompositions we will pursue a graphical method based on a construction of Fischer and Switzer [10, 11] known as a chi-plot or χ-plot. There are other more formal selection techniques [6, 16, 18] that should be used in conjunction with this method.

The χ-plot is a powerful graphical representation to help us extract information about the dependence between two random variables. Traditionally the scatterplot has been used to detect patterns (or lack of patterns) of association between two variables.

We know that if two random variables are independent, then the scatterplot should show a random arrangement of points. Unfortunately, the human eye is not very good at identifying randomness. We are all too eager to find some sort of pattern in the data. The χ-plot was designed as an auxiliary display in which independence is itself manifested in a characteristic way.

The essence of the χ-plot is to compare the empirical bivariate distribution against the null hypothesis of independence at each point in the scatterplot. To construct this plot from a set of points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ we calculate three empirical distribution functions: the bivariate distribution H and the two marginal distributions F and G. For each point (x_i, y_i) let H_i be the proportion of points below and to the left of (x_i, y_i). Also let F_i and G_i be the proportion of points to the left and below of the point (x_i, y_i), respectively. Figure 5.1 shows graphically how to calculate H_i, F_i and G_i.

Each point (χ_i, λ_i) of the χ-plot is then defined by

$$\chi_i = \frac{H_i - F_i G_i}{\sqrt{F_i(1 - F_i)G_i(1 - G_i)}} \tag{24}$$

and

$$\lambda_i = 4S_i \max\left\{\left(F_i - \frac{1}{2}\right)^2, \left(G_i - \frac{1}{2}\right)^2\right\}, \tag{25}$$

where

$$S_i = \operatorname{sign}\left\{\left(F_i - \frac{1}{2}\right)\left(G_i - \frac{1}{2}\right)\right\}. \tag{26}$$

The formal definitions for H, F and G are

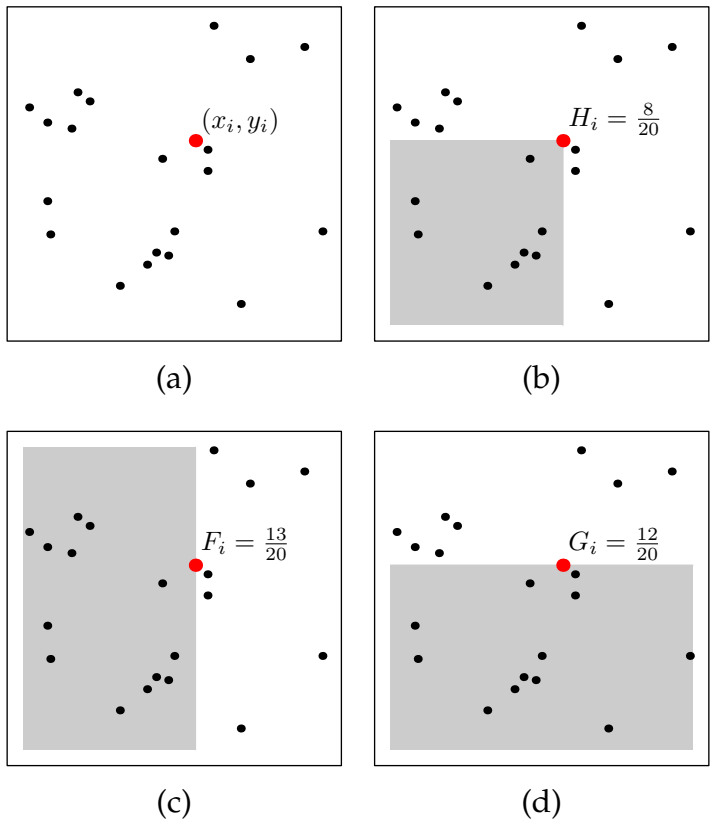
$$H_i = \frac{1}{n-1} \sum_{j \neq i} I(x_j \leq x_i, y_j \leq y_i), \tag{27}$$

$$F_i = \frac{1}{n-1} \sum_{j \neq i} I(x_j \leq x_i), \tag{28}$$

$$G_i = \frac{1}{n-1} \sum_{j \neq i} I(y_j \leq y_i), \tag{29}$$

where $I(A)$ is the indicator function of the event A. To avoid some erratic behavior at the edges of the dataset only the points that satisfy $|\lambda_i| < 4\{1/(n-1) - 1/2\}^2$ are included in the display. This restriction will eliminate at most eight points [10, p. 256].

Figure 5.1
Chi-plot Construction I



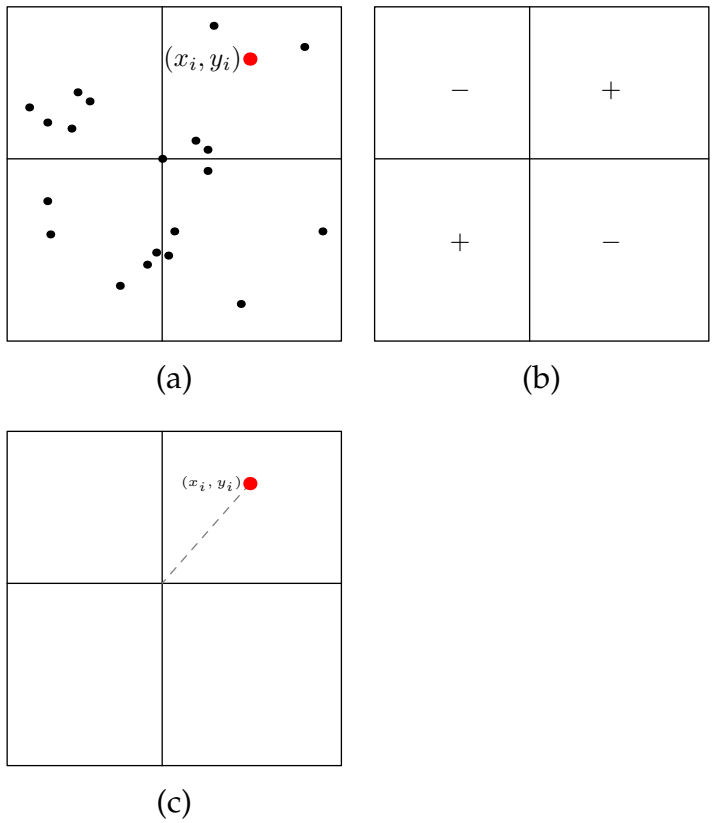
Panel (a) shows the scatterplot of $n = 21$ points with point (x_i, y_i) highlighted. Panel (b) shows the region where we need to count the number of points to calculate H_i. Panels (c) and (d) show the regions for the calculation of F_i and G_i, respectively. In all cases we do not count the point (x_i, y_i) as being inside the region.

The value of λ_i is a measure of the distance of the point (x_i, y_i) from the center of the dataset and the value of χ_i is a measure of the distance of the distribution H to the distribution of independent pairs of random variables (X, Y). Note that the functions H_i, F_i, G_i are the empirical distribution functions of the joint distribution and the marginal distributions of (X, Y) and depend only on the ranks of the observations. Figures 5.1 and 5.2 graphically show how to calculate the points necessary to construct a χ-plot.

If X and Y are independent, then the numerator of χ, which equals $H - F \cdot G$, is equal to zero. In practice we have a sample from the bivariate distribution H and if X and Y are independent we expect the χ-plot to show most points close to the line $\chi = 0$. In our χ-plots we have included a 95 percent confidence band around the null hypothesis of independence. If most points lie within these control limits, then we have strong evidence that the variables are indeed independent.

To understand what a χ-plot has to offer let's consider three examples where we know what the dependence between X and Y is and see how it is manifested in the χ-plot.

Figure 5.2
Chi-plot Construction II

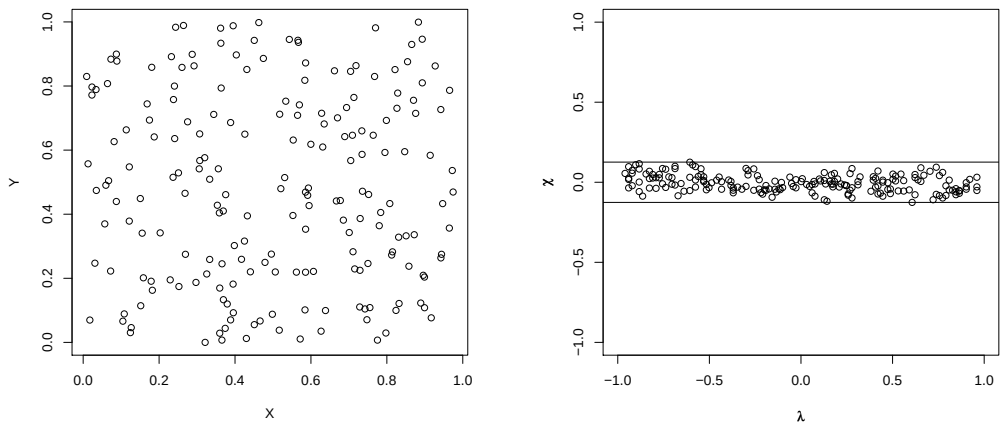


Panel (a) shows the scatterplot of $n = 21$ points with point (x_i, y_i) highlighted. The center of the dataset is defined as the point with coordinates equal to the medians of the marginal distributions and is shown as the intersection of the vertical and horizontal lines. Panel (b) shows the sign that λ_i would have depending on which quadrant the point (x_i, y_i) is located. In panel (c) we want to calculate the distance from the point (x_i, y_i) to the center of the distribution. This distance is not the usual Euclidean distance but rather the maximum of the squares of the distance from the marginal distributions F_i and G_i.

In Figure 5.3 we have 200 random points (x_i, y_i) taken from a bivariate normal distribution with mean $\mu = (0, 0)$, variance $\sigma^2 = (1, 1)$ and correlation coefficient equal to 0. In this case, we know that X and Y are independent and so the χ-plot (see Figure 5.3) should show that most of the points fall within the control bands.

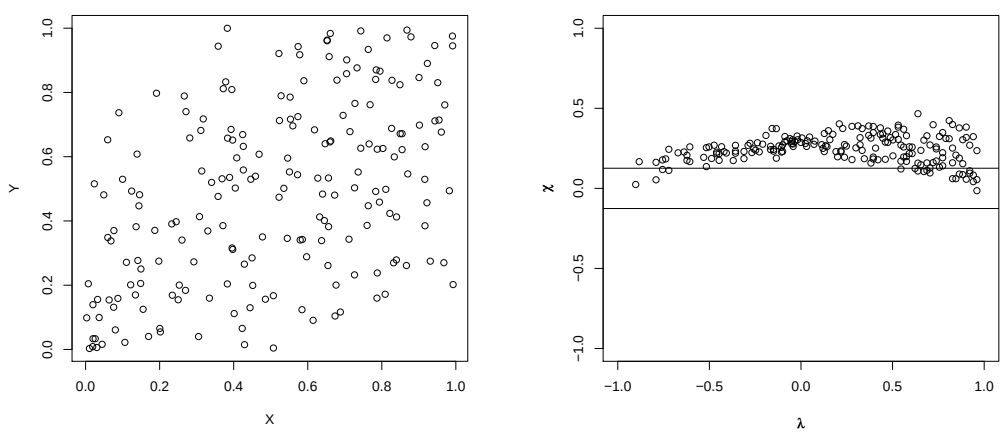
Now consider again 200 points sampled at random from a standard bivariate normal distribution with correlation coefficient equal to 0.5 as shown in Figure 5.4. These points now have a monotone positive association and the χ-plot shows that as a pattern that increases from the point $(-1, 0)$ towards the point $(0, 0.5)$ and then decreases as λ reaches the value +1. The majority of the points are now above the line $\chi = 0$ signaling positive association.

Figure 5.3
Bivariate Normal with Zero Correlation



The left panel displays the scatterplot of 200 random points from a bivariate normal distribution with mean $\mu = (0, 0)$, variance $\sigma^2 = (1, 1)$ and correlation coefficient equal to 0. The right-hand panel shows the corresponding χ-plot. Note that the majority of the points fall within the 95 percent control limits.

Figure 5.4
Bivariate Normal with 0.5-Correlation



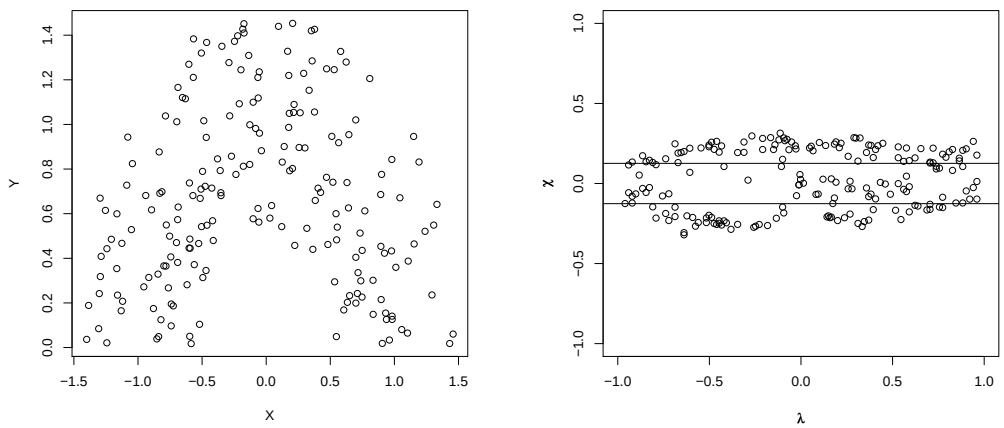
The left panel displays the scatterplot of 200 random points from a bivariate normal distribution with mean $\mu = (0, 0)$, variance $\sigma^2 = (1, 1)$ and correlation coefficient equal to 0.5. The right-hand panel shows the corresponding χ-plot. Most of the points are outside the control lines and above the line $\chi = 0$ which indicates positive dependence. Note how the peak is near the point $(0, 0.5)$.

For the next example, taken from [10, example 4], consider a set of points where there is no monotone association present.[7] The data consists of 200 points taken from the

[7] We have slightly modified the example by plotting twice as many points as in the original.

standard bivariate normal distribution with zero correlation and satisfying the restrictions $Y \geq 0$ and $|X^2 + Y^2 - 1| \leq 1/2$. The scatterplot and the χ-plot are shown in Figure 5.5.

Figure 5.5
Non-Monotone Association

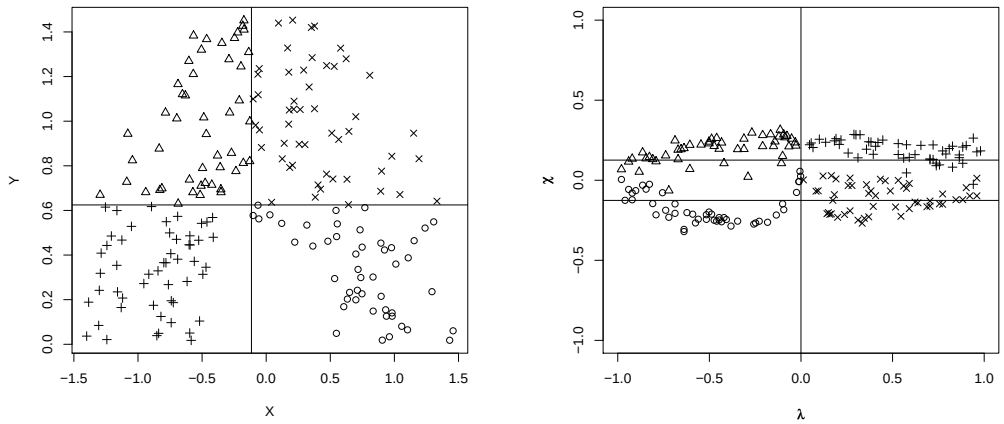


The left panel displays the scatterplot of 200 random points from a bivariate normal distribution with mean $\mu = (0, 0)$, variance $\sigma^2 = (1, 1)$ and zero correlation satisfying the constraints $Y \geq 0$ and $|X^2 + Y^2 - 1| \leq 1/2$. The right-hand panel shows the corresponding χ-plot.

The χ-plot shows that there are many points outside the control limits. This indicates that the original variables are not independent. It also shows a more complex pattern compared to the previous examples. In particular, there are four distinct regions in the χ-plot roughly corresponding to the four quadrants defined by the median point of the X and Y marginal distributions. In Figure 5.6 we have added a vertical and horizontal line passing through the median of the marginal distributions and used different symbols to plot in each quadrant. Note that in the lower-left and upper-left quadrants the association between X and Y is positive and these points appear predominantly above the line $\chi = 0$ in the χ-plot. Similarly, the points in the lower-right and upper-right quadrants of the scatterplot appear mainly below the line $\chi = 0$ in the χ-plot.

Appendices A–E show scatter- and χ-plots for various copulas under different parameters. This catalog serves as a good starting point to compare against an empirical dataset. In general, to determine the best copula for a given application we would use both formal and informal selection methods. For some of the formal techniques consult [18, 6, 20]. One of the informal techniques is to look at the χ-plot of the empirical data and compare it against the simulated χ-plots for various copula families to find an appropriate match.

Figure 5.6
Non-Monotone Association



This is the same as Figure 5.5 but we have added horizontal and vertical lines at the medians of the marginal distributions. In each of the four quadrants we have used different plotting symbols. These symbols have been carried over to the χ-plot.

5.2 Parameter Estimation via Maximum Likelihood

Once we have selected the appropriate pair-copula families we can proceed with the estimation of the parameters via maximum likelihood. Suppose we have an n-dim distribution function along with T observations. Let $\mathbf{x}_s$ denote the vector of observations for the s-th point with $s = 1, 2, \ldots, T$. The likelihood function for a canonical vine decomposition is

$$L(\mathbf{x};\Theta) = \prod_{s=1}^{T}\prod_{k=1}^{n} f(x_{s,k}) \cdot \prod_{j=1}^{n-1}\prod_{i=1}^{n-j} c_{j,j+i|1,\ldots,j-1}(F(x_{s,j}|x_{s,1},\ldots,x_{s,j-1}), F(x_{s,j+i}|x_{s,1},\ldots,x_{s,j-1})). \tag{30}$$

By taking logarithms and assuming that each of the n marginals is uniform on the unit interval[8]; that is, $f(x_{s,k}) = 1$ for all s and k, the log-likelihood function is

$$\ell(\mathbf{x};\Theta) = \sum_{s=1}^{T}\sum_{j=1}^{n-1}\sum_{i=1}^{n-j} \log\Big(c_{j,j+i|1,\ldots,j-1}(F(x_{s,j}|x_{s,1},\ldots,x_{s,j-1}), F(x_{s,j+i}|x_{s,1},\ldots,x_{s,j-1}))\Big) \tag{31}$$

Similarly, the likelihood function for a D-vine is

[8] For simplicity we are using uniform margins. Extending our discussion to non-uniform margins is straightforward.

$$L(\mathbf{x};\Theta) = \prod_{s=1}^{T}\prod_{k=1}^{n} f(x_{s,k}) \cdot$$
$$\prod_{j=1}^{n-1}\prod_{i=1}^{n-j} c_{i,j+i|i+1,\ldots,i+j-1}(F(x_{s,i}|x_{s,i+1},\ldots,x_{s,i+j-1}), F(x_{s,j+i}|x_{s,i+1},\ldots,x_{s,i+j-1})), \tag{32}$$

and the log-likelihood (assuming uniform marginals) is

$$\ell(\mathbf{x};\Theta) = \sum_{s=1}^{T}\sum_{j=1}^{n-1}\sum_{i=1}^{n-j} \log\Big(c_{i,i+j|i+1,\ldots,i+j-1}(F(x_{s,i}|x_{s,i+1},\ldots,x_{s,i+j-1}), F(x_{s,i+j}|x_{s,i+1},\ldots,x_{s,i+j-1}))\Big). \tag{33}$$

For each copula term in the log-likelihood (31) or (33) we have at least one parameter to estimate.[9]

The conditional distributions for a canonical vine

$$F(x_{s,j}|x_{s,1},\ldots,x_{s,j-1}) \quad \text{and} \quad F(x_{s,j+i}|x_{s,1},\ldots,x_{s,j-1})$$

or the conditional distributions for a D-vine

$$F(x_{s,i}|x_{s,i+1},\ldots,x_{s,i+j-1}) \quad \text{and} \quad F(x_{s,i+j}|x_{s,i+1},\ldots,x_{s,i+j-1})$$

are again determined by using the recursive relation (16) and the appropriate $h(\cdot)$ function (21). Then we can use numerical optimization techniques to maximize the log-likelihood over all parameters simultaneously.

Algorithm 3 (shown in Appendix F), from [1], computes the log-likelihood function of a canonical vine for a given set of observations. The numerical maximization can be carried out via the Nelder-Mead algorithm [21] or another optimization technique.

Starting values for the numerical maximization of the log-likelihood can be determined as follows:

1. Estimate the parameters of the copulas in tree 1 from the original data. That is, fit each bivariate copula to the observations. This estimation is easy to do since we are only looking at two dimensions at a time.
2. Compute the implied observations for tree 2 using the copula parameters from tree 1 and the appropriate $h(\cdot)$ functions.
3. Estimate the parameters of the copulas in tree 2 from the observations in step 2.

[9] The number of parameters depends on the bivariate copula chosen. Many bivariate copula families have one parameter, but others have two or more, such as the t-copula.

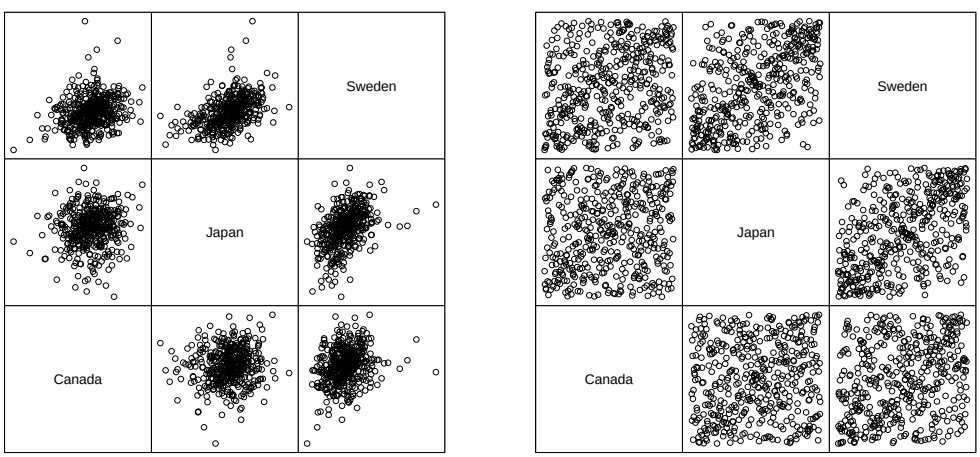
4. Compute the implied observations for tree 3 using the copula parameters from step 3 and the appropriate $h(\cdot)$ functions.
5. Continue along the sequence of trees in the pair-copula decomposition.

These starting values can then be passed to the full numerical maximization routine.

6. Illustration Based on Currency Rate Changes

In [26] various multivariate copula models were fitted to currency rate changes. We will use the same currencies, but for a longer observation period, to illustrate the pair-copula construction. The raw data consist of monthly rates of exchange between the Canadian, Japanese and Swedish currencies and the U.S. dollar. This raw data spans from January 1971 to July 2007, and the source is the Fred database of the Federal Reserve Bank of St. Louis [8]. As in [26] we will apply the pair-copula construction to the monthly changes in the rate of exchange. The left-hand panel of Figure 6.1 shows the scatterplot matrix of this dataset.

Figure 6.1
Changes in Currency Exchange Rates



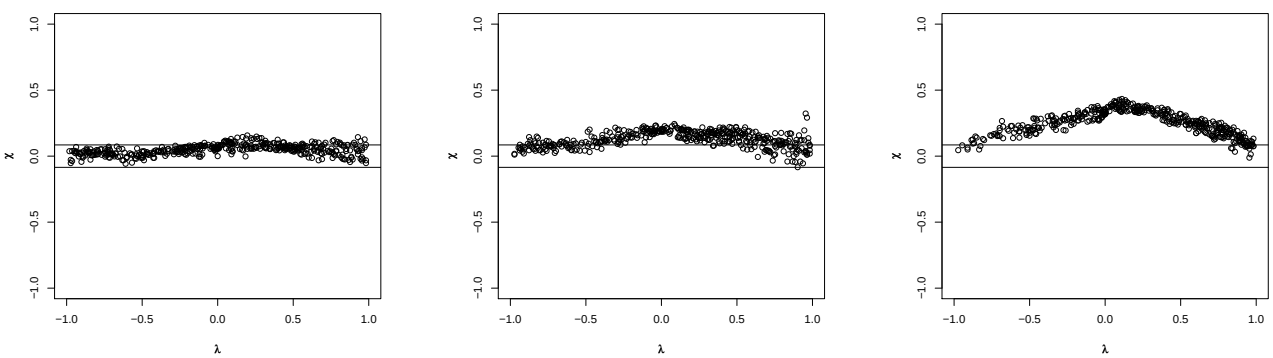
The left-hand panel shows the monthly changes in the foreign exchange rates between the Canadian, Japanese and Swedish currencies against the U.S. dollar for the period January 1971 to July 2007. The right-hand panel shows the ranked transformed data.

We are interested in understanding the dependence between these variables and so before we proceed any further we will remove the marginal distributions from our analysis; that is, we are only interested in the ranks of our dataset. Thus in the right-hand panel of Figure 6.1 we have transformed our data via its empirical distribution function. If (x_i, y_i) is a point in one of the scatterplots on the left-hand panel, then the corresponding point

on the right-hand panel is $\left(\widehat{F}_x(x_i), \widehat{F}_y(y_i)\right)$; where $\widehat{F}_x$ and $\widehat{F}_y$ are the empirical marginal distribution functions.

Figure 6.2 shows the three pairwise χ-plots based on the currency data shown in Figure 6.1. Note that all three plots are very different. Most of the points on the Canada–Japan χ-plot (left-most panel) are within the control bands implying that these two variables are slightly positively dependent but not too far from being independent. But notice that within the control bands the points are not randomly scattered. Rather they seem to steadily increase from about $\lambda = -0.5$ to $\lambda = 0.15$ and then they decrease as λ continues towards the right. Also note that as we reach the right edge the points seem to scatter more than before.

Figure 6.2
χ-plots for Changes in Currency Exchange Rates



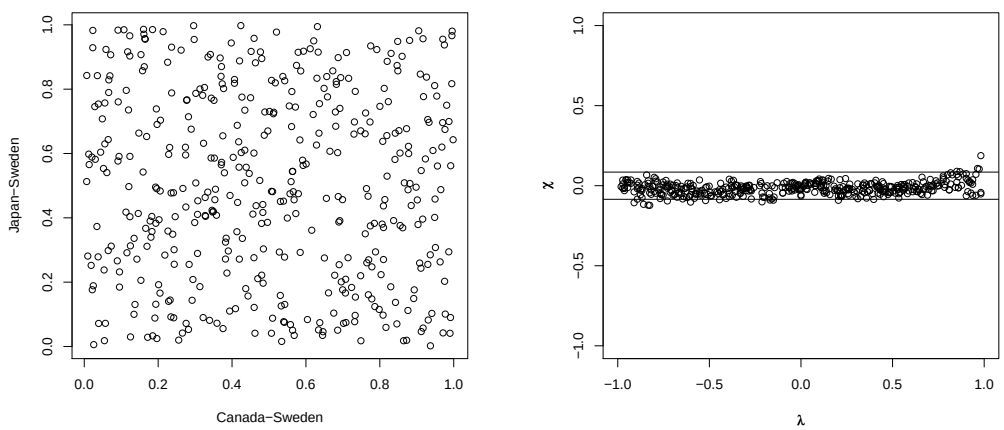
Canada–Japan Canada–Sweden Japan–Sweden

The left-most panel shows that there is a mild association between the Canadian and Japanese data. Note that there is a slight up and down pattern with increasing spread as λ approaches 1 along the horizontal axis. The middle panel clearly shows that the Canadian and Swedish data are positively associated and note the increasing spread as we approach $\lambda = 1$ on the horizontal axis. In the right most panel the dependence between the Japanese and Swedish data is strong. Note the peak around $\lambda = 0.1$ and how as λ approaches 1 on the horizontal axis there is no appreciable spread. This behavior is opposite from the other two panels.

On the middle panel of Figure 6.2 we clearly see that the monthly changes in the U.S. exchange rate between the Canadian and Swedish currencies are positively associated as most points are outside the control band and above the line $\chi = 0$. Again we have an increasing trend from $\lambda = -1$ to $\lambda = 0$ and then a decreasing trend as λ reaches the right-hand side of the plot. Note also the increasing spread as λ moves from 0 to 1. The highest point is approximately at $(\lambda, \chi) = (0, 0.15)$, and so we expect Kendall's τ between these two variables to be approximately equal to 0.15.

The last panel in Figure 6.2 shows that the strongest positive association is between the Japanese and Swedish currencies. Note that the highest point is approximately at $\chi = 0.35$, and so we suspect that Kendall's measure of concordance would be about that value. As with the previous panel we see the steady increase as we approach $\lambda = 0$ and then the decrease as we continue towards the right. But notice that we do not see the increasing spread as λ approaches 1. This indicates that for values far away from the center of distribution (λ near 1) there seems to be no dependence among them. This feature of the χ-plot rules out copulas such as the Gumbel or Galambos families.

Figure 6.3
Implied Data for Tree 2 Construction



The left-hand panel shows the scatterplot for the implied Canada–Sweden, Japan–Sweden data. The right-hand panel shows the corresponding χ-plot.

The currency pairs Canada–Sweden and Sweden–Japan have strong positive dependence and so we will use them to build our pair-copula construction. We need to select a copula family that adequately represents the dependence between these pairs. Comparing the empirical χ-plot for the pair Canada–Sweden with the copula families in the appendices we can see that the Gumbel or Galambos families might be good candidates as their χ-plots resemble the empirical one. We further explored this via simulation (and other more formal techniques [6, 13]) and concluded that the Galambos family has features that are not consistent with the empirical data. The Gumbel simulations provided evidence that this family adequately describes the empirical data. The maximum likelihood estimate for the parameter for the Gumbel family is 1.17. This estimate is based only on the bivariate data Canada–Sweden and not on the full dataset. We will only use it as a starting point for a full maximum likelihood estimation.

A similar investigation between the Japanese and Swedish data led us to model their dependence with a Frank copula. The maximum likelihood estimate of the parameter is 3.45. Again this estimate will only be used as a starting point for the full maximum likelihood estimation. We now have initial parameter estimates for the first tree in our canonical vine decomposition. The second tree consists of two nodes and one edge. The conditional pair-copula that we need to specify here is for the *Canadian–Japanese given the Swedish* data. To this end we compute the implied observations for tree 2 from the Gumbel and Frank copulas used in tree 1. Figure 6.3 shows the scatterplot and χ-plot for the Canada–Sweden, Japan–Sweden implied data.

From the χ-plot in Figure 6.3 it is clear that the relationship Canada–Japan *given Swedish* data should be modeled by the independence copula.

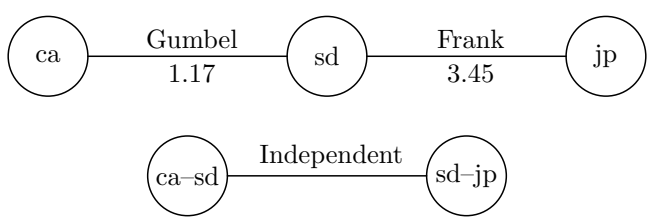
The canonical vine structure for the currency rate changes dataset is shown in Figure 6.4. These parameter estimates are only starting values to be used in a full maximum likelihood estimation. Using Algorithm 3 with our dataset and our chosen pair-copulas in an optimization routine to maximize the log-likelihood we arrive at the maximum likelihood parameter estimates in Table 6.1.

Table 6.1
Maximum Likelihood Parameter Estimates

Pair-copula	Family	ML estimate
Canada–Sweden	Gumbel	1.11
Japan–Sweden	Frank	1.62
Canada–Japan *given* Sweden	Independent	

Full model maximum likelihood parameter estimates for the canonical vine used to model the currency rate changes dataset.

Figure 6.4
Initial Parameter Estimates for Maximum Likelihood



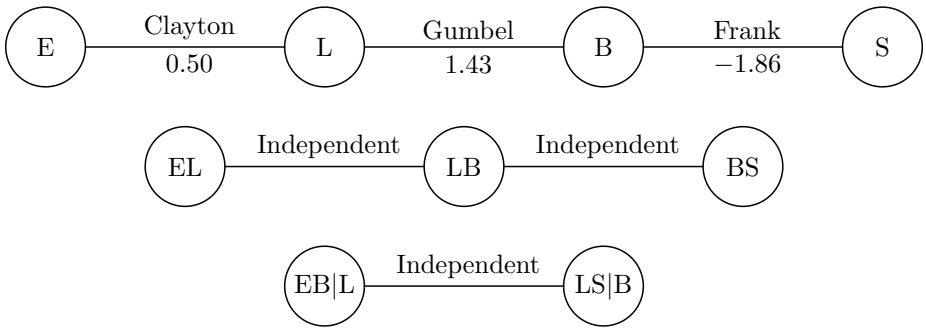
The canonical vine used to model the Canadian (ca), Japanese (jp), and Swedish (sd) currency rate changes along with the chosen pair-copula family and the individually estimated maximum likelihood parameters. These parameters are only used as the starting values for the full model maximum likelihood estimation procedure.

7. Illustration Based on Simulated Insurance Assets and Liabilities

In the context of enterprise risk management (ERM) one would like to either understand the dependencies among given risk factors or be able to simulate data that exhibits specific dependence traits. In this section we will illustrate how one can use the simulation algorithms of Appendix F to generate complex dependencies.

Suppose we would like to generate data for two asset classes (bonds and stocks) and a liability portfolio (say, losses and expenses). For simplicity and because we want to highlight the dependency structure we will work with uniform margins but extending our discussion to other marginal distributions is straightforward. Let us use a simple D-vine structure as our starting point. Figure 7.1 shows the copula families chosen and their parameters.

Figure 7.1
Initial D-Vine Structure

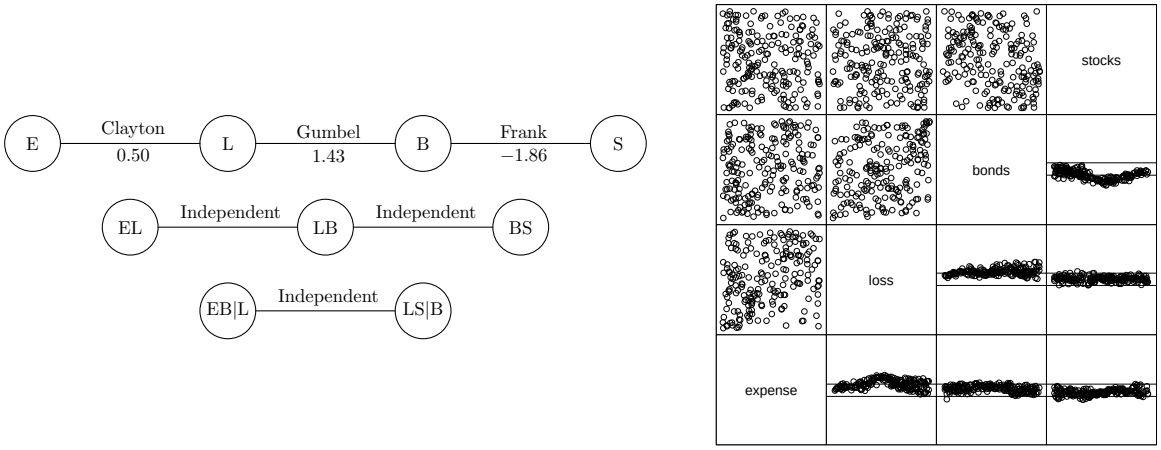


This is a four-dimensional D-vine structure coupling one liability portfolio and two asset classes. The liabilities are losses (L) and expenses (E) and the assets are bonds (B) and stocks (S). We have deliberately kept this structure simple by choosing the independent copula for trees two and three. We have chosen to model the dependence between the loss and expense with a Clayton copula. The asset classes are linked via a Frank copula, and the losses are coupled with a Gumbel copula to the bonds. The parameters of these copulas appear beneath the names. The independence copula does not have any parameters.

The copula used to link the expenses to the losses is Clayton with a parameter chosen so that the Kendall's τ between these two variables is equal to 0.2. Losses are linked to bonds via a Gumbel copula with parameter chosen so that Kendall's τ is equal to 0.3, and finally bonds and stocks use a Frank copula with parameter -1.86; that is, the Kendall τ coefficient is approximately equal to -0.2.

The left-hand panel of Figure 7.2 displays the D-vine structure, and the right-hand panel shows scatterplots and corresponding χ-plots for a random sample of 200 points. Notice how the χ-plots next to the main diagonal have the expected behavior given the D-vine structure. The remaining three χ-plots also follow the assumptions of independence.

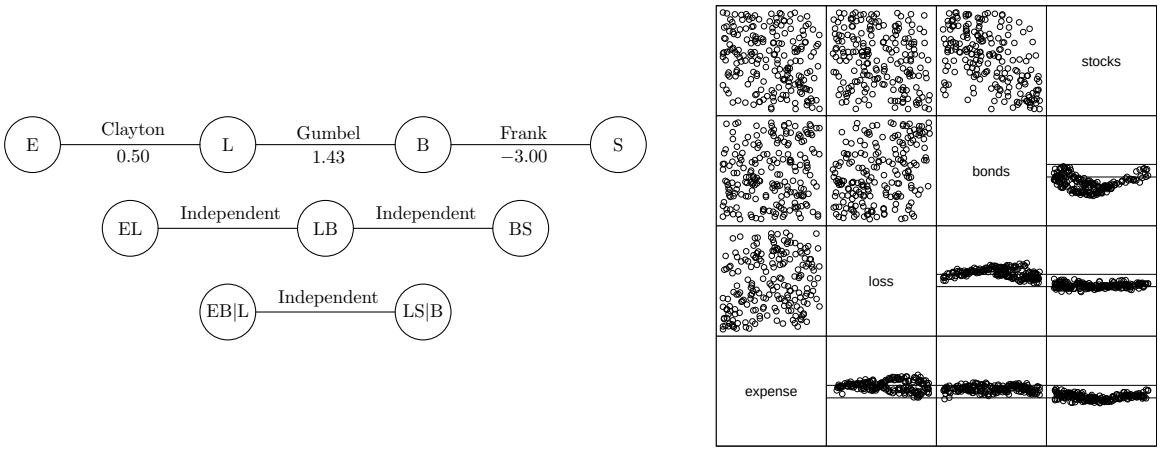
Figure 7.2
Simulated Insurance Assets and Liabilities I



The left-hand panel shows the D-vine structure along with the parameters for each copula and the right-hand panel has a random sample of 200 points from this D-vine. The upper portion of the graph show the pairwise scatterplots and the lower portion displays the corresponding χ-plot.

In the next figure we change one of the parameters of our D-vine structure. For the Frank copula (linking bonds and stocks) let us increase the strength of the dependence from about a Kendall's τ value of -0.2 to a value of -0.3. Figure 7.3 shows the new D-vine structure and the simulated data.

Figure 7.3
Simulated Insurance Assets and Liabilities II

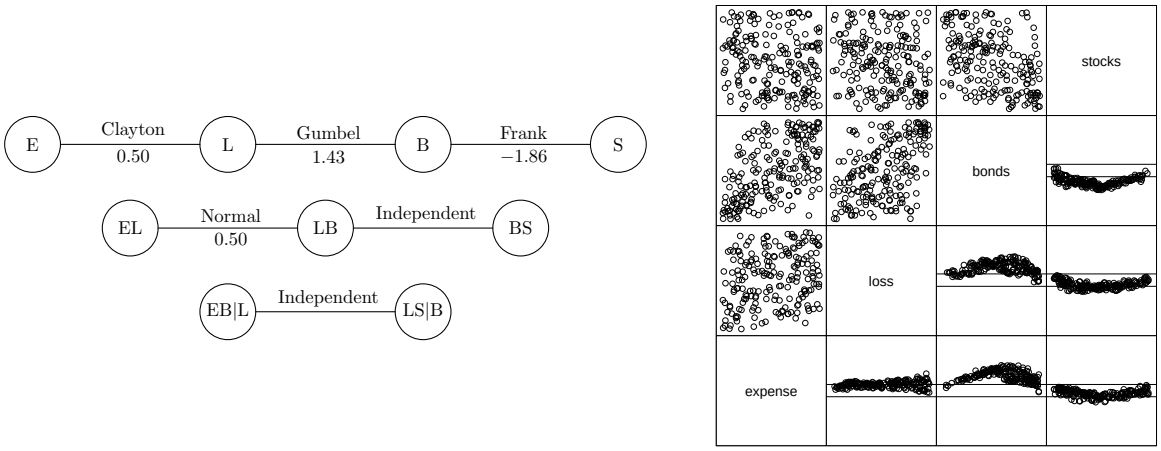


The left-hand panel shows the D-vine structure along with the parameters for each copula. The right-hand panel displays a random sample of 200 points from this D-vine. The upper portion of the graph shows the pairwise scatterplots and the lower portion displays the corresponding χ-plots.

First notice that, as expected, the bonds–stocks χ-plot shows a stronger negative dependence. Also this change has influenced the other χ-plots; in particular, the loss–expense pair the conditional expense–stocks pair given loss and bonds.

Now let us make a change to one of the conditional copulas in our structure; that is, one of the copulas linking two variables given others. In Figure 7.4 we have changed the copula linking expenses to bonds given losses from independent to normal with a parameter equal to 0.5. Again comparing this new simulated data to the one in Figure 7.2 we see that the loss–bonds dependence has increased as well as the loss–stocks given bonds.

Figure 7.4
Simulated Insurance Assets and Liabilities III



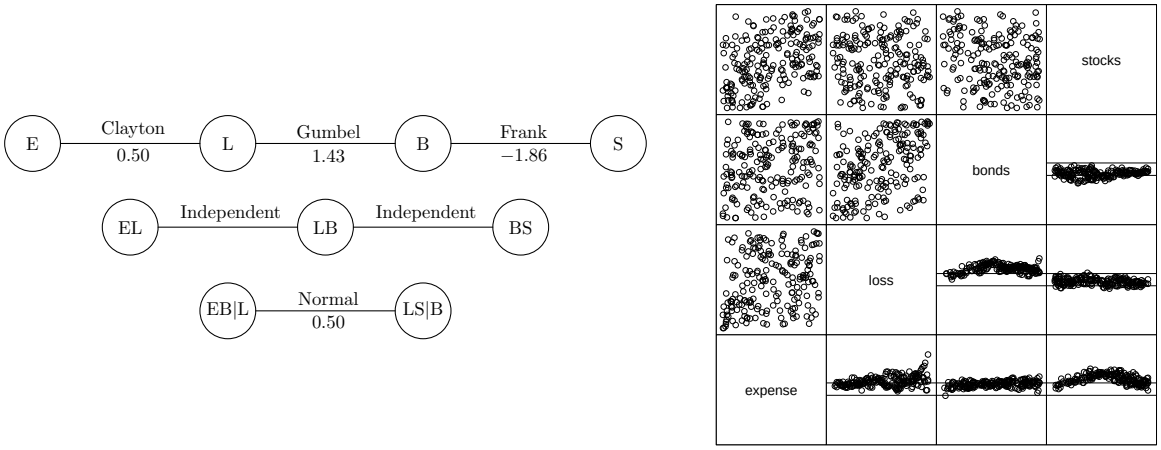
The left-hand panel shows the D-vine structure along with the parameters for each copula, and the right-hand panel has a random sample of 200 points from this D-vine. The upper portion of the graph show the pairwise scatterplots, and the lower portion displays the corresponding χ-plot.

Finally, the last change involves the conditional copula linking expenses and stocks given loss and bonds. We remove the independent copula and put in its place a normal copula with parameter equal to 0.5.

In this last case, the bottom-right χ-plot shows the expected behavior. But the adjacent χ-plots do not differ significantly from those in the original Figure 7.2. Surprisingly the loss–expense χ-plot now has a different feel to it. For λ near 1 we see increasing spread. Also the dependence between bonds and stocks seems to be weaker.

As we gain more experience with the canonical and D-vines, the selection of pair-copulas, and how they interact with each other, we will be able to build high-dimensional models that better capture the most relevant aspects of the problem in question.

Figure 7.5
Simulated Insurance Assets and Liabilities IV



The left-hand panel shows the D-vine structure along with the parameters for each copula, and the right-hand panel has a random sample of 200 points from this D-vine. The upper portion of the graph show the pairwise scatterplots, and the lower portion displays the corresponding χ-plot.

References

[1] Aas, K., Czado, C., Frigessi, A., and Bakken, H. 2006. "Pair-copula constructions of multiple dependence" Tech. Rep. SAMBA/24/06, Oslo, Norway: Norwegian Computing Center.
http://www.nr.no/files/samba/bff/SAMBA2406.pdf

[2] Anscombe, F. J. 1973. "Graphs in statistical analysis" The American Statistician 27(1): 17–21.

[3] Brehm, P. J., et al. 2007. *Enterprise Risk Analysis for Property & Liability Insurance Companies*. Guy Carpenter & Company.

[4] Bedford, T., and Cooke, R. M. 2002. "Vines: A new graphical model for dependent random variables" The Annals of Statistics 30(4): 1031–1068.

[5] Charpentier, A., and Segers, J. 2006. "Lower tail dependence for archimedean copulas: Characterizations and pitfalls" Discussion Paper 29, Tilburg University, Center for Economic Research.
http://ideas.repec.org/p/dgr/kubcen/200629.html

[6] Durrleman, V., Nikeghbali, A., and Roncally, T. 2000. "Which copula is the right one?" Tech. Rep., Crédit Lyonnais.
http://ssrn.com/abstract=103245

[7] Embrechts, P., Lindskog, F., and McNeil, A. 2003. "Modelling dependence with copulas and applications to risk management" in *Handbook of Heavy Tailed Distributions in Finance*, Rachev, S. T., ed. North Holland: Elsevier.

[8] Federal Reserve Bank of St. Louis.
http://research.stlouisfed.org/fred2

[9] Faivre, F. 2003. "Copula: A new vision for economic capital and applications to a four line of business company" Tech. Rep., ASTIN Colloquium.
http://www.actuaries.org/ASTIN/Colloquia/Berlin/Faivre.pdf

[10] Fisher, N. I., and Switzer, P. 1985. "Chi-plots for assessing dependence" Biometrika 72(2): 253–265.

[11] Fisher, N. I., and Switzer, P. 2001. "Graphical assessment of dependence: Is a picture worth 100 tests?" The American Statistician 55(3): 233–239.

[12] Fox, J. 1991. "Regression diagnostics: An introduction" Vol. 07–079 of *Sage University Paper Series on Quantitative Applications in the Social Sciences.* Newbury Park, Calif.: Sage Publications.

[13] Frees, E. W., and Valdez, E. A. 1998. "Understanding relationships using copulas" North American Actuarial Journal 2(1): 1–25.
http://www.soa.org/library/journals/north-american-actuarial-journal/1998/january/naaj9801_1.pdf

[14] Genest, C., and Favre, A.-C. 2007. "Everything you always wanted to know about copula modeling but were afraid to ask" Journal of Hydrologic Engineering 12(4): 347–368.
http://archimede.mat.ulaval.ca/pages/genest/publi/JHE-2007.pdf

[15] Genest, C., and MacKay, J. 1986. "The joy of copulas: Bivariate distributions with uniform marginals" The American Statistician 40(4): 280–283.
http://archimede.mat.ulaval.ca/pages/genest/publi/TAS-1986.pdf

[16] Genest, C., and Rivest, L.-P. 1993. "Statistical inference procedures for bivariate archimedean copulas" Journal of the American Statistical Association 88: 1034–1043.

[17] Joe, H. 1997a. "Families of m-variate distributions with given margins and $m(m-1)/2$ bivariate dependence parameters," in *Distributions with Fixed Marginals and Related Topics*, Rüschendorf, L., Schweizer, B., and Taylor, M. D., eds.

[18] Joe, H. 1997b. *Multivariate Models and Dependence Concepts*, Vol. 73 of *Monographs on Statistics and Applied Probability*. Boca Raton, Fla.: Chapman & Hall/CRC.

[19] Klugman, S. A., and Parsa, R. 1999. "Fitting bivariate loss distributions with copulas" Insurance: Mathematics and Economics 24(1–2): 139–148.
http://ideas.repec.org/a/eee/insuma/v24y1999i1-2p139-148.html

[20] Mari, D. D., and Kotz, S. 2004. *Correlation and Dependence*. London: Imperial College Press.

[21] Nelder, J. A., and Mead, R. 1965. "A simplex algorithm for function minimization" Computer Journal 7: 308–313.

[22] Nelsen, R. B. 1999. *An Introduction to Copulas*, Vol. 139 of *Lecture Notes in Statistics*. New York: Springer-Verlag.

[23] Schweizer, B., and Sklar, A. 2005. *Probabilistic Metric Spaces*. Mineola, N.Y.: Dover Publications.

[24] Sklar, A. 1959. "Fonctions dé repartition á n dimensions et leurs marges" Publ. Inst. Stat. Univ. Paris 8: 229–231.

[25] Valdez, E. A., and Tang, A. 2005. "Economic capital and the aggregation of risks using copulas" Tech. Rep., University of New South Wales.
http://www.gloriamundi.org/detailpopup.asp?ID=453058136

[26] Venter, G., Barnett, J., Kreps, R., and Major, J. 2007. "Multivariate copulas for financial modeling" Variance 1(1): 103–119.

[27] Venter, G. G. 2002. "Tails of copulas" Proceedings of the Casualty Actuarial Society 89: 68–113.
http://www.casact.org/pubs/proceed/proceed02/02068.pdf

[28] Venter, G. G. 2003. "Quantifying correlated reinsurance exposures with copulas" Casualty Actuarial Society Forum: 215–229.
http://www.casact.org/pubs/forum/03spforum/03spf215.pdf

Appendix A. Bivariate Clayton Copulas

The bivariate Clayton copulas are given for $0 \leq \delta < \infty$ by

$$C(u, v; \delta) = (u^{-\delta} + v^{-\delta} - 1)^{-1/\delta}. \tag{34}$$

The density function is

$$c(u, v; \delta) = (1 + \delta)[uv]^{-\delta-1}(u^{-\delta} + v^{-\delta} - 1)^{-2-1/\delta}. \tag{35}$$

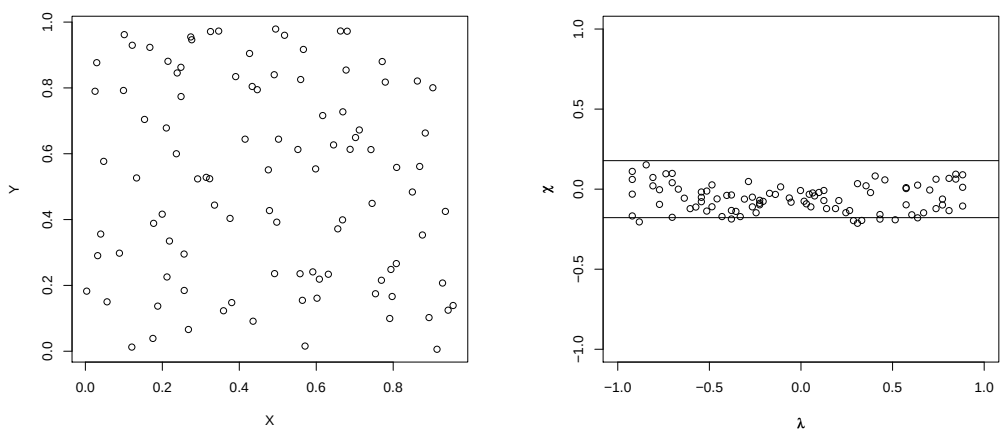
The h and h^{-1} functions are

$$h(u, v; \delta) = v^{-\delta-1}\left(u^{-\delta} + v^{-\delta} - 1\right)^{-1-1/\delta}$$

$$h^{-1}(u, v; \delta) = \left[\left(uv^{\delta+1}\right)^{-\delta/(1+\delta)} + 1 - v^{-\delta}\right]^{-1/\delta}.$$

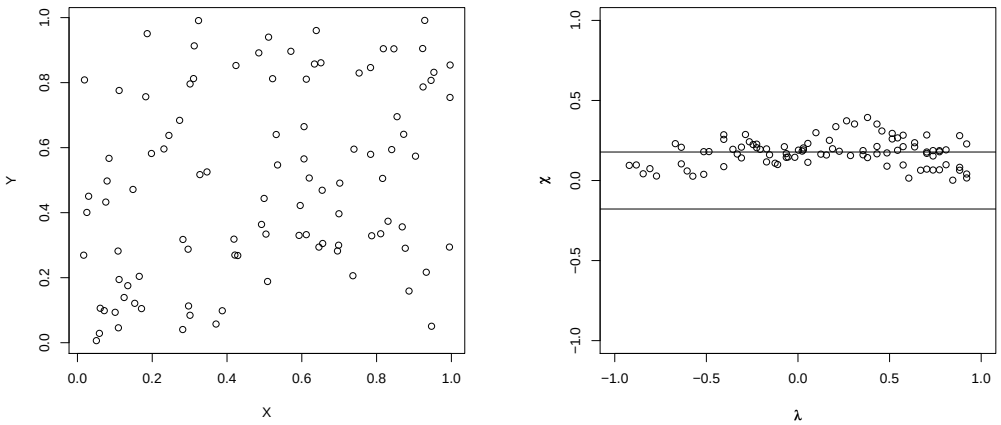
The derivation of these formulas is in [1].

Figure A.1
Clayton Copula 0-Correlation



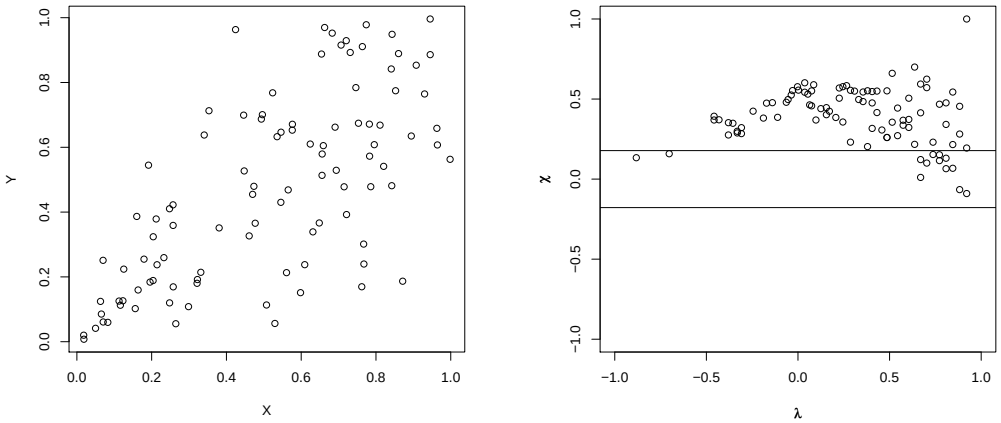
The left panel displays the scatterplot of a Clayton copula with a correlation coefficient of zero. The right-hand panel shows the corresponding χ-plot. Notice that most of the points are within the 95 percent control bands.

Figure A.2
Clayton Copula 0.2-Correlation



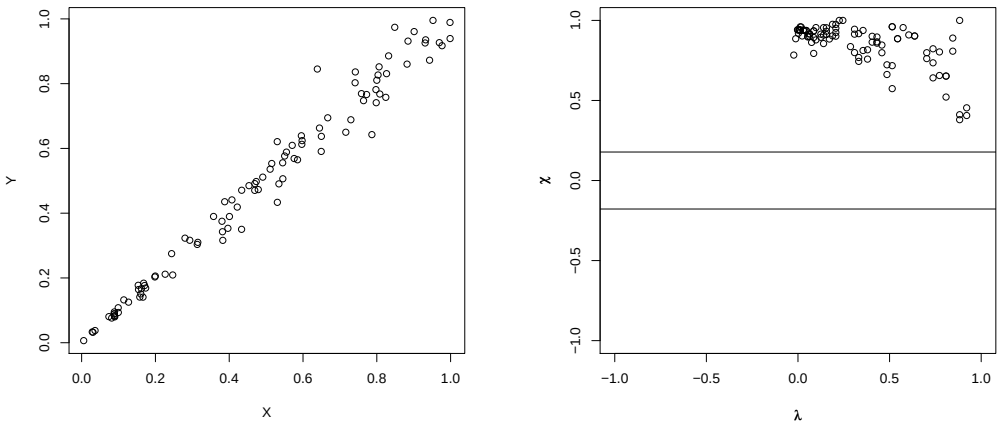
The left panel displays the scatterplot of a Clayton copula with a correlation coefficient equal to 0.2. The right-hand panel shows the corresponding χ-plot.

Figure A.3
Clayton Copula 0.5-Correlation



The left panel displays the scatterplot of a Clayton copula with a correlation coefficient equal to 0.5. The right-hand panel shows the corresponding χ-plot.

Figure A.4
Clayton Copula 0.9-Correlation



The left panel displays the scatterplot of a Clayton copula with a correlation coefficient equal to 0.9. The right-hand panel shows the corresponding χ-plot.

Appendix B. Bivariate Frank Copulas

The bivariate Frank copula has parameter space $0 \leq \delta < \infty$ and distribution function

$$C(u, v; \delta) = -\delta^{-1} \log([\eta - (1 - e^{\delta u})(1 - e^{\delta v})]/\eta), \tag{36}$$

where $\eta = 1 - e^{-\delta}$. The density is

$$c(u, v; \delta) = \frac{\delta \eta e^{-\delta(u+v)}}{[\eta - (1 - e^{\delta u})(1 - e^{\delta v})]^2}, \tag{37}$$

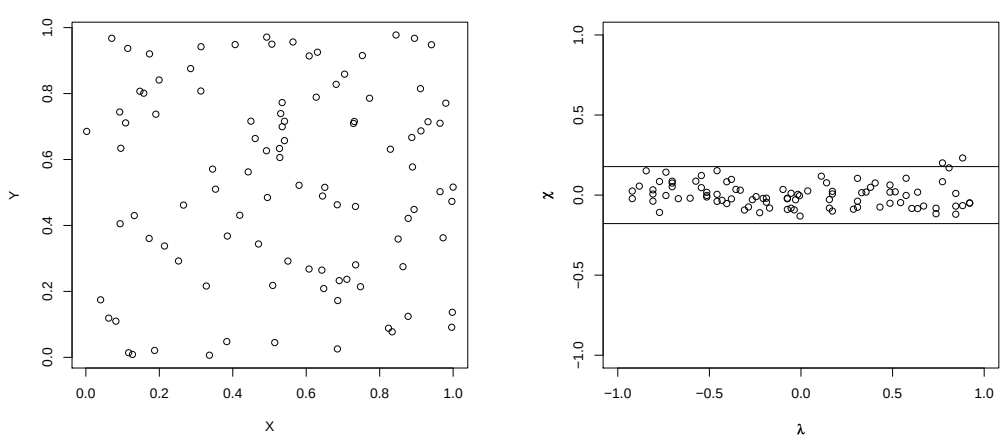
and the h and h^{-1} functions are

$$h(u, v; \delta) = \frac{e^{-\delta v}}{\frac{1-e^{-\delta}}{1-e^{-\delta u}} + e^{-\delta v} - 1}$$

$$h^{-1}(u, v; \delta) = -\log\left\{1 - \frac{1 - e^{-\delta}}{(u^{-1} - 1)e^{-\delta v} + 1}\right\} \Big/ \delta.$$

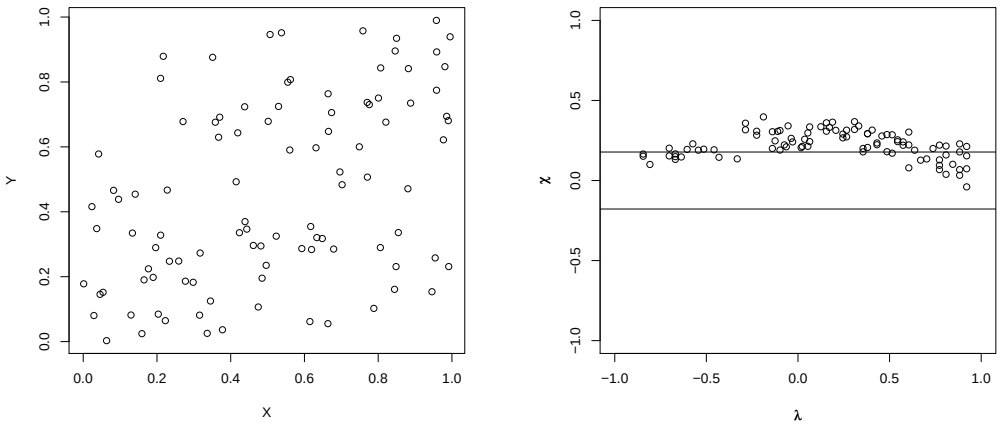
The derivation of these formulas is in [1].

Figure B.1
Frank Copula 0-Correlation



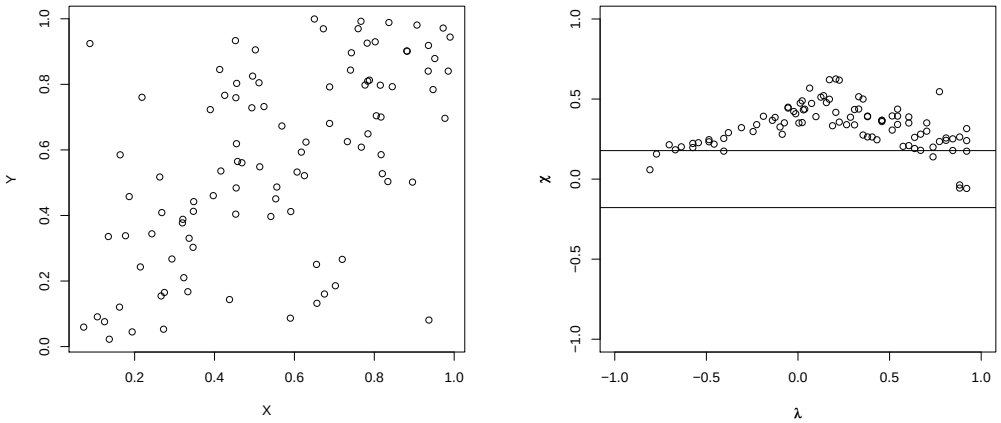
The left panel displays the scatterplot of a Frank copula with a correlation coefficient of zero. The right-hand panel shows the corresponding χ-plot. Notice that most of the points are within the 95 percent control bands.

Figure B.2
Frank Copula 0.2-Correlation



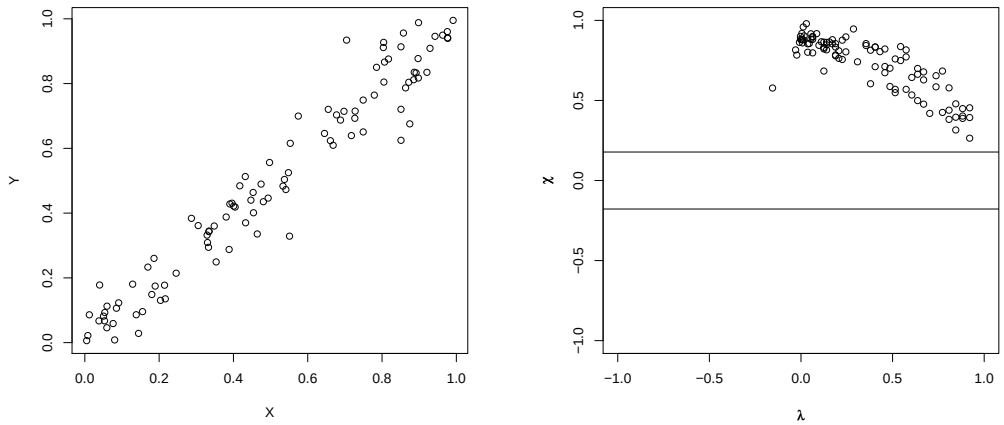
The left panel displays the scatterplot of a Frank copula with a correlation coefficient equal to 0.2. The right-hand panel shows the corresponding χ-plot.

Figure B.3
Frank Copula 0.5-Correlation



The left panel displays the scatterplot of a Frank copula with a correlation coefficient equal to 0.5. The right-hand panel shows the corresponding χ-plot.

Figure B.4
Frank Copula 0.9-Correlation



The left panel displays the scatterplot of a Frank copula with a correlation coefficient equal to 0.9. The right-hand panel shows the corresponding χ-plot.

Appendix C. Bivariate Galambos Copulas

Let $\tilde{u} = -\log u$ and $\tilde{v} = -\log v$ and for $0 \leq \delta < \infty$ the distribution function is

$$C(u, v; \delta) = uv \exp\left\{(\tilde{u}^{-\delta} + \tilde{v}^{-\delta})^{-1/\delta}\right\}, \tag{38}$$

and the density function is

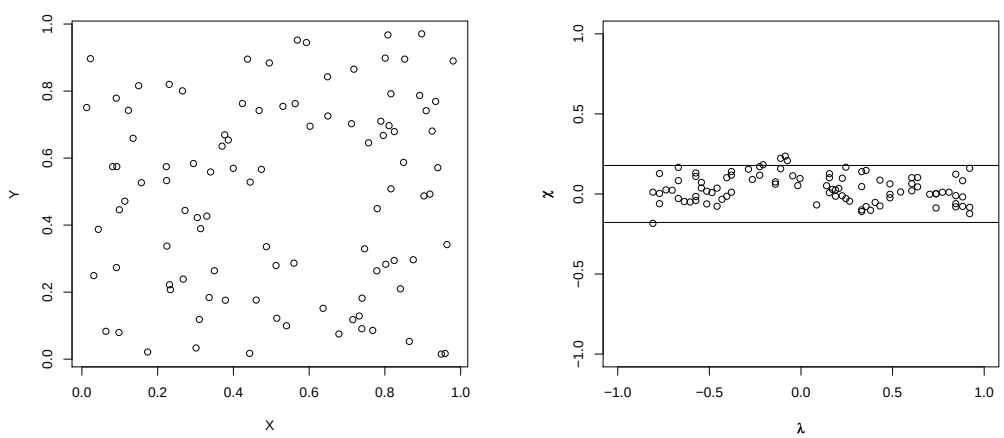
$$\begin{aligned} c(u, v; \delta) = C(u, v; \delta)(uv)^{-1}\Big[1 - \left(\tilde{u}^{-\delta} + \tilde{v}^{-\delta}\right)^{-1-1/\delta}\left(\tilde{u}^{-\delta-1} + \tilde{v}^{-\delta-1}\right) \\ + \left(\tilde{u}^{-\delta} + \tilde{v}^{-\delta}\right)^{-2-1/\delta}(\tilde{u}\tilde{v})^{-\delta-1}\left\{1 + \delta + \left(\tilde{u}^{-\delta} + \tilde{v}^{-\delta}\right)^{-1/\delta}\right\}\Big]. \end{aligned}$$

The conditional density function $h(u, v; \delta)$ is given by

$$h(u, v; \delta) = \frac{C(u, v; \delta)}{v}\left(1 - \left[1 + \left(\tilde{v}/\tilde{u}\right)^{\delta}\right]^{-1-1/\delta}\right). \tag{39}$$

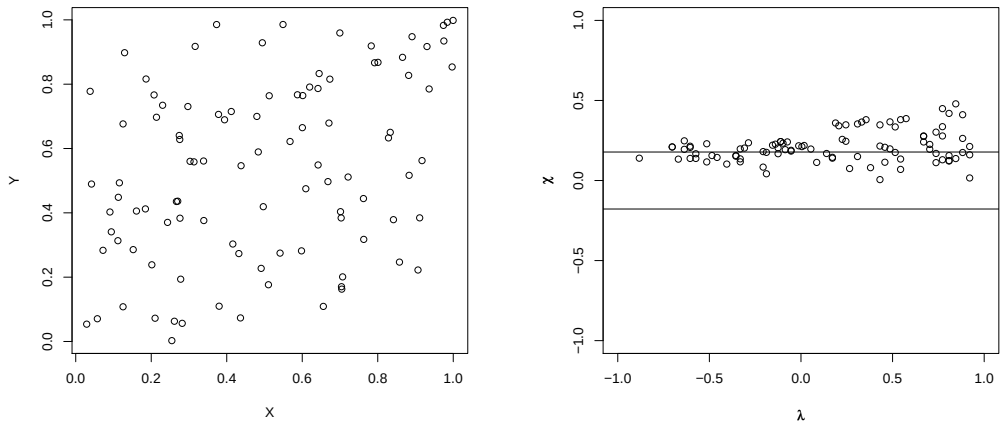
The inverse function, h^{-1}, cannot be written down in closed form but we can find its inverse numerically. The derivation of these formulas is in [1].

Figure C.1
Galambos Copula 0-Correlation



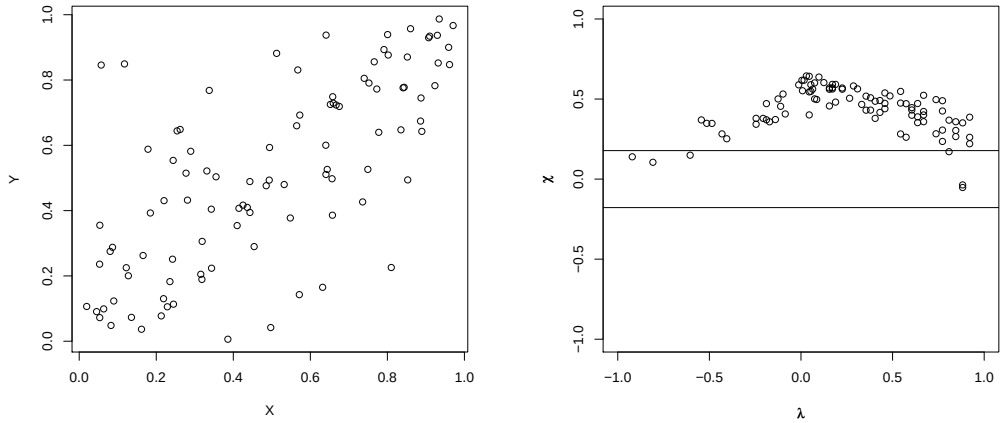
The left panel displays the scatterplot of a Galambos copula with a correlation coefficient of zero. The right-hand panel shows the corresponding χ-plot. Notice that most of the points are within the 95 percent control bands.

Figure C.2
Galambos Copula 0.2-Correlation



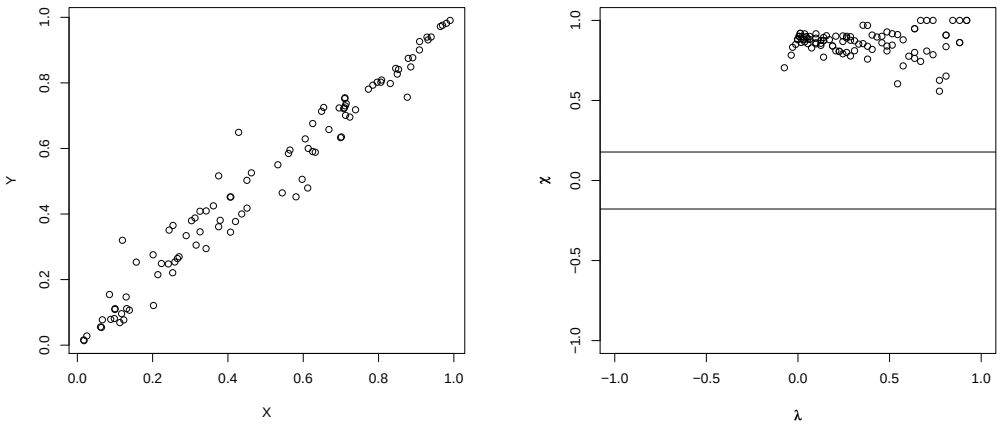
The left panel displays the scatterplot of a Galambos copula with a correlation coefficient equal to 0.2. The right-hand panel shows the corresponding χ-plot.

Figure C.3
Galambos Copula 0.5-Correlation



The left panel displays the scatterplot of a Galambos copula with a correlation coefficient equal to 0.5. The right-hand panel shows the corresponding χ-plot.

Figure C.4
Galambos Copula 0.9-Correlation



The left panel displays the scatterplot of a Galambos copula with a correlation coefficient equal to 0.9. The right-hand panel shows the corresponding χ-plot.

Appendix D. Bivariate Gumbel Copulas

Let $\tilde{u} = -\log u$ and $\tilde{v} = -\log v$. For $1 \leq \delta < \infty$ the Gumbel copula distribution is

$$C(u, v; \delta) = \exp\left\{-(\tilde{u}^{\delta} + \tilde{v}^{\delta})^{1/\delta}\right\} \tag{40}$$

and its density function is

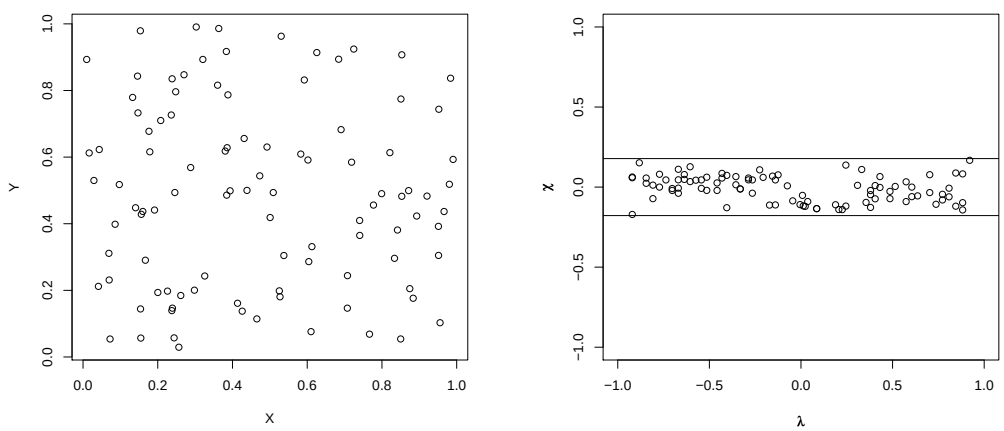
$$c(u, v; \delta) = C(u, v; \delta)(uv)^{-1} \frac{(\tilde{u}\tilde{v})^{\delta-1}}{(\tilde{u}^{\delta} + \tilde{v}^{\delta})^{2-1/\delta}} \left[(\tilde{u}^{\delta} + \tilde{v}^{\delta})^{1/\delta} + \delta - 1\right]. \tag{41}$$

The h function is

$$h(u, v; \delta) = v^{-1} \exp\left(-(\tilde{v}^{\delta} + \tilde{u}^{\delta})^{1/\delta}\right) \cdot \left(1 + \left(\frac{\tilde{u}}{\tilde{v}}\right)^{\delta}\right)^{-1+1/\delta} \tag{42}$$

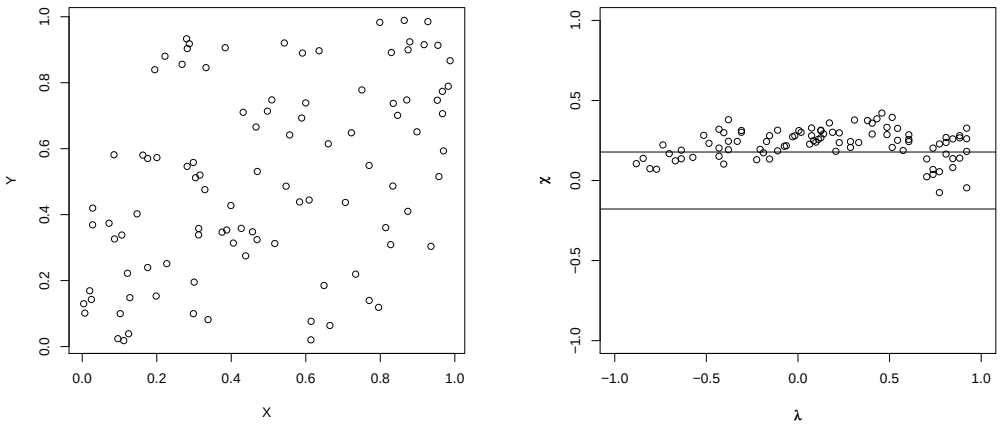
and the function $h^{-1}(u, v; \delta)$ cannot be written in closed form; therefore, we need to use a numerical routine to invert it. The derivation of these formulas is in [1].

Figure D.1
Gumbel Copula 0-Correlation



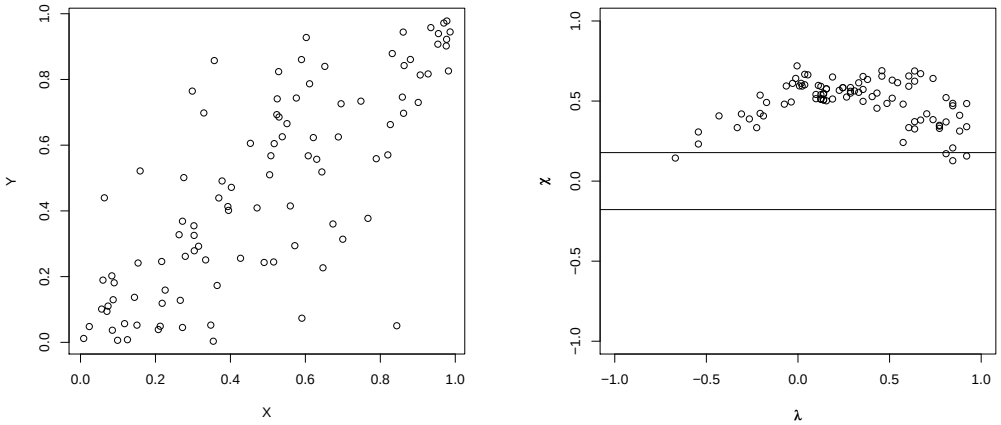
The left panel displays the scatterplot of a Gumbel copula with a correlation coefficient of zero. The right-hand panel shows the corresponding χ-plot. Notice that most of the points are within the 95 percent control bands.

Figure D.2
Gumbel Copula 0.2-Correlation



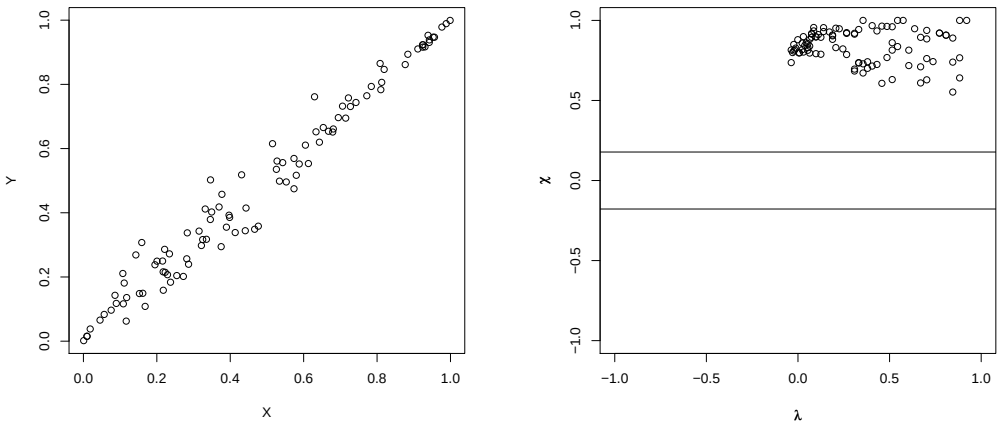
The left panel displays the scatterplot of a Gumbel copula with a correlation coefficient equal to 0.2. The right-hand panel shows the corresponding χ-plot.

Figure D.3
Gumbel Copula 0.5-Correlation



The left panel displays the scatterplot of a Gumbel copula with a correlation coefficient equal to 0.5. The right-hand panel shows the corresponding χ-plot.

Figure D.4
Gumbel Copula 0.9-Correlation



The left panel displays the scatterplot of a Gumbel copula with a correlation coefficient equal to 0.9. The right-hand panel shows the corresponding χ-plot.

Appendix E. Bivariate Normal Copulas

The *bivariate normal* copula is given by

$$C(u, v; \delta) = \Phi_\delta(\Phi^{-1}(u), \Phi^{-1}(v)), \tag{43}$$

where Φ is the standard normal distribution $N(0, 1)$ with mean zero and unit variance, Φ^{-1} is its inverse, and Φ_δ is the bivariate standard normal distribution with correlation δ.

The density is given by

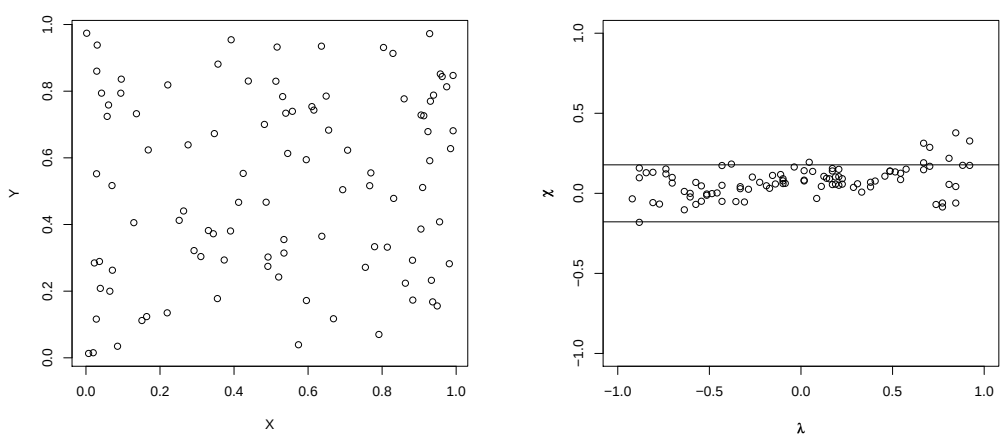
$$c(u, v; \delta) = \frac{1}{\sqrt{1-\delta^2}} \exp\left(-\frac{\delta^2(u^2+v^2) - 2\delta uv}{2(1-\delta^2)}\right), \tag{44}$$

and the h and h^{-1} functions are

$$h(u, v; \delta) = \Phi\left(\frac{\Phi^{-1}(u) - \delta\Phi^{-1}(v)}{\sqrt{1-\delta^2}}\right)$$
$$h^{-1}(u, v; \delta) = \Phi\left(\Phi^{-1}(u)\sqrt{1-\delta^2} + \delta\Phi^{-1}(v)\right).$$

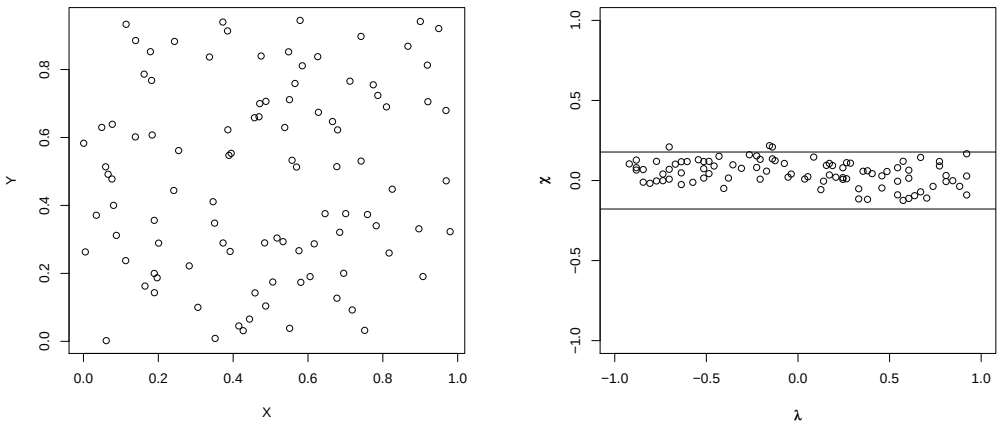
The derivation of these formulas is in [1].

Figure E.1
Normal Copula 0-Correlation



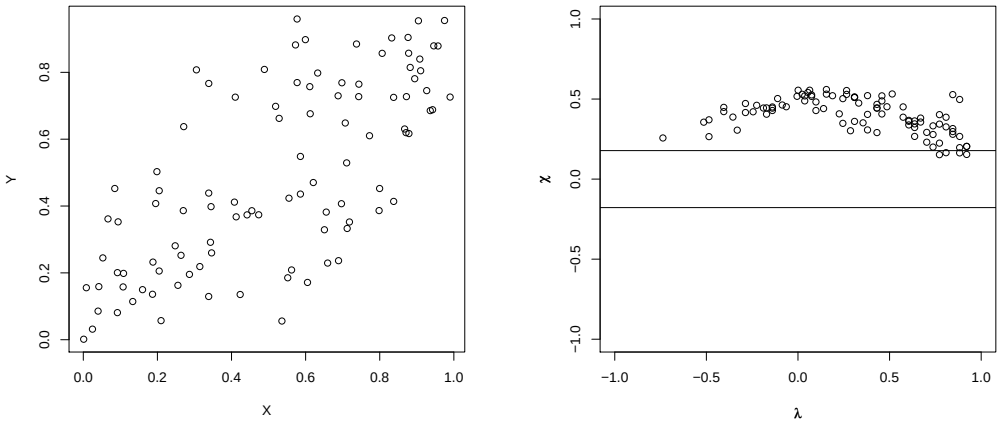
The left panel displays the scatterplot of a normal copula with a correlation coefficient of zero. The right-hand panel shows the corresponding χ-plot. Notice that most of the points are within the 95 percent control bands.

Figure E.2
Normal Copula 0.2-Correlation



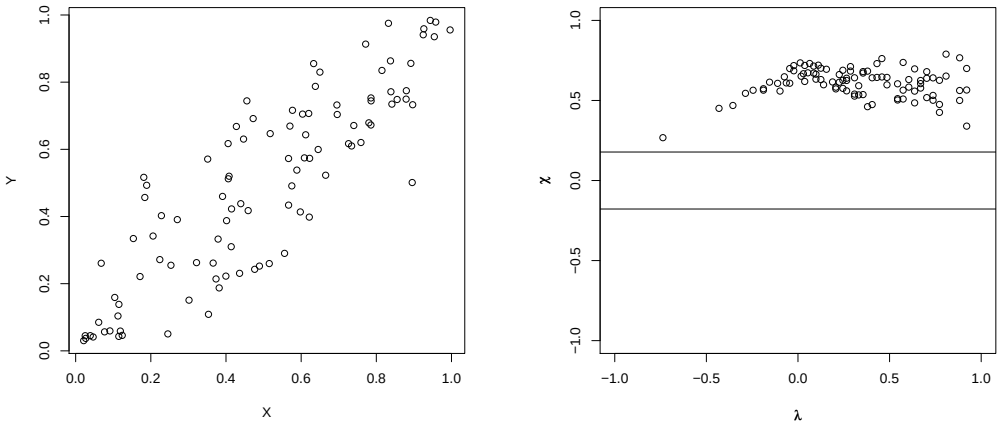
The left panel displays the scatterplot of a normal copula with a correlation coefficient equal to 0.2. The right-hand panel shows the corresponding χ-plot.

Figure E.3
Normal Copula 0.5-Correlation



The left panel displays the scatterplot of a normal copula with a correlation coefficient equal to 0.5. The right-hand panel shows the corresponding χ-plot.

Figure E.4
Normal Copula 0.9-Correlation



The left panel displays the scatterplot of a normal copula with a correlation coefficient equal to 0.9. The right-hand panel shows the corresponding χ-plot.

Appendix F. Canonical and D-Vine Algorithms

Algorithm 1 Simulation for a Canonical Vine.

This algorithm will generate a sample $x_1, x_2, \ldots, x_n$ from a canonical vine. The parameter $\Theta_{j,i}$ represents the necessary parameters for the copula $c_{j,j+1|1,\ldots,j-1}$ used in the construction.

```
For i = 1, 2, ..., n, let w_i be independent uniform random numbers on [0, 1].
x_1 ← v_{1,1} ← w_1
for i ← 2, 3, ..., n
  v_{i,1} ← w_i
  for k ← i − 1, i − 2, ..., 1
    v_{i,1} ← h^{−1}(v_{i,1}, v_{k,k}; Θ_{k,i−k})
  end for
  x_i ← v_{i,1}
  if i = n then
    Stop
  end if
  for j ← 1, 2, ..., i − 1
    v_{i,j+1} ← h(v_{i,j}, v_{j,j}; Θ_{j,i−j})
  end for
end for
```

Algorithm 2 Simulation for a D-Vine.

This algorithm will generate a sample $x_1, x_2, \ldots, x_n$ from a D-vine. The parameter $\Theta_{i,j}$ represents the necessary parameters for the i, j-th copula in the construction.

```
For i = 1, 2, ..., n, let w_i be independent uniform random numbers on [0, 1].
x_1 ← v_{1,1} ← w_1
x_2 ← v_{2,1} ← h^{-1}(w_2, v_{1,1}, Θ_{1,1})
v_{2,2} ← h(v_{1,1}, v_{2,1}, Θ_{1,1})
for i ← 3, 4, ..., n
  v_{i,1} ← w_i
    for k ← i − 1, i − 2, ..., 2
      v_{i,1} ← h^{-1}(v_{i,1}, v_{i−1,2k−2}, Θ_{k,i−k})
    end for
    v_{i,1} ← h^{-1}(v_{i,1}, v_{i−1,1}, Θ_{1,i−1})
    x_i ← v_{i,1}
    if i = n then
      Stop
    end if
    v_{i,2} ← h(v_{i−1,1}, v_{i,1}, Θ_{1,i−1})
    v_{i,3} ← h(v_{i,1}, v_{i−1,1}, Θ_{1,i−1})
    if i > 3 then
      for j ← 2, 3, ..., i − 2
        v_{i,2j} ← h(v_{i−1,2j−2}, v_{i,2j−1}, Θ_{j,i−j})
        v_{i,2j+1} ← h(v_{i,2j−1}, v_{i−1,2j−2}, Θ_{j,i−j})
      end for
    end if
  v_{i,2i−2} ← h(v_{i−1,2i−4}, v_{i,2i−3}, Θ_{i−1,1})
end for
```

Algorithm 3 Log-likelihood for Canonical Vine.

Evaluation of the log-likelihood function for a canonical vine.

```
log-likelihood ← 0
for i ← 1, 2, ..., n
  v_{0,i} ← x_i
end for
for j ← 1, 2, ..., n − 1
  for i ← 1, 2, ..., n − j
    log-likelihood ← log-likelihood + ℓ(v_{j−1,1}, v_{j−1,i+1}; Θ_{j,i})
  end for
  if j = n − 1 then
    Stop
  end if
  for i ← 1, 2, ..., n − j
    v_{j,i} = h(v_{j−1,i+1}, v_{j−1,1}; Θ_{j,i})
  end for
end for
```

Algorithm 4 Log-likelihood for D-Vine.

Evaluation of the log-likelihood function for a D-vine decomposition.

```
log-likelihood ← 0
for i ← 1, 2, ..., n
    v_{0,i} ← x_i
end for
for i ← 1, 2, ..., n − 1
    log-likelihood ← log-likelihood + ℓ(v_{0,i}, v_{0,i+1}; Θ_{1,i})
end for
v_{1,1} ← h(v_{0,1}, v_{0,2}; Θ_{1,1})
for k ←1, 2, ..., n-3
    v_{1,2k} ← h(v_{0,k+2}, v_{0,k+1}; Θ_{1,k+1})
    v_{1,2k+1} ← h(v_{0,k+1}, v_{0,k+2}; Θ_{1,k+1})
end for
v_{1,2n−4} ← h(v_{0,n}, v_{0,n−1}; Θ_{1,n−1})
for j ← 2, 3, ..., n − 1
    for i ← 1, 2, ..., n − j
        log-likelihood ← log-likelihood + ℓ(v_{j−1,2i−1}, v_{j−1,2i}; Θ_{j,i})
    end for
    if j = n − 1 then
        Stop
    end if
    v_{j,1} ← h(v_{j−1,1}, v_{j−1,2}; Θ_{j,1})
    if n > 4 then
        for i ← 1, 2, ..., n − j − 2
            v_{j,2i} ← h(v_{j−1,2i}, v_{j−1,2i+1}; Θ_{j,i+1})
            v_{j,2i+1} ← h(v_{j−1,2i+1}, v_{j−1,2i+2}; Θ_{j,i+1})
        end for
    end if
    v_{j,2n−2j−2} ← h(v_{j−1,2n−2j}, v_{j−1,2n−2j−1}; Θ_{j,n−j})
end for
```