8

Copulas

8.1 Introduction

Copulas are a popular method for modeling multivariate distributions. A copula models the dependence—and only the dependence—between the variates in a multivariate distribution and can be combined with any set of univariate distributions for the marginal distributions. Consequently, the use of copulas allows us to take advantage of the wide variety of univariate models that are available.

A *copula* is a multivariate CDF whose univariate marginal distributions are all Uniform(0,1). Suppose that $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ has a multivariate CDF F_Y with continuous marginal univariate CDFs $F_{Y_1}, \ldots, F_{Y_d}$. Then, by equation (A.9) in Section A.9.2, each of $F_{Y_1}(Y_1), \ldots, F_{Y_d}(Y_d)$ is Uniform(0,1) distributed. Therefore, the CDF of $\{F_{Y_1}(Y_1), \ldots, F_{Y_d}(Y_d)\}$ is a copula. This CDF is called the copula of $\boldsymbol{Y}$ and denoted by C_Y. C_Y contains all information about dependencies among the components of $\boldsymbol{Y}$ but has no information about the marginal CDFs of $\boldsymbol{Y}$.

It is easy to find a formula for C_Y. To avoid technical issues, in this section we will assume that all random variables have continuous, strictly increasing CDFs. More precisely, the CDFs are assumed to be increasing on their support. For example, the exponential CDF

$$F(y) = \begin{cases} 1 - e^{-y}, & y \geq 0, \\ 0, & y < 0, \end{cases}$$

has support $[0, \infty)$ and is strictly increasing on that set. The assumption that the CDF is continuous and strictly increasing is avoided in more mathematically advanced texts; see Section 8.8.

Since C_Y is the CDF of $\{F_{Y_1}(Y_1), \ldots, F_{Y_d}(Y_d)\}$, by the definition of a CDF we have

$$C_Y(u_1, \ldots, u_d) = P\left\{F_{Y_1}(Y_1) \leq u_1, \ldots, F_{Y_d}(Y_d) \leq u_d\right\}$$

D. Ruppert, *Statistics and Data Analysis for Financial Engineering*, Springer Texts in Statistics, DOI 10.1007/978-1-4419-7787-8_8,

$$
\begin{aligned}
&= P\left\{Y_1 \le F_{Y_1}^{-1}(u_1), \ldots, Y_d \le F_{Y_d}^{-1}(u_d)\right\} \\
&= F_Y\left\{F_{Y_1}^{-1}(u_1), \ldots, F_{Y_d}^{-1}(u_d)\right\}. \qquad (8.1)
\end{aligned}
$$

Next, letting $u_j = F_{Y_j}(y_j)$, $j = 1, \ldots, d$, in (8.1) we see that

$$
F_Y(y_1, \ldots, y_d) = C_Y\left\{F_{Y_1}(y_1), \ldots, F_{Y_d}(y_d)\right\}. \qquad (8.2)
$$

Equation (8.2) is part of a famous theorem due to Sklar which states that the F_Y can be decomposed into the copula C_Y, which contains all information about the dependencies among $(Y_1, \ldots, Y_d)$, and the univariate marginal CDFs $F_{Y_1}, \ldots, F_{Y_d}$, which contain all information about the univariate marginal distributions.

Let

$$
c_Y(u_1, \ldots, u_d) = \frac{\partial^d}{\partial u_1 \cdots \partial u_d} C_Y(u_1, \ldots, u_d) \qquad (8.3)
$$

be the density of C_Y. By differentiating (8.2), we find that the density of $\boldsymbol{Y}$ is equal to

$$
f_Y(y_1, \ldots, y_d) = c_Y\{F_{Y_1}(y_1), \ldots, F_{Y_d}(y_d)\} f_{Y_1}(y_1) \cdots f_{Y_d}(y_d). \qquad (8.4)
$$

One important property of copulas is that they are invariant to strictly increasing transformations of the variables. More precisely, suppose that g_j is strictly increasing and $X_j = g_j(Y_j)$ for $j = 1, \ldots, d$. Then $\boldsymbol{X} = (X_1, \ldots, X_d)$ and $\boldsymbol{Y}$ have the same copulas. To see this, first note that the CDF of $\boldsymbol{X}$ is

$$
\begin{aligned}
F_X(x_1, \ldots, x_d) &= P\left\{g_1(Y_1) \le x_1, \ldots, g_d(Y_d) \le x_d\right\} \\
&= P\left\{Y_1 \le g_1^{-1}(x_1), \ldots, Y_d \le g_d^{-1}(x_d)\right\} \\
&= F_Y\left\{g_1^{-1}(x_1), \ldots, g_d^{-1}(x_d)\right\} \qquad (8.5)
\end{aligned}
$$

and therefore the CDF of X_j is

$$
F_{X_j}(x_j) = F_{Y_j}\left\{g_j^{-1}(x_j)\right\}.
$$

Consequently,

$$
F_{X_j}^{-1}(u) = g_j\left\{F_{Y_j}^{-1}(u)\right\} \qquad (8.6)
$$

and by (8.1) applied to $\boldsymbol{X}$, (8.5), (8.6), and then (8.1) applied to $\boldsymbol{Y}$, the copula of $\boldsymbol{X}$ is

$$
\begin{aligned}
C_X(u_1, \ldots, u_d) &= F_X\left\{F_{X_1}^{-1}(u_1), \ldots, F_{X_d}^{-1}(u_d)\right\} \\
&= F_Y\left[g_1^{-1}\left\{F_{X_1}^{-1}(u_1)\right\}, \ldots, g_d^{-1}\left\{F_{X_d}^{-1}(u_d)\right\}\right] \\
&= F_Y\left\{F_{Y_1}^{-1}(u_1), \ldots, F_{Y_d}^{-1}(u_d)\right\} \\
&= C_Y(u_1, \ldots, u_d).
\end{aligned}
$$

To use copulas to model multivariate dependencies, we need parametric families of copulas. We turn to that topic next.

8.2 Special Copulas

There are three copulas of special interest because they represent independence and the two extremes of dependence.

The d-dimensional *independence copula* is the copula of d independent uniform(0,1) random variables. It equals

$$C^{\text{ind}}(u_1, \ldots, u_d) = u_1 \cdots u_d, \tag{8.7}$$

and has a density that is uniform on $[0,1]^d$, that is, its density is $c^{\text{ind}}(u_1, \ldots, u_d) = 1$ on $[0,1]^d$.

The d-dimensional *co-monotonicity copula* C^{M} has perfect positive dependence. Let U be Uniform(0,1). Then, the co-monotonicity copula is the CDF of $\boldsymbol{U} = (U, \ldots, U)$; that is, $\boldsymbol{U}$ contains d copies of U so that all of the components of $\boldsymbol{U}$ are equal. Thus,

$$\begin{aligned} C^{\text{M}}(u_1, \ldots, u_d) &= P(U \leq u_1, \ldots, U \leq u_d) = P\{Y \leq \min(u_1, \ldots, u_d)\} \\ &= \min(u_1, \ldots, u_d). \end{aligned}$$

The two-dimensional *counter-monotonicity copula* C^{CM} copula is the CDF of $(U, 1-U)$, which has perfect negative dependence. Therefore,

$$\begin{aligned} C^{\text{CM}}(u_1, u_2) &= P(U \leq u_1 \ \&\ 1 - U \leq u_2) \\ &= P(1 - u_2 \leq U \leq u_1) = \max(u_1 + u_2 - 1, 0). \end{aligned} \tag{8.8}$$

It is easy to derive the last equality in (8.8). If $1 - u_2 > u_1$, then the event $\{1 - u_2 \leq U \leq u_1\}$ is impossible so the probability is 0. Otherwise, the probability is the length of the interval $(1 - u_2, u_1)$, which is $u_1 + u_2 - 1$. It is not possible to have a counter-monotonicity copula with $d > 2$. If, for example, U_1 is counter-monotonic to U_2 and U_2 is counter-monotonic to U_3, then U_1 and U_3 will be co-monotonic, not counter-monotonic.

8.3 Gaussian and t-Copulas

Multivariate normal and t-distributions offer a convenient way to generate families of copulas. Let $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ have a multivariate normal distribution. Since $C_{\boldsymbol{Y}}$ depends only on the dependencies within $\boldsymbol{Y}$, not the univariate marginal distributions, $C_{\boldsymbol{Y}}$ depends only on the correlation matrix of $\boldsymbol{Y}$, which will be denoted by $\boldsymbol{\Omega}$. Therefore, there is a one-to-one correspondence between correlation matrices and Gaussian copulas. The Gaussian copula with correlation matrix $\boldsymbol{\Omega}$ will be denoted $C^{\text{Gauss}}(\,\cdot\,|\boldsymbol{\Omega})$.

If a random vector $\boldsymbol{Y}$ has a Gaussian copula, then $\boldsymbol{Y}$ is said to have a *meta-Gaussian distribution*. This does not, of course, mean that $\boldsymbol{Y}$ has a multivariate Gaussian distribution, since the univariate marginal distributions of $\boldsymbol{Y}$ could be any distributions at all. A d-dimensional Gaussian copula whose correlation matrix is the identity matrix, so that all correlations are zero, is the d-dimensional independence copula. A Gaussian copula will converge to the co-monotonicity copula if all correlations in $\boldsymbol{\Omega}$ converge to 1. In the bivariate case, as the correlation converges to -1, the copula converges to the counter-monotonicity copula.

Similarly, let $C^{\mathrm{t}}(\cdot\,|\nu, \boldsymbol{\Omega})$ be the copula of a multivariate t-distribution with correlation matrix $\boldsymbol{\Omega}$ and degrees of freedom ν.[1] The shape parameter ν affects both the univariate marginal distributions and the copula, so ν is a parameter of the copula. We will see in Section 8.6 that ν determines the amount of tail dependence in a t-copula. A distribution with a t-copula is called a *t-meta distribution*.

8.4 Archimedean Copulas

An *Archimedean copula* with a strict generator has the form

$$C(u_1, \ldots, u_d) = \phi^{-1}\{\phi(u_1) + \cdots + \phi(u_d)\}, \tag{8.9}$$

where the function ϕ is the generator of the copula and satisfies

1. ϕ is a continuous, strictly decreasing, and convex function mapping $[0, 1]$ onto $[0, \infty]$,
2. $\phi(0) = \infty$, and
3. $\phi(1) = 0$.

Figure 8.1 is a plot of a generator and illustrates these properties. It is possible to relax assumption 2, but then the generator is not called strict and construction of the copula is more complex. There are many families of Archimedean copulas, but we will only look at three, the Clayton, Frank, and Gumbel copulas.

Notice that in (8.9), the value of $C(u_1, \ldots, u_d)$ is unchanged if we permute $u_1, \ldots, u_d$. A distribution with this property is called *exchangeable*. One consequence of exchangeability is that both Kendall's and Spearman's rank correlation introduced later in Section 8.5 are the same for all pairs of variables. Archimedean copulas are most useful in the bivariate case or in applications where we expect all pairs to have similar dependencies.

8.4.1 Frank Copula

The Frank copula has generator

$$\phi^{\mathrm{Fr}}(u) = -\log\left\{\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right\}, \quad -\infty < \theta < \infty.$$

[1] There is a minor technical issue here if $\nu \leq 2$. In this case, the t-distribution does not have covariance and correlation matrices. However, it still has a scale matrix and we will assume that the scale matrix is equal to some correlation matrix $\boldsymbol{\Omega}$.

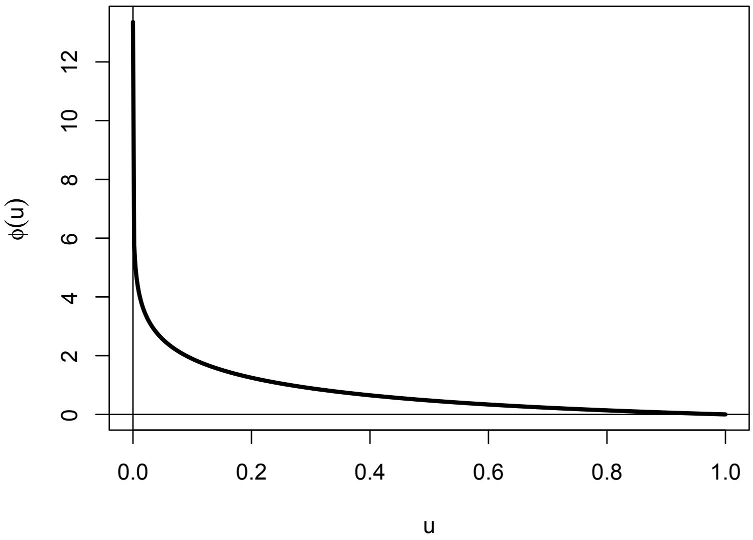

Fig. 8.1. *Generator of the Frank copula with* $\theta = 1$.

The inverse generator is

$$(\phi^{\mathrm{Fr}})^{-1}(y) = -\frac{\log\left[e^{-y}\{e^{-\theta} - 1\} + 1\right]}{\theta}.$$

Therefore, by (8.9), the bivariate Frank copula is

$$C^{\mathrm{Fr}}(u_1, u_2) = -\frac{1}{\theta}\log\left\{1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right\}. \qquad (8.10)$$

The case $\theta = 0$ requires some care, since plugging this value into (8.10) gives $0/0$. Instead, one must evaluate the limit of (8.10) as $\theta \to 0$. Using the approximations $e^x - 1 \approx x$ and $\log(1 + x) \approx x$ as $x \to 0$, one can show that as $\theta \to 0$, $C^{\mathrm{Fr}}(u_1, u_2) \to u_1 u_2$, the bivariate independence copula. Therefore, for $\theta = 0$ we define the Frank copula to be the independence copula.

It is interesting to study the limits of $C^{\mathrm{Fr}}(u_1, u_2)$ as $\theta \to \pm\infty$. As $\theta \to -\infty$, the bivariate Frank copula converges to the counter-monotonicity copula. To see this, first note that as $\theta \to -\infty$,

$$C^{\mathrm{Fr}}(u_1, u_2) \sim -\frac{1}{\theta}\log\left\{1 + e^{-\theta(u_1+u_2-1)}\right\}. \qquad (8.11)$$

If $u_1 + u_2 - 1 > 0$, then as $\theta \to -\infty$, the exponent $-\theta(u_1 + u_2 - 1)$ in (8.11) converges to ∞ and

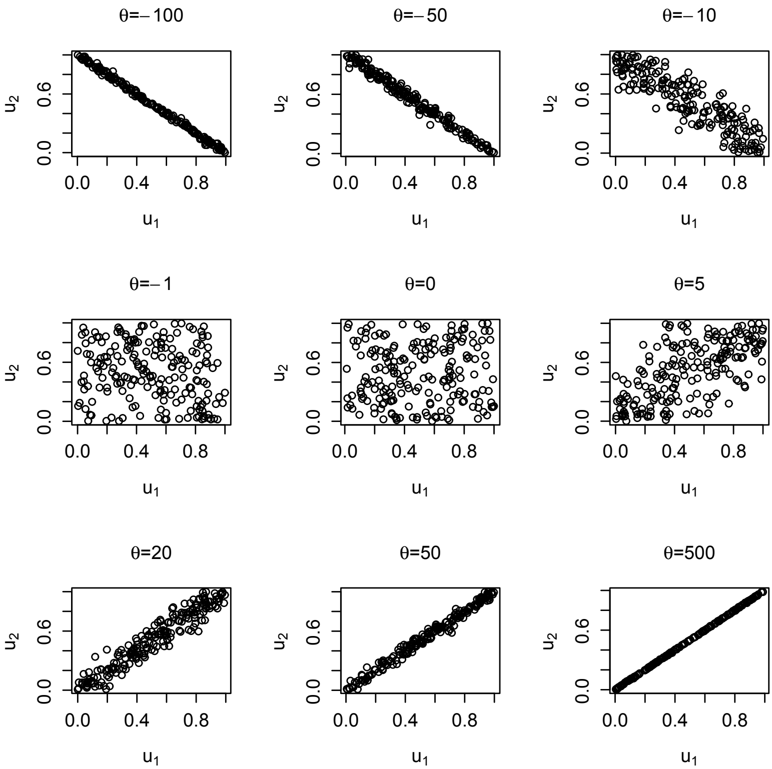

Fig. 8.2. *Random samples from Frank copulas.*

$$\log\left\{1+e^{-\theta(u_1+u_2-1)}\right\} \sim -\theta(u_1+u_2-1)$$

so that $C^{\mathrm{Fr}}(u_1,u_2) \to u_1+u_2-1$. If $u_1+u_2-1<0$, then $-\theta(u_1+u_2-1) \to -\infty$ and $C^{\mathrm{Fr}}(u_1,u_2) \to 0$. Putting these results together, we see that $C^{\mathrm{Fr}}(u_1,u_2)$ converges to $\max(0, u_1+u_2-1)$, the counter-monotonicity copula, as $\theta \to -\infty$.

As $\theta \to \infty$, $C^{\mathrm{Fr}}(u_1,u_2) \to \min(u_1,u_2)$, the co-monotonicity copula. Verification of this is left as an exercise for the reader.

Figure 8.2 contains scatterplots of bivariate samples from nine Frank copulas, all with a sample size of 200 and with values of θ that give dependencies ranging from strongly negative to strongly positive. The convergence to the counter-monotonicity (co-monotonicity) copula as $\theta \to -\infty$ $(+\infty)$ can be seen in the scatterplots.

8.4.2 Clayton Copula

The *Clayton copula*, with generator $(t^{-\theta}-1)/\theta$, $\theta > 0$, is

$$C^{\text{Cl}}(u_1, \ldots, u_d) = (u_1^{-\theta} + \cdots + u_d^{-\theta} - d + 1)^{-1/\theta}.$$

We define the Clayton copula for $\theta = 0$ as the limit

$$\lim_{\theta \downarrow 0} C^{\text{Cl}}(u_1, \ldots, u_d) = u_1 \cdots u_d$$

which is the independence copula. There is another way to derive this result. As $\theta \downarrow 0$, l'Hôpital's rule shows that the generator $(t^{-\theta} - 1)/\theta$ converges to $\phi(t) = -\log(t)$ with inverse $\phi^{-1}(t) = \exp(-t)$. Therefore,

$$\begin{aligned} C^{\text{Cl}}(u_1, \ldots, u_d) &= \phi^{-1}\{\phi(u_1) + \cdots + \phi(u_d)\} \\ &= \exp\{-(-\log u_1 - \cdots - \log u_d)\} = u_1 \cdots u_d. \end{aligned}$$

It is possible to extend the range of θ to include $-1 \leq \theta < 0$, but then the generator $(t^{-\theta} - 1)/\theta$ is finite at $t = 0$ in violation of assumption 2. of strict generators. Thus, the generator is not strict if $\theta < 0$. As a result, it is necessary to define $C^{\text{Cl}}(u_1, \ldots, u_d)$ to equal 0 for small values of u_i. To appreciate this, consider the bivariate case. If $-1 \leq \theta < 0$, then $u_1^{-\theta} + u_2^{-\theta} - 1 < 0$ occurs when u_1 and u_2 are both small. In these cases, $C^{\text{Cl}}(u_1, u_2)$ is set equal to 0. Therefore, there is no probability in the region $u_1^{-\theta} + u_2^{-\theta} - 1 < 0$. In the limit, as $\theta \to -1$, there is no probability in the region $u_1 + u_2 < 1$.

As $\theta \to -1$, the bivariate Clayton copula converges to the counter-monotonicity copula, and as $\theta \to \infty$, the Clayton copula converges to the co-monotonicity copula.

Figure 8.3 contains scatterplots of bivariate samples from Clayton copulas, all with a sample size of 200 and with values of θ that give dependencies ranging from counter-monotonicity to co-monotonicity. Comparing Figures 8.2 and 8.3, we see that the Frank and Clayton copulas are rather different when the amount of dependence is somewhere between these two extremes. In particular, the Clayton copula's exclusion of the region $u_1^{-\theta} + u_2^{-\theta} - 1 < 0$ when $\theta < 0$ is evident, especially in the example with $\theta = -0.7$. In contrast, the Frank copula has positive probability on the entire unit square. The Frank copula is symmetric about the diagonal from $(0, 1)$ to $(1, 0)$, but the Clayton copula does not have this symmetry.

8.4.3 Gumbel Copula

The Gumbel copula has generator $\{-\log(t)\}^\theta$, $\theta \geq 1$, and consequently is equal to

$$C^{\text{Gu}}(u_1, \ldots, u_d) = \exp\left[-\left\{(\log u_1)^\theta + \cdots + (\log u_d)^\theta\right\}^{1/\theta}\right].$$

The Gumbel copula is the independence copula when $\theta = 1$ and converges to the co-monotonicity copula as $\theta \to \infty$, but the Gumbel copula cannot have negative dependence.

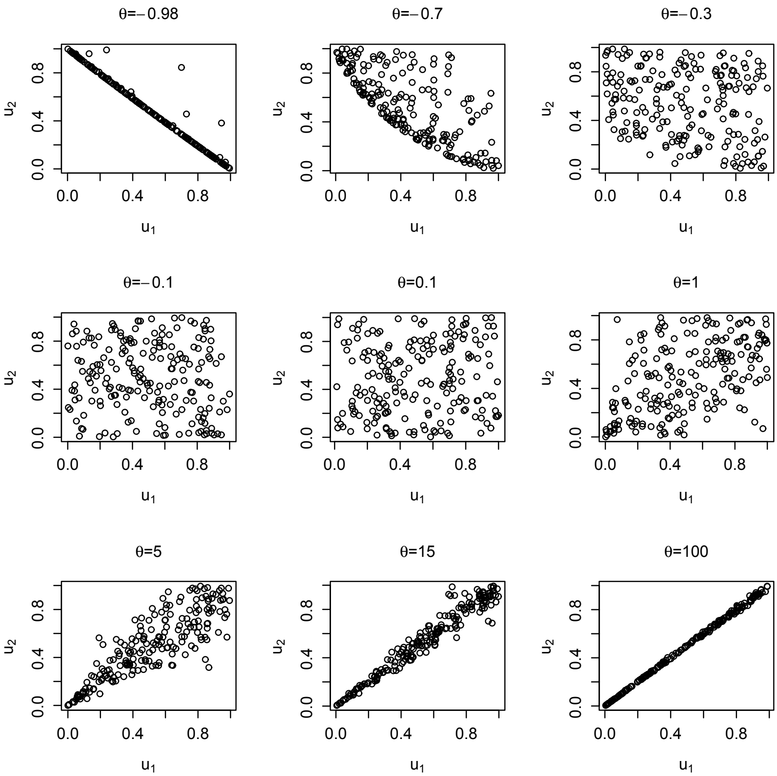

Fig. 8.3. *Random samples of size 200 from Clayton copulas.*

Figure 8.4 contains scatterplots of bivariate samples from Gumbel copulas, with a sample size of 200 and with values of θ that give dependencies ranging from near independence to strong positive dependence.

In applications, it is useful that the different copula families have different properties, since this increases the likelihood of finding a copula that fits the data.

8.5 Rank Correlation

The Pearson correlation coefficient defined by (4.3) is not convenient for fitting copulas to data, since it depends on the univariate marginal distributions as well as the copula. Rank correlation coefficients remedy this problem, since they depend only on the copula.

For each variable, the ranks of that variable are determined by ordering the observations from smallest to largest and giving the smallest rank 1, the next-smallest rank 2, and so forth. In other words, if $Y_1, \ldots, Y_n$ is a sample,

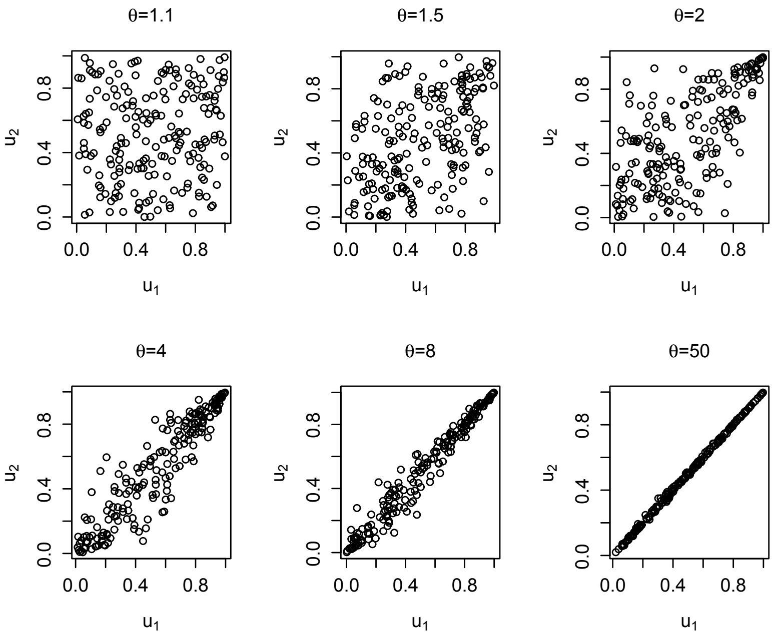

Fig. 8.4. *Random samples from Gumbel copulas.*

then the *rank* of Y_i in the sample is equal to 1 if Y_i is the smallest observation, is 2 if Y_2 is the second smallest, and so forth. More mathematically, the rank of Y_i can be defined also by the formula

$$\text{rank}(Y_i) = \sum_{j=1}^{n} I(Y_j \leq Y_i), \tag{8.12}$$

which counts the number of observations (including Y_i itself) that are less than or equal to Y_i. A *rank statistic* is a statistic that depends on the data only through the ranks. A key property of ranks is that they are unchanged by strictly monotonic transformations. In particular, the ranks are unchanged by transforming each variable by its CDF, so the distribution of any rank statistic depends only on the copula of the data, not on the univariate marginals.

We will be concerned with rank statistics that measure statistical association between pairs of variables. These statistics are called *rank correlations.* There are two rank correlation coefficients in widespread usage, Kendall's tau and Spearman's rho.

8.5.1 Kendall's Tau

Let (Y_1, Y_2) be a bivariate random vector and let (Y_1^*, Y_2^*) be an independent copy of (Y_1, Y_2). Then (Y_1, Y_2) and (Y_1^*, Y_2^*) are called a *concordant pair* if the ranking of Y_1 relative to Y_1^* is the same as the ranking of Y_2 relative to Y_2^*, that is, either $Y_1 > Y_1^*$ and $Y_2 > Y_2^*$ or $Y_1 < Y_1^*$ and $Y_2 < Y_2^*$. In either case, $(Y_1 - Y_1^*)(Y_2 - Y_2^*) > 0$. Similarly, (Y_1, Y_2) and (Y_1^*, Y_2^*) are called a *discordant pair* if $(Y_1 - Y_1^*)(Y_2 - Y_2^*) < 0$. *Kendall's tau* is the probability of a concordant pair minus the probability of a discordant pair. Therefore, Kendall's tau for (Y_1, Y_2) is

$$\begin{aligned} \rho_\tau(Y_1, Y_2) &= P\{(Y_1 - Y_1^*)(Y_2 - Y_2^*) > 0\} - P\{(Y_1 - Y_1^*)(Y_2 - Y_2^*) < 0\} \\ &= E\left[\text{sign}\{(Y_1 - Y_1^*)(Y_2 - Y_2^*)\}\right], \qquad (8.13) \end{aligned}$$

where the *sign function* is

$$\text{sign}(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0, \\ 0, & x = 0. \end{cases}$$

It is easy to check that if g and h are increasing functions, then

$$\rho_\tau\{g(Y_1), h(Y_2)\} = \rho_\tau(Y_1, Y_2). \qquad (8.14)$$

Stated differently, Kendall's tau is invariant to monotonically increasing transformations. If g and h are the marginal CDFs of Y_1 and Y_2, then the left-hand side of (8.14) is the value of Kendall's tau for the copula of (Y_1, Y_2). This shows that Kendall's tau depends only on the copula of a bivariate random vector. For a random vector $\boldsymbol{Y}$, we define the *Kendall tau correlation matrix* to be the matrix whose (j, k) entry is Kendall's tau for the jth and kth components of $\boldsymbol{Y}$.

If we have a bivariate sample $\boldsymbol{Y}_i = (Y_{i,1}, Y_{i,2})$, $i = 1, \ldots, n$, then the sample Kendall's tau is

$$\widehat{\rho}_\tau(Y_1, Y_2) = \binom{n}{2}^{-1} \sum_{1 \le i < j \le n} \text{sign}\left\{(Y_{i,1} - Y_{j,1})(Y_{i,2} - Y_{j,2})\right\}. \qquad (8.15)$$

Note that $\binom{n}{2}$ is the number of summands in (8.15), so $\widehat{\rho}$ is $\text{sign}\{(Y_{i,1} - Y_{j,1})(Y_{i,2} - Y_{j,2})\}$ averaged across all distinct pairs and is a sample version of (8.13).

8.5.2 Spearman's Correlation Coefficient

For a sample, Spearman's correlation coefficient is simply the usual Pearson correlation calculated from the ranks of the data. For a distribution (that is, an infinite population rather than a finite sample), both variables are transformed by their CDFs and then the Pearson correlation is computed from the transformed variables. Transforming a random variable by its CDF is analogous to computing the ranks of a variable in a finite sample.

Stated differently, Spearman's correlation coefficient, also called *Spearman's rho*, for a bivariate random vector (Y_1, Y_2) will be denoted by $\rho_S(Y_1, Y_2)$ and is defined to be the Pearson correlation coefficient of $\{F_{Y_1}(Y_1), F_{Y_2}(Y_2)\}$:

$$\rho_S(Y_1, Y_2) = \mathrm{Corr}\{F_{Y_1}(Y_1), F_{Y_2}(Y_2)\}.$$

Since the distribution of $\{F_{Y_1}(Y_1), F_{Y_2}(Y_2)\}$ is the copula of (Y_1, Y_2), Spearman's rho, like Kendall's tau, depends only on the copula.

The sample version of Spearman's correlation coefficient can be computed from the ranks of the data and for a bivariate sample $\boldsymbol{Y}_i = (Y_{i,1}, Y_{i,2})$, $i = 1, \ldots, n$, is

$$\widehat{\rho}_S(Y_1, Y_2) = \frac{12}{n(n^2-1)} \sum_{i=1}^{n} \left\{\mathrm{rank}(Y_{i,1}) - \frac{n+1}{2}\right\} \left\{\mathrm{rank}(Y_{i,2}) - \frac{n+1}{2}\right\}. \tag{8.16}$$

The set of ranks for any variable is, of course, the integers 1 to n and $(n+1)/2$ is the mean of its ranks. It can be shown that $\widehat{\rho}_S(Y_1, Y_2)$ is the sample Pearson correlation between the ranks of $Y_{i,1}$ and the ranks of $Y_{i,2}$.[2]

If $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ is a random vector, then the *Spearman correlation matrix* of $\boldsymbol{Y}$ is the correlation matrix of $\{F_{Y_1}(Y_1), \ldots, F_{Y_d}(Y_d)\}$ and contains the Spearman correlation coefficients for all pairs of coordinates of $\boldsymbol{Y}$. The sample Spearman correlation matrix is defined analogously.

8.6 Tail Dependence

Tail dependence measures association between the extreme values of two random variables and depends only on their copula. We will start with lower tail dependence, which uses extremes in the lower tail. Suppose that $\boldsymbol{Y} = (Y_1, Y_2)$ is a bivariate random vector with copula $C_{\boldsymbol{Y}}$. Then the *coefficient of lower tail dependence* is denoted by λ_l and defined as

$$\lambda_l := \lim_{q \downarrow 0} P\left\{Y_2 \le F_{Y_2}^{-1}(q) \,|\, Y_1 \le F_{Y_1}^{-1}(q)\right\} \tag{8.17}$$

$$= \lim_{q \downarrow 0} \frac{P\left\{Y_2 \le F_{Y_2}^{-1}(q) \text{ and } Y_1 \le F_{Y_1}^{-1}(q)\right\}}{P\left\{Y_1 \le F_{Y_1}^{-1}(q)\right\}} \tag{8.18}$$

$$= \lim_{q \downarrow 0} \frac{P\left\{F_{Y_2}(Y_2) \le q \text{ and } F_{Y_1}(Y_1) \le q\right\}}{P\left\{F_{Y_1}(Y_1) \le q\right\}} \tag{8.19}$$

$$= \lim_{q \downarrow 0} \frac{C_{\boldsymbol{Y}}(q,q)}{q}. \tag{8.20}$$

---

[2] If there are ties, then ranks are averaged among tied observations. For example, if there are two observations tied for smallest, then they each get a rank of 1.5. When there are ties, then these results must be modified.

It is helpful to look at these equations individually. As elsewhere in this chapter, for simplicity we are assuming that F_{Y_1} and F_{Y_2} are strictly increasing on their supports and therefore have inverses.

First, (8.17) defines λ_l as the limit as $q \downarrow 0$ of the conditional probability that Y_2 is less than or equal to its qth quantile, given that Y_1 is less than or equal to its qth quantile. Since we are taking a limit as $q \downarrow 0$, we are looking at the extreme left tail. What happens if Y_1 and Y_2 are independent? Then $P(Y_2 \le y_2 \,|\, Y_1 \le y_1) = P(Y_2 \le y_2)$ for all y_1 and y_2. Therefore, the conditional probability in (8.17) equals the unconditional probability $P(Y_2 \le F_{Y_2}^{-1}(q))$ and this probability converges to 0 as $q \downarrow 0$. Therefore, $\lambda_l = 0$ implies that in the extreme left tail, Y_1 and Y_2 behave as if they were independent.

Equation (8.18) is just the definition of conditional probability. Equation (8.19) is simply (8.18) after applying the probability transformation to both variables.

The numerator in equation (8.20) is just the definition of a copula and the denominator is the result of $F_{Y_1}(Y_1)$ being Uniform(0,1) distributed; see (A.9).

Deriving formulas for λ_l for Gaussian and t-copulas is a topic best left for more advanced books. Here we give only the results; see Section 8.8 for further reading. For any Gaussian copula with $\rho \neq 1$, $\lambda_l = 0$, that is, Gaussian copulasdo not have tail dependence except in the extreme case of perfect positive correlation. For a t-copula with ν degrees of freedom and correlation ρ,

$$\lambda_l = 2F_{t,\nu+1}\left\{-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right\}, \tag{8.21}$$

where $F_{t,\nu+1}$ is the CDF of the t-distribution with $\nu + 1$ degrees of freedom.

Since $F_{t,\nu+1}(-\infty) = 0$, we see that $\lambda_l \to 0$ as $\nu \to \infty$, which makes sense since the t-copula converges to a Gaussian copula as $\nu \to \infty$. Also, $\lambda_l \to 0$ as $\rho \to -1$, which is also not too surprising, since $\rho = -1$ is perfect *negative* dependence and λ_l measures *positive* tail dependence.

The *coefficient of upper tail dependence*, λ_u, is

$$\lambda_u := \lim_{q \uparrow 1} P\left\{Y_2 \ge F_{Y_2}^{-1}(q) \,|\, Y_1 \ge F_{Y_1}^{-1}(q)\right\} \tag{8.22}$$

$$= 2 + \lim_{q \uparrow 1} \frac{1 - C_Y(q,q)}{1-q}. \tag{8.23}$$

We see that λ_u is defined analogously to λ_l; λ_u is the limit as $q \uparrow 1$ of the conditional probability that Y_2 is greater than or equal to its qth quantile, given that Y_1 is greater than or equal to its qth quantile. Deriving (8.23) is left as an exercise for the interested reader.

For Gaussian and t-copula, $\lambda_u = \lambda_l$, so that $\lambda_u = 0$ for any Gaussian copula and for a t-copula, λ_l is given by the right-hand side of (8.21). Coefficients of tail dependence for t-copulas are plotted in Figure 8.5. One can see $\lambda_l = \lambda_u$ depends strongly on both ρ and ν.

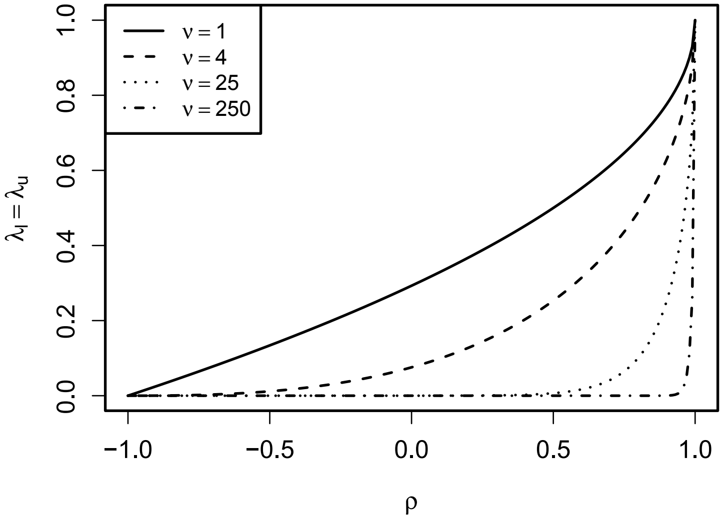

Fig. 8.5. *t-copulas coefficients of tail dependence as functions of* ρ *for* $\nu = 1$, 4, 25, *and* 250.

For the independence copula, λ_l and λ_u are both equal to 0, and for the co-monotonicity copula both are equal to 1.

Knowing whether or not there is tail dependence is important for risk management. If there are no tail dependencies among the returns on the assets in a portfolio, then there is little risk of clusters of very negative returns, and the risk of an extreme negative return on the portfolio is low. Conversely, if there are tail dependencies, then the likelihood of extreme negative returns occurring simultaneously on several assets in the portfolio can be high.

8.7 Calibrating Copulas

Assume that we have an i.i.d. sample $\boldsymbol{Y}_i = (Y_{i,1}, \ldots, Y_{i,d})$, $i = 1, \ldots, n$, and we wish to estimate the copula of $\boldsymbol{Y}_i$ and perhaps its marginal distributions as well.

An important task is choosing a copula model. The various copula models differ notably from each other. For example, some have tail dependence and others do not. The Gumbel copula allows only positive dependence or independence. The Clayton copula with negative dependence excludes the region where both u_1 and u_2 are small. As will be seen in this section, an appropriate copula model can be selected using graphical techniques as well as with AIC.

8.7.1 Maximum Likelihood

Suppose we have parametric models $F_{Y_1}(\cdot \mid \boldsymbol{\theta}_1), \ldots, F_{Y_d}(\cdot \mid \boldsymbol{\theta}_d)$ for the marginal CDFs as well as a parametric model $c_Y(\cdot|\boldsymbol{\theta}_C)$ for the copula density. By taking logs of (8.4), we find that the log-likelihood is

$$L(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d, \boldsymbol{\theta}_C) = \sum_{i=1}^{n} \Bigg(\log \left[c_Y \left\{ F_{Y_1}(Y_{i,1}|\boldsymbol{\theta}_1), \ldots, F_{Y_d}(Y_{i,d}|\boldsymbol{\theta}_d) \Big| \boldsymbol{\theta}_C \right\} \right] + \log \left\{ f_{Y_1}(Y_{i,1}|\boldsymbol{\theta}_1) \right\} + \cdots + \log \left\{ f_{Y_d}(Y_{i,d}|\boldsymbol{\theta}_d) \right\} \Bigg). \quad (8.24)$$

Maximum likelihood estimation finds the maximum of $L(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d, \boldsymbol{\theta}_C)$ over the entire set of parameters $(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d, \boldsymbol{\theta}_C)$.

There are two potential problems with maximum likelihood estimation. First, because of the large number of parameters, especially for large values of d, maximizing $L(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d, \boldsymbol{\theta}_C)$ can be a challenging numerical problem. This difficulty can be ameliorated by the use of starting values that are close to the MLEs. The pseudo-maximum likelihood estimates discussed in the next section are easier to compute than the MLE and can used either as an alternative to the MLE or as starting values for the MLE.

Second, maximum likelihood estimation requires parametric models for both the copula and the marginal distributions. If any of the marginal distributions are not well fit by a convenient parametric family, this may cause biases in the estimated parameters of both the marginal distributions and the copula. The semiparametric approach to pseudo-maximum likelihood estimation, where the marginal distributions are estimated nonparametrically, provides a remedy to this problem.

8.7.2 Pseudo-Maximum Likelihood

Pseudo-maximum likelihood estimation is a two-step process. In the first step, each of the d marginal distribution functions is estimated, one at a time. Let $\widehat{F}_{Y_j}$ be the estimate of the jth marginal CDF, $j = 1, \ldots, d$. In the second step,

$$\sum_{i=1}^{n} \log \left[c_Y \left\{ \widehat{F}_{Y_1}(Y_{i,1}), \ldots, \widehat{F}_{Y_d}(Y_{i,d}) \Big| \boldsymbol{\theta}_C \right\} \right] \quad (8.25)$$

is maximized over $\boldsymbol{\theta}_C$. Note that (8.25) is obtained from (8.24) by deleting terms that do not depend on $\boldsymbol{\theta}_C$ and replacing the marginal CDFs by estimates. By estimating parameters in the marginal distributions and in the copula separately, the pseudo-maximum likelihood approach avoids a high-dimensional optimization.

There are two approaches to step 1, parametric and nonparametric. In the parametric approach, parametric models $F_{Y_1}(\cdot \mid \boldsymbol{\theta}_1), \ldots, F_{Y_d}(\cdot \mid \boldsymbol{\theta}_d)$ for the marginal CDFs are assumed as in maximum likelihood estimation. The data $Y_{1,j}, \ldots, Y_{n,j}$ for the jth variate are used to estimate $\boldsymbol{\theta}_j$, usually by maximum likelihood as discussed in Chapter 5. Then, $\widehat{F}_{Y_j}(\cdot) = F_{Y_j}(\cdot|\widehat{\boldsymbol{\theta}}_j)$. In the nonparametric approach, $\widehat{F}_{Y_j}$ is estimated by the empirical CDF of $Y_{1,j}, \ldots, Y_{n,j}$, except that the divisor n in (4.1) is replaced by $n+1$ so that

$$\widehat{F}_{Y_j}(y) = \frac{\sum_{i=1}^{n} I\{Y_{i,j} \le y\}}{n+1}. \tag{8.26}$$

With this modified divisor, the maximum value of $\widehat{F}_{Y_j}(Y_{i,j})$ is $n/(n+1)$ rather than 1. Avoiding a value of 1 is essential when, as is often the case, $c_Y(u_1, \ldots, u_d|\boldsymbol{\theta}_C) = \infty$ if some of $u_1, \ldots, u_d$ are equal to 1.

When both steps are parametric, the estimation method is called *parametric pseudo-maximum likelihood*. The combination of a nonparametric step 1 and a parametric step 2 is called *semiparametric pseudo-maximum likelihood*.

In the second step of pseudo-maximum likelihood, the maximization can be difficult when $\boldsymbol{\theta}_C$ is high-dimensional. For example, if one uses a Gaussian or t-copula, then there are $d(d-1)/2$ correlation parameters. One way to solve this problem is to assume some structure to the correlation. An extreme case of this is the *equi-correlation model* where all nondiagonal elements of the correlation matrix have a common value, call it ρ. If one is reluctant to assume some type of structured correlation matrix, then it is essential to have good starting values for the correlation matrix when maximizing (8.25). For Gaussian and t-copulas, starting values can be obtained via rank correlations as discussed in the next section.

The values $\widehat{F}_{Y_j}(Y_{i,j})$, $i = 1, \ldots, n$ and $j = 1, \ldots, d$, will be called the *uniform-transformed variables*, since they should have approximately Uniform(0,1) distributions. The multivariate empirical CDF [see equation (A.38)] of the uniform-transformed variables is called the *empirical copula* and is a nonparametric estimate of the copula. The empirical copula is useful for checking the goodness of fits of parametric copula models; see Example 8.2.

8.7.3 Calibrating Meta-Gaussian and Meta-*t*-Distributions

Gaussian Copulas

Rank correlation can be useful for estimating the parameters of a copula. Suppose $\boldsymbol{Y}_i = (Y_{i,1}, \ldots, Y_{i,d})$, $i = 1, \ldots, n$, is an i.i.d. sample from a meta-Gaussian distribution. Then its copula is $C^{\text{Gauss}}(\,\cdot\,|\boldsymbol{\Omega})$ for some correlation matrix $\boldsymbol{\Omega}$. To estimate the distribution of $\boldsymbol{Y}$, we need to estimate the univariate marginal distributions and $\boldsymbol{\Omega}$. The marginal distribution can be estimated by the methods discussed in Chapter 5. Result (8.28) in the following theorem shows that $\boldsymbol{\Omega}$ can be estimated by the sample Spearman correlation matrix.

Theorem 8.1. *Let $\boldsymbol{Y} = (Y_1, \ldots, Y_d)$ have a meta-Gaussian distribution with continuous marginal distributions and copula $C^{\text{Gauss}}(\,\cdot\,|\boldsymbol{\Omega})$ and let $\Omega_{i,j}$ be the i,jth entry of $\boldsymbol{\Omega}$. Then*

$$\rho_\tau(Y_i, Y_j) = \frac{2}{\pi}\arcsin(\Omega_{i,j}), \text{ and} \tag{8.27}$$

$$\rho_S(Y_i, Y_j) = \frac{6}{\pi}\arcsin(\Omega_{i,j}/2) \approx \Omega_{i,j}. \tag{8.28}$$

Suppose, instead, that $\boldsymbol{Y}_i$*,* $i = 1, \ldots, n$*, has a meta t-distribution with continuous marginal distributions and copula* $C^{\,\mathrm{t}}(\,\cdot\,|\nu, \boldsymbol{\Omega})$*. Then (8.27) still holds, but (8.28) does not hold.*

The approximation in (8.28) uses the result that

$$\frac{6}{\pi}\arcsin(x/2) \approx x \text{ for } |x| \le 1. \tag{8.29}$$

The left- and right-hand sides of (8.29) are equal when $x = -1, 0, 1$ and their maximum difference over the range $-1 \le x \le 1$ is 0.018. However, the relative error $\left\{\frac{6}{\pi}\arcsin(x/2) - x\right\}/\frac{6}{\pi}\arcsin(x/2)$ can be larger, as much as 0.047, and is largest near $x = 0$.

By (8.28), the sample Spearman rank correlation matrix $\boldsymbol{Y}_i$, $i = 1, \ldots, n$, can be used as an estimate of the correlation matrix $\boldsymbol{\Omega}$ of $C^{\,\mathrm{Gauss}}(\,\cdot\,|\boldsymbol{\Omega})$. This estimate could be the final one or could be used as a starting value for maximum likelihood or pseudo-maximum likelihood estimation.

t-Copulas

If $\{\boldsymbol{Y}_i = (Y_{i,1}, \ldots, Y_{i,d}), i = 1, \ldots, n\}$ is a sample from a distribution with a t-copula, $C^{\,\mathrm{t}}(\,\cdot\,|\nu, \boldsymbol{\Omega})$, then we can use (8.27) and the sample Kendall's taus to estimate $\boldsymbol{\Omega}$. Let $\widehat{\rho}_\tau(Y_j, Y_k)$ be the sample Kendall's tau calculated using the samples $\{Y_{1,j}, \ldots, Y_{n,j}\}$ and $\{Y_{1,k}, \ldots, Y_{n,k}\}$ of the jth and kth variables, and let $\widetilde{\boldsymbol{\Omega}}^{**}$ be the matrix whose j,kth entry is $\sin\{\frac{\pi}{2}\widehat{\rho}_\tau(Y_j, Y_k)\}$. Then $\widetilde{\boldsymbol{\Omega}}^{**}$ will have two of the three properties of a correlation matrix; it will be symmetric with all diagonal entries equal to 1. However, it may not be positive definite, or even semidefinite, because some of its eigenvalues may be negative.

If all its eigenvalues are positive, then we will use $\widetilde{\boldsymbol{\Omega}}^{**}$ to estimate $\boldsymbol{\Omega}$. Otherwise, we alter $\widetilde{\boldsymbol{\Omega}}^{**}$ slightly to make it positive definite. By (A.47),

$$\widetilde{\boldsymbol{\Omega}}^{**} = \boldsymbol{O}\,\mathrm{diag}(\lambda_i)\,\boldsymbol{O}^\mathsf{T}$$

where $\boldsymbol{O}$ is an orthogonal matrix whose columns are the eigenvectors of $\widetilde{\boldsymbol{\Omega}}^{**}$ and $\lambda_1, \ldots, \lambda_d$ are the eigenvalues. We then define

$$\widetilde{\boldsymbol{\Omega}}^{*} = \boldsymbol{O}\,\mathrm{diag}\{\max(\epsilon, \lambda_i)\}\,\boldsymbol{O}^\mathsf{T},$$

where ϵ is some small positive quantity, for example, $\epsilon = 0.001$. Now, $\widetilde{\boldsymbol{\Omega}}^{*}$ is symmetric and positive definite, but its diagonal elements, $\widetilde{\Omega}^{*}_{i,i}$, $i = 1, \ldots, p$, may not be equal to 1. This problem is easily fixed; multiple the ith row and the ith column of $\widetilde{\boldsymbol{\Omega}}^*$ by $\left(\widetilde{\Omega}^*_{i,i}\right)^{-1/2}$, for $i = 1, \ldots, d$. The final result, which we will call $\widetilde{\boldsymbol{\Omega}}$, is a bona fide correlation matrix; that is, it is symmetric and positive definite and it has all diagonal entries equal to 1.

After $\boldsymbol{\Omega}$ has been estimated by $\widetilde{\boldsymbol{\Omega}}$, an estimate of ν is still needed. One can be obtained by plugging $\widetilde{\boldsymbol{\Omega}}$ into the log-likelihood (8.25) and then maximizing over ν.

Example 8.2. Flows in pipelines

In this example, we will continue the analysis of the pipeline flows data introduced in Example 4.3. Only the flows in the first two pipelines will be used.

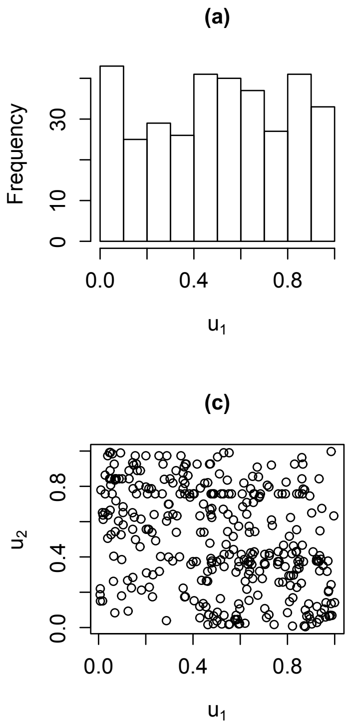

Fig. 8.6. *Pipeline data. Histograms (a) and (b) and a scatterplot (c) of the uniform-transformed flows. The empirical copula is the empirical CDF of the data in (c).*

In a fully parametric pseudo-likelihood analysis, the univariate skewed t-model will be used for flows 1 and 2. Let $U_{1,j}, \ldots, U_{n,j}$ be the flows in pipeline j, $j = 1, 2$, transformed by their estimated skewed-t CDFs. We will call the $U_{i,j}$ "uniform-transformed flows." Define $Z_{i,j} = \Phi^{-1}(U_{i,j})$, where Φ^{-1} is the standard normal quantile function. The $Z_{i,j}$ should be approximately $N(0,1)$-distributed and we will call them "normal-transformed flows."

Both sets of uniform-transformed flows should be Uniform(0,1). Figure 8.6 shows histograms of both samples of uniform-transformed flows as well as their scatterplot. The histograms show some deviations from uniform distributions, which suggests that the skewed t may not provide excellent fits and that a semiparametric pseudo-maximum likelihood approach might be tried—this will be done soon. However, the deviations may be due to random variation.

The scatterplot in Figure 8.6 shows some negative correlation as the data are somewhat concentrated along the diagonal from top left to bottom right. Thus, we can expect that the Gumbel copula, which cannot have negative dependence, will not fit well. Also, the Clayton copula may not fit well either, since the scatterplot shows data in the region where both u_1 and u_2 have small values, but this region is excluded by a Clayton copula with negative dependence. We will soon see that AIC agrees with these conclusions from a graphical analysis, since both the Clayton and Gumbel have higher (worse) AIC values compared to the Gaussian, t, and Frank copula models.

Figure 8.7 shows that the normal-transformed flows have approximately linear normal plots, as is to be expected, and their scatterplot again shows negative correlation.

We will assume for now that the two flows have a meta-Gaussian distribution. There are three ways to estimate the correlation in their Gaussian copula. The first, Spearman's sample rank correlation, is -0.357. The second, which uses (8.27) is $\sin(\pi\widehat{\tau}/2)$, where $\widehat{\tau}$ is the sample Kendall rank correlation; its value is -0.359. The third way, Pearson's correlation of the normal-transformed flows, is -0.335. There is reasonably close agreement among the three values, especially relative to their uncertainties; for example, the 95% confidence interval for the Pearson correlation of the normal-transformed flows is $(-0.426, -0.238)$, and the other two estimate are well within this interval.

Five parametric copulas were fit to the uniform-transformed flows: t, Gaussian, Gumbel, Frank, and Clayton. Since we used parametric estimates to transform the flows, we are fitting the copulas by parametric maximum pseudo-likelihood. The results are in Table 8.1. Looking at the maximized log-likelihood values, we see that the Gumbel copula fits poorly, which was to be expected since that copula only allows positive dependence and these data show negative dependence. The Frank copula fits best since it minimizes AIC, but the t and Gaussian fit reasonably well. Figure 8.8 plots uniform-transformed flows and contours of the distribution functions of five copulas: the empirical copula and four estimated parametric copulas. The t-copula is similar to the Gaussian since $\widehat{\nu}$ is large, specifically 22.3, so the t-copula was

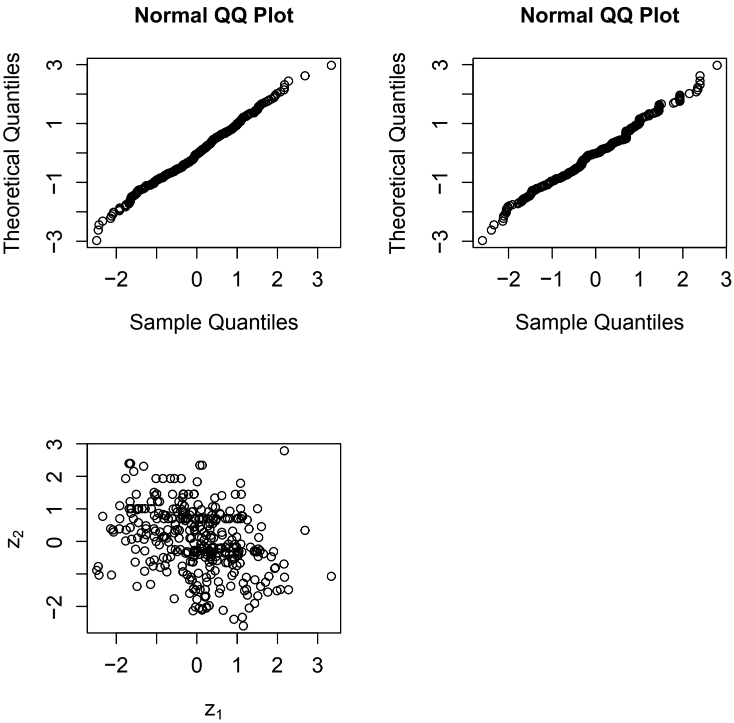

Fig. 8.7. *Pipeline data. Normal plots (a) and (b) and a scatterplot (c) of the normal-transformed flows.*

not included in the figure. The Frank copula fits best in the sense that its contours are closest to those of those of the empirical copula. This is in agreement with the AIC values.

The analysis in the previous paragraph was repeated with the flows transformed by their empirical CDFs. Doing this yielded the semiparametric pseudo-maximum likelihood estimates. Since the results were very similar to those for parametric pseudo-maximum likelihood estimates, they are not presented here.

□

8.8 Bibliographic Notes

For discussion of Archimedean copula with nonstrict generators, see McNeil, Frey, and Embrechts (2005). These authors discuss a number of other topics in more detail than is done here. They discuss methods defining nonexchangeable

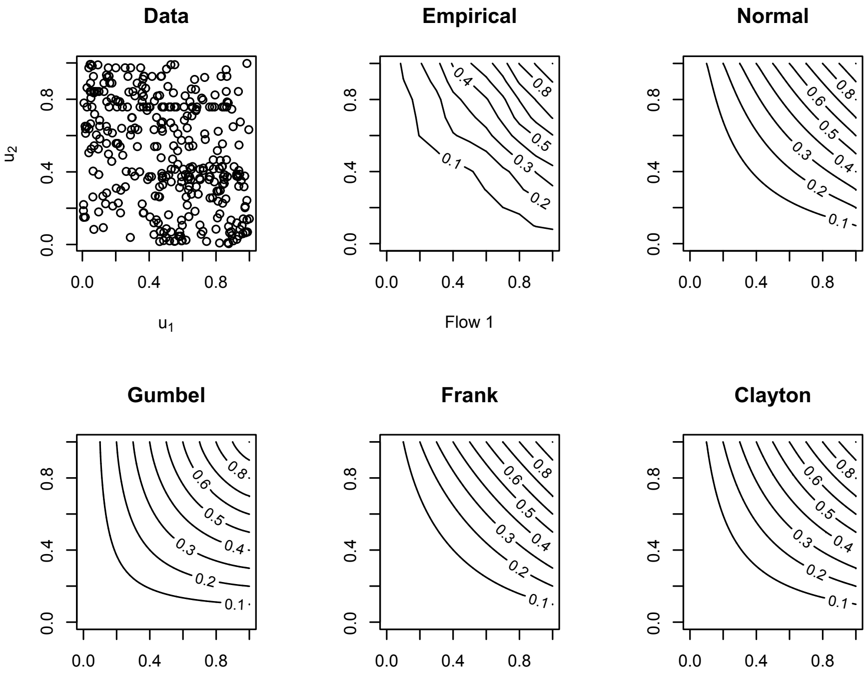

Fig. 8.8. *Uniform-transformed flows for pipeline data. Scatterplot, empirical copula, and fitted copulas using four parametric models.*

Table 8.1. *Estimates of copula parameters using the uniform-transformed pipeline flow data.*

Copula family	Estimates	Maximized log-likelihood	AIC
t	$\widehat{\rho} = -0.34$	21.0	−38.0
	$\widehat{\nu} = 22.3$		
Gaussian	$\widehat{\rho} = -0.331$	20.4	−38.8
Gumbel	$\widehat{\theta} = 0.988$	1.06	−0.06
Frank	$\widehat{\theta} = -2.25$	23.1	−44.1
Clayton	$\widehat{\theta} = -0.167$	9.87	−17.7

Archimedean copulas. The coefficients of tail dependence for Gaussian and t-copulas are derived in their Section 5.2. The theorem and calibration methods in Section 8.7.3 are discussed in their Section 5.5.

Cherubini, Luciano, and Vecchiato (2004) treat the application of copulas to finance. Joe (1997) and Nelsen (2007) are standard references on copulas.

Li (2000) developed a well-known but controversial model for credit risk using exponentially distributed default times with a Gaussian copula. An article in *Wired* magazine states that Li's Gaussian copula model was "a quick—and fatally flawed—way to assess risk" (Salmon, 2009). Duffie and Singleton's (2003) Section 10.4 also discusses copula-based methods for modeling dependent default times.

8.9 References

Cherubini, U., Luciano, E., and Vecchiato, W. (2004) *Copula Methods in Finance*, John Wiley, New York.

Duffie, D., and Singleton, K. J. (2003) *Credit Risk*, Princeton University Press, Princeton and Oxford.

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

Li, D (2000) On default correlation: A copula function approach, *Journal of Fixed Income*, **9**, 43–54.

McNeil, A., Frey, R., and Embrechts, P. (2005) *Quantitative Risk Management*, Princeton University Press, Princeton and Oxford.

Nelsen, R. B. (2007) *An Introduction to Copulas,* 2nd ed., Springer, New York.

Salmon, F. (2009) Recipe for Disaster: The Formula That Killed Wall Street, *Wired* `http://www.wired.com/techbiz/it/magazine/17-03/wp_quant?currentPage=all`

8.10 Problems

8.11 R Lab

8.11.1 Simulating Copulas

Run the R code that appears on the next page to generate data from a copula. The first line loads the `copula` library. The second line defines a copula. At this point, nothing is done with the copula—it is simply defined. However, the copula is used in the fourth line to generate a random sample. The remaining lines create a scatterplot matrix of the sample and print its sample correlation matrix.

```
library(copula)
cop_t_dim3 = tCopula(c(-.6,.75,0), dim = 3, dispstr = "un",
   df = 1)
set.seed(5640)
rand_t_cop = rcopula(cop_t_dim3,500)
pairs(rand_t_cop)
cor(rand_t_cop)
```

You can use `R`'s help to learn more about the functions `tCopula` and `rCopula`.

Problem 1 *(a) What type of copula has been sampled? (Give the copula family, the correlation matrix, and any other parameters that specify the copula.)*
(b) What is the sample size?

Problem 2 *Examine the scatterplot matrix and answer the questions below. Include the scatterplot matrix with your work.*

(a) Var 2 and Var 3 are uncorrelated. Do they seem independent? Why or why not?
(b) Do you see signs of tail dependence? If so, where?
(c) What are the effects of correlation upon the plots?
(d) The nonzero correlations in the copula do not have the same values as the corresponding sample correlations. Do you think this is just due to random variation or is something else going on? If there is another cause besides random variation, what might that be? To help answer this question, you can get confidence intervals for correlation: For example,
`cor.test(rand_t_cop[,1],rand_t_cop[,2])`
will give a confidence interval for the correlation between Var 1 and Var 2. Does this confidence interval include -0.6*?*

The first line of the following `R` code defines a normal copula. The second line defines a multivariate distribution by specifying its copula and its marginal distributions—the copula is the one just defined. The fourth line generates a random sample of size 1000 from this distribution, and the variable are labeled "Var 1," "Var 2," and "Var 3." The remaining lines create a scatterplot matrix and kernel estimates of the marginal densities.

```
cop_normal_dim3 = normalCopula(c(-.6,.75,0), dim = 3, dispstr = "un")
mvdc_normal <- mvdc(cop_normal_dim3, c("exp", "exp","exp"),
                  list(list(rate=2), list(rate = 3), list(rate=4)) )
set.seed(5640)
rand_mvdc = rmvdc(mvdc_normal,1000)
pairs(rand_mvdc)
par(mfrow=c(2,2))
plot(density(rand_mvdc[,1]))
```

```
plot(density(rand_mvdc[,2]))
plot(density(rand_mvdc[,3]))
```

Run the code above to generate the random sample.

Problem 3 *(a) What are the marginal distributions of the three variables in* `rand_mvdc`*? What are their expected values?*
(b) Are the second and third variables independent? Why or why not?

8.11.2 Fitting Copulas to Returns Data

In this section, you will fit copulas to a bivariate data set of returns on IBM and the CRSP index.

First, you will fit a model with univariate t-distributions and a t-copula. The model has three degrees-of-freedom parameters, one each for the two univariate models and a third for the copula. This means that the univariate distributions can have different tail weights and that their tail weights are independent of the tail dependence in the copula.

Run the following `R` code to load the data and necessary libraries, fit univariate t-distributions to the two variables, and convert estimated scale parameters to estimated standard deviations:

```
library(Ecdat)    #  need for the data
library(copula)   #  for copula functions
library(fGarch)   #  need for standardized t density
library(MASS)     #  need for fitdistr and kde2d
library(fCopulae)  #  additional copula functions (pempiricalCopula
                   #  and ellipticalCopulaFit)
data(CRSPday,package="Ecdat")
ibm = CRSPday[,5]
crsp = CRSPday[,7]
est.ibm = as.numeric(fitdistr(ibm,"t")$estimate)
est.crsp = as.numeric(fitdistr(crsp,"t")$estimate)
est.ibm[2] = est.ibm[2]*sqrt(est.ibm[3]/(est.ibm[3]-2))
est.crsp[2] = est.crsp[2]*sqrt(est.crsp[3]/(est.crsp[3]-2))
```

The univariate estimates will be used as starting values when the meta t-distribution is fit by maximum likelihood. You also need an estimate of the correlation coefficient in the t-copula. This can be obtained using Kendall's tau. Run the following code and complete the second line so that `omega` is the estimate of the correlation based on Kendall's tau.

```
cor_tau = cor(ibm,crsp,method="kendall")
omega =
```

Problem 4 *How did you complete the second line of code? What was the computed value of* `omega`*?*

Next, define the t-copula using `omega` as the correlation parameter and 4 as the degrees-of-freedom parameter.

```
cop_t_dim2 = tCopula(omega, dim = 2, dispstr = "un", df = 4)
```

Now fit copulas to the uniform-transformed data.

```
n = length(ibm)
data1 = cbind(pstd(ibm,mean=est.ibm[1],sd=est.ibm[2],nu=est.ibm[3]),
   pstd(crsp,mean=est.crsp[1],sd=est.crsp[2],nu=est.crsp[3]))
data2 = cbind(rank(ibm)/(n+1), rank(crsp)/(n+1))
ft1 = fitCopula(cop_t_dim2, method="L-BFGS-B", data=data1,
   start=c(omega,5),lower=c(0,2.5),upper=c(.5,15) )
ft2 = fitCopula(cop_t_dim2, method="L-BFGS-B", data=data2,
   start=c(omega,5),lower=c(0,2.5),upper=c(.5,15) )
```

Problem 5

(a) Explain the difference between methods used to obtain the two estimates `ft1` *and* `ft2`*.*
(b) Do the two estimates seem significantly different (in a practical sense)?

The next step defines a meta t-distribution by specifying its t-copula and its univariate marginal distributions. Values for the parameters in the univariate margins are also specified. The values of the copula parameter were already defined in the previous step.

```
mvdc_t_t = mvdc( cop_t_dim2, c("std","std"),
           list(list(mean=est.ibm[1],sd=est.ibm[2],nu=est.ibm[3]),
           list(mean=est.crsp[1],sd=est.crsp[2],nu=est.crsp[3]) ) )
```

Now fit the meta t-distribution. Be patient. This takes awhile; for instance, it took over four minutes on my laptop. The elapsed time in minutes will be printed.

```
start=c(est.ibm,est.crsp,ft1@est)
objFn = function(param)
{
-loglikMvdc(param, cbind(ibm,crsp), mvdc_t_t)
}
t1 = proc.time()
fit_cop = optim(start,objFn,method="L-BFGS-B",
   lower = c(-.1,.001,2.5, -.1,.001,2.5,  .2,2.5),
   upper = c(.1,.03,15,    .1,.03,15,     .8,15)
  )
t2 = proc.time()
total_time = t2-t1
total_time[3]/60
```

Lower and upper bounds are used to constrain the algorithm to stay inside a region where the log-likelihood is defined and finite. The function `fitMvdc` in the `copula` package does not allow setting lower and upper bounds and did not converge on this problem.

Problem 6

(a) What are the estimates of the copula parameters in `fit_cop`*?*
(b) What are the estimates of the parameters in the univariate marginal distributions?
(c) Was the estimation method maximum likelihood, parametric pseudo-maximum likelihood, or semiparametric pseudo-maximum likelihood?
(d) Estimate the coefficient of lower tail dependence for this copula.

Now fit normal, Gumbel, Frank, and Clayton copulas to the data.

```
fnorm = fitCopula(data=data1,copula=normalCopula(-.3,dim=2),
    method="BFGS",start=.5)
fgumbel = fitCopula(data=data1,method="BFGS",
   copula=gumbelCopula(3,dim=2),start=1)
ffrank = fitCopula(data=data1,method="BFGS",
   copula=frankCopula(3,dim=2),start=1)
fclayton = fitCopula(data=data1,method="BFGS",
    copula=claytonCopula(1,dim=2),start=1)
```

The estimated copulas (CDFs) will be compared with the empirical copula.

```
u1 = data1[,1]
u2 = data1[,2]
dem = pempiricalCopula(u1,u2)
par(mfrow=c(3,2))
contour(dem$x,dem$y,dem$z,main="Empirical")
contour(tCopula(param=ft2@est[1],df=ft2@est[2]),
   pcopula,main="t")
contour(normalCopula(fnorm@est),pcopula,main="Normal")
contour(gumbelCopula(fgumbel@est,dim=2),pcopula,
   main="Gumbel")
contour(frankCopula(ffrank@est,dim=2),pcopula,main="Frank")
contour(claytonCopula(fclayton@est,dim=2),pcopula,
   main="Clayton")
```

Problem 7 *Do you see any difference between the parametric estimates of the copula? If so, which seem closest to the empirical copula? Include the plot with your work.*

A two-dimensional KDE of the copula's density will be compared with the parametric density estimates.

```
par(mfrow=c(3,2))
contour(kde2d(u1,u2),main="KDE")
contour(tCopula(param=ft2@est[1],df=ft2@est[2]),
   dcopula,main="t",nlevels=25)
contour(normalCopula(fnorm@est),dcopula,
   main="Normal",nlevels=25)
contour(gumbelCopula(fgumbel@est,dim=2),
   dcopula,main="Gumbel",nlevels=25)
contour(frankCopula(ffrank@est,dim=2),
   dcopula,main="Frank",nlevels=25)
contour(claytonCopula(fclayton@est,dim=2),
   dcopula,main="Clayton",nlevels=25)
```

Problem 8 *Do you see any difference between the parametric estimates of the copula density? If so, which seem closest to the KDE? Include the plot with your work.*

Problem 9 *Find AIC for the t, normal, Gumbel, Frank, and Clayton copulas. Which copula model fits best by AIC? (Hint: The* `fitCopula` *function returns the log-likelihood.)*

8.12 Exercises

1. Kendall's tau rank correlation between X and Y is 0.55. Both X and Y are positive. What is Kendall's tau between X and $1/Y$? What is the Kendall's tau between $1/X$ and $1/Y$?
2. Suppose that X is Uniform(0,1) and Y^2. Then the Spearman rank correlation and the Kendall's tau between X and Y will both equal 1, but the Pearson correlation between X and Y will be less than 1. Explain why.
3. Show that the generator of a Frank copula
$$\phi^{\text{Fr}}(u) = -\log\left\{\frac{e^{-\theta u}-1}{e^{-\theta}-1}\right\}, \quad -\infty < \theta < \infty,$$
satisfies assumptions 1–3 of a strict generator.
4. Show that as $\theta \to \infty$, $C^{\text{Fr}}(u_1, u_2) \to \min(u_1, u_2)$, the co-monotonicity copula.