Coping with Copulas

Thorsten Schmidt[1]

Department of Mathematics, University of Leipzig
Dec 2006

Forthcoming in Risk Books "Copulas - From Theory to Applications in Finance"

Contents

1 Introdcution

Copulas are tools for modelling dependence of several random variables. The term *copula* was first used in the work of Sklar (1959) and is derived from the latin word *copulare*, to connect or to join. The main purpose of copulas is to describe the interrelation of several random variables.

The outline of this chapter is as follows: as a starting point we explore a small example trying to grasp an idea about the problem. The first section then gives precise definitions and fundamental relationships as well as first examples. The second section explores the most important examples of copulas. In the following section we describe measures of dependence

[1]Dep of Mathematics, Johannisgasse 26, 04081 Leipzig, Germany. Email: thorsten.schmidt@math.uni-leipzig.de
The author thanks S. L. Zhanyong for pointing out a typo in Formula (15).

as correlation and tail dependence and show how they can be applied to copulas. The fourth section shows how to simulate random variables (rvs) from the presented copulas. The final section resumes and gives a word of caution on the problems arising by the use of copulas. Finally, Table 1 at the very end of this chapter shows all the introduced copulas and their definitions.

The important issue of fitting copulas to data is examined in the next chapter of this book, “The Estimation of Copulas: Theory and Practice”, by Charpentier, Fermanian and Scaillet.

Let us start with an explanatory example: Consider two real-valued random variables X_1 and X_2 which shall give us two numbers out of $\{1, 2, \ldots, 6\}$. These numbers are the outcome of a more or less simple experiment or procedure. Assume that X_1 is communicated to us and we may enter a bet on X_2. The question is, how much information can be gained from the observation of X_1, or formulated in a different way, what is the interrelation or dependence of these two random variables (rvs).

The answer is easily found if the procedure is just that a dice is thrown twice and the outcome of the first throw is X_1 and the one of the second is X_2. In this case the variables are independent: the knowledge of X_1 gives us no information about X_2. The contrary example is when both numbers are equal, such that with X_1 we have full information on X_2.

A quite different answer will be given if X_1 is always the number of the smaller throw and X_2 the larger one. Then we have a strict monotonic relation between these two, namely $X_1 \leq X_2$. In the case where $X_1 = 6$ we also know X_2. If $X_1 = 5$, we would guess that X_2 is either 5 or 6, both with a chance of 50%, and so on.

For a deeper analysis we will need some tools which help us to describe the possible dependency of these two rvs. Observe that each rv is fully described by its cumulative distribution function (cdf) $F_i(x) := P(X_i \leq x)$ (the so-called *marginals*). In the case of throwing the dice twice we would typically have $F_1 = F_2 =: F$. However, and this is important to note, the cdfs give us no information about the joint behaviour. If we have *independence* as in the first example, the joint distribution function is simply the product of the marginals,

$$P(X_1 \leq x_1, X_2 \leq x_2) = F(x_1) \cdot F(x_2). \tag{1}$$

Hence, to obtain a full description of X_1 and X_2 together we used *two* ingredients: the marginals and the type of interrelation, in this case independence. The question is if this kind of separation between marginals and dependence can also be realized in a more general framework. Luckily the answer is yes, and the right concept for this is copulas.

To get a feeling about this, consider the third case in the above example, where X_1 (X_2) is the minimum (maximum) of the thrown dice, respectively. It is not difficult to deduce the joint distribution function,

$$P(X_1 \leq x_1, X_2 \leq x_2) = 2F(\min\{x_1, x_2\})F(x_2) - F(\min\{x_1, x_2\})^2.$$

Now, if the dice were numbered $11, 12, \ldots, 16$ instead of $1, 2, \ldots, 6$, the dependence structure obviously wouldn't change, but the joint distribution function would be totally different. This is due to the different marginal distribution. We therefore aim at finding a way to somehow disentangle marginal distribution and dependence structure. The goal is to transform the rvs X_i into uniformly distributed rvs U_i. As will be shown in Proposition 2.2, a rv X with cdf F can always be represented as $X = F^{\leftarrow}(U)$, where $F^{\leftarrow}$ denotes the generalized inverse of F as defined below. Therefore the joint distribution function can be restated, using two independent and standard uniformly distributed rvs U_1 and U_2, as

$$P\big(F_1^{\leftarrow}(U_1) \leq x_1, F_2^{\leftarrow}(U_2) \leq x_2\big) = P\big(U_1 \leq F_1(x_1), U_2 \leq F_2(x_2)\big).$$

For example, comparing the expression for the independent case (1) and the above representations of the joint distribution function we realize that the dependence structure itself – stripped from the marginals – may be expressed[2] setting $u_i = F_i(x_i) = F(x_i)$ via

$$C(u_1, u_2) = u_1 \cdot u_2$$

and this function is a *copula*. This function describes the dependence structure separated from the marginals. The intuition behind this is that the marginal distributions are transformed to uniform ones which are used as the reference case. The copula then expresses the dependence structure according to this reference case.

In the third example where we considered minimum and maximum, the two marginals are $F_1(x) = 2F(x) - F(x)^2$ and $F_2(x) = F(x)^2$. With these results at hand one receives the following copula that describes the dependence structure of the third example:

$$C(u_1, u_2) = 2\min\{1 - \sqrt{1-u_1}, \sqrt{u_2}\}\sqrt{u_2} - \min\{1 - \sqrt{1-u_1}, \sqrt{u_2}\}^2$$

2 Copulas: first definitions and examples

In this section, we give precise definitions and necessary fundamental relationships as well as first examples. To this, we mainly follow McNeil, Frey, and Embrechts (2005). Proofs as well as further details may be found therein. Starting point is the definition of a copula in d dimensions. The copula in the explanatory example was simply the distribution function of rvs with uniform marginals. It turns out that this concept is rich enough to be used for defining copulas.

Definition 2.1. A d-dimensional copula $C : [0,1]^d :\rightarrow [0,1]$ is a function which is a cumulative distribution function with uniform marginals.

In the following the notation $C(\mathbf{u}) = C(u_1, \dots, u_d)$ will always be used for a copula. The condition that C is a distribution function immediately leads to the following properties:

- As cdfs are always increasing, $C(u_1, \dots, u_d)$ is increasing in each component u_i.
- The marginal in component i is obtained by setting $u_j = 1$ for all $j \neq i$ and as it must be uniformly distributed,

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i.$$

- Clearly, for $a_i \leq b_i$ the probability $P(U_1 \in [a_1, b_1], \dots, U_d \in [a_d, b_d])$ must be nonnegative, which leads to the so-called *rectangle inequality*

$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} C(u_{1,i_1}, \dots, u_{d,i_d}) \geq 0,$$

  where $u_{j,1} = a_j$ and $u_{j,2} = b_j$.

On the other side, every function which satisfies these properties is a copula. Furthermore, if we have a d-dimensional copula then $C(1, u_1, \dots, u_{d-1})$ is again a copula and so are all k-dimensional marginals with $2 \leq k < d$. Hence, for many theoretical questions one may concentrate on two-dimensional copulas.

[2]It may be noted, that for cdfs which are not continuous the copula is not unique, and therefore in this example, also other copulas exist, which express the same dependence structure. We discuss this issue in more detail in Sklar's theorem below.

As mentioned in the explanatory example, the main goal of copulas is to disentangle marginals and dependence structure. The key to this is the following result on quantile transformations[3].

For a distribution function F we define its *generalized inverse* by

$$F^{\leftarrow}(y) := \inf\{x : F(x) \geq y\}.$$

If a rv U is uniformly distributed on $[0,1]$, we write $U \sim U[0,1]$. Throughout we use $\sim$ as notation for "is distributed as".

Proposition 2.2. *If $U \sim U[0,1]$ and F is a cumulative distribution function, then*

$$P\big(F^{\leftarrow}(U) \leq x\big) = F(x). \tag{2}$$

On the contrary, if the real-valued rv Y has a distribution function F and F is continuous, then

$$F(Y) \sim U[0,1]. \tag{3}$$

The relation (2) is typically used for simulating random variables with arbitrary cdf from uniformly distributed ones.

2.1 Sklar's theorem

Given the above result on quantile transformations, it is not surprising that every distribution function on $\mathbb{R}^d$ inherently embodies a copula function. On the other side, if we choose a copula and some marginal distributions and entangle them in the right way, we will end up with a proper multivariate distribution function. This is due to the following theorem.

Theorem 2.3. Sklar (1959) *Consider a d-dimensional cdf F with marginals $F_1, \dots, F_d$. There exists a copula C, such that*

$$F(x_1, \dots, x_d) = C\big(F_1(x_1), \dots, F_d(x_d)\big) \tag{4}$$

for all x_i in $[-\infty, \infty]$, $i = 1, \dots, d$. If F_i is continuous for all $i = 1, \dots, d$ then C is unique; otherwise C is uniquely determined only on $\operatorname{Ran} F_1 \times \dots \times \operatorname{Ran} F_d$, where $\operatorname{Ran} F_i$ denotes the range of the cdf F_i. In the explanatory example with dice, the range was $\{\frac{1}{6}, \frac{2}{6}, \dots, \frac{6}{6}\}$, while for a continuous rv this is always $[0,1]$.

On the other hand, consider a copula C and univariate cdfs $F_1, \dots, F_d$. Then F as defined in (4) is a multivariate cdf with marginals $F_1, \dots, F_d$.

It is interesting to examine the consequences of representation (4) for the copula itself. Using that $F_i \circ F_i^{\leftarrow}(y) \geq y$, we obtain

$$C(\mathbf{u}) = F\big(F_1^{\leftarrow}(u_1), \dots, F_d^{\leftarrow}(u_d)\big). \tag{5}$$

While relation (4) usually is the starting point for simulations that are based on a given copula and given marginals, relation (5) rather proves as a theoretical tool to obtain the copula from a multivariate distribution function. This equation also allows to extract a copula directly from a multivariate distribution function. So, if we refer in the sequel to a copula of rvs $X_1, \dots, X_d$, we actually refer to the copula given by (5).

It is an easy guess that for discrete distributions copulas are not as natural as they are for continuous distributions. For example, in the explanatory example at the beginning, the copula will be uniquely defined only on $\operatorname{Ran} F_1 \times \operatorname{Ran} F_2 = \{\frac{1}{6}, \dots, \frac{6}{6}\}^2$. For the independent case this was

$$C\Big(\frac{i}{6}, \frac{j}{6}\Big) = \frac{i}{6} \cdot \frac{j}{6}, \qquad i, j = 1, \dots, 6$$

and any copula satisfying this constraint is a considerable choice, for example $C(u,v) = u \cdot v$.

[3] See, for example McNeil, Frey, and Embrechts (2005), Proposition 5.2.

2.2 Copula densities

According to its definition, a copula is a cumulative distribution function. It is quite typical for these monotonically increasing functions, that albeit being theoretically very powerful their graphs are hard to interpret. Because of this, typically plots of densities are used to illustrate distributions, rather than plots of the cdf. Of course, not in all cases copulas do have densities and we will see examples very soon. However, if the copula is sufficiently differentiable the *copula density* can be computed:

$$c(\mathbf{u}) := \frac{\partial^d C(u_1, \ldots, u_d)}{\partial u_1 \cdots \partial u_d}.$$

In contrast to this, it may be the case that besides an *absolutely continuous* component (represented by the density) the copula also has a *singular* component. We will discuss this issue in the section on Marshall-Olkin copulas.

If the copula is given in form (5) we obtain the copula density in terms of the joint density together with marginal cdfs as well as marginal densities. Note that as it is necessary that the cdf is differentiable, we have $F_i^{\leftarrow} = F_i^{-1}$. Denoting the joint density by f and the marginal densities by f_i, $i = 1, \ldots, d$, it follows by the chain rule that

$$c(\mathbf{u}) = \frac{f\big(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)\big)}{f_1(F_1^{-1}(u_1)) \cdots f_d(F_d^{-1}(u_d))}.$$

In the following we will typically plot copula densities to illustrate the copulas under consideration, see for example Figures 2 and 3. However, also a realised sample is shown in Figure 4 on the right.

2.3 Conditional distributions

As pointed out in the introductory example, dependency is an important concept inferring outcomes from a rv based on the knowledge of a related factor. For an illustration, consider two uniform rvs U_1 and U_2 with known copula C and U_1 is observed. The goal is to deduce the conditional distribution which can then be used for predicting or estimating U_2. Assuming sufficient regularity, we obtain for the conditional cdf

$$\begin{aligned} P(U_2 \leq u_2 | U_1 = u_1) &= \lim_{\delta \to 0} \frac{P(U_2 \leq u_2, U_1 \in (u_1 - \delta, u_1 + \delta])}{P(U_1 \in (u_1 - \delta, u_1 + \delta])} \\ &= \lim_{\delta \to 0} \frac{C(u_1 + \delta, u_2) - C(u_1 - \delta, u_2)}{2\delta} \\ &= \frac{\partial}{\partial u_1} C(u_1, u_2). \end{aligned}$$

Hence, the conditional cdf may be derived directly from the copula itself. The conditional density function is obtained by deriving once more with respect to u_2. In most cases, the best estimator of U_2 will be the conditional expectation which of course is directly obtained from the conditional density function.

2.4 Bounds of copulas

It was Hoeffding's idea (already in the 1940's) to study multivariate distributions under "arbitrary changes of scale", and although he did not introduce copulas directly, his works contribute a lot of interesting results, see for example Fisher (1995). In this spirit is the following result:

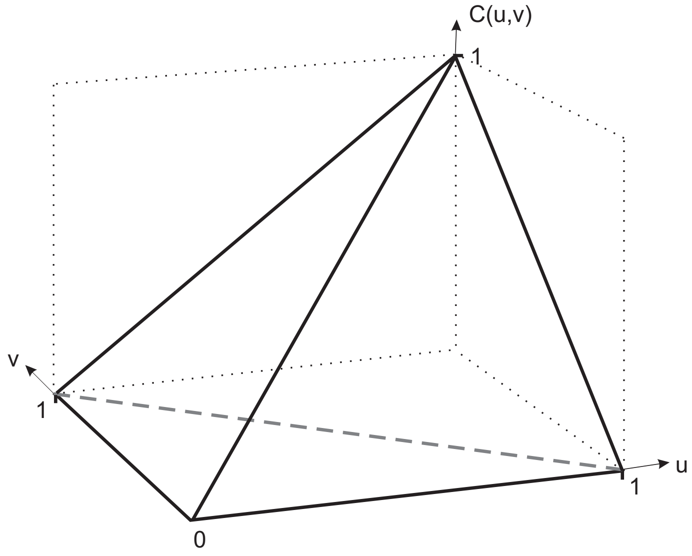

Figure 1: According to the Fréchet-Hoeffding bounds every copula has to lie inside of the pyramid shown in the graph. The surface given by the bottom and back side of the pyramid (the lower bound) is the countermonotonicity-copula $C(u, v) = \max\{u + v - 1, 0\}$, while the front side is the upper bound, $C(u, v) = \min(u, v)$.

Proposition 2.4. *Consider the rvs $X_1, \dots, X_d$ whose dependence structure is given by a copula C. Let $T_i : \mathbb{R} \to \mathbb{R}, i = 1, \dots, d$ be strictly increasing functions. Then the dependence structure of the random variables*

$$T_1(X_1), \dots, T_d(X_d)$$

is also given by the copula C.

Basically, this means that strictly increasing transformations do not change the dependence structure. On first sight, this seems to be counterintuitive. Of course, monotone transformations do change the dependence. Just after removing the effect of the marginals we end up with the same dependence structure, i.e. copula. Consider for example two normally distributed random variables X_1 and X_2 and let the covariance be the measure of dependence we look at. For $c > 0$, we could look at X_1 and cX_1 and obtain the covariance $c\,\mathrm{Cov}(X_1, X_2)$, thus the covariance has changed under this transformation. However, the correlation stayed the same, and similarly the copula did not change. We will discuss the importance of correlation for elliptical distributions in the section "measures of dependence".

Hoeffding and Fréchet independently derived that a copula always lies in between certain bounds. The reason for this is the existence of some extreme cases of dependency. To get some insight, consider two uniform rvs U_1 and U_2. Certainly, if $U_1 = U_2$ these two variables are extremely dependent on each other. In this case, the copula is given by

$$C(u_1, u_2) = P(U_1 \le u_1, U_1 \le u_2) = \min(u_1, u_2). \tag{6}$$

In terms of rvs, this copula is always attained if $X_2 = T(X_1)$, where T is a monotonic transformation. Rvs of this kind are called *comonotonic.*

Certainly, independence is quite contrary to this, and for two independent rvs the copula equals $C(u_1, u_2) = u_1 \cdot u_2$. However, independence just serves as an intermediate step on the way to the contrary extreme of comonotonicity, namely the case of *countermonotonic* rvs. In terms of uniform rvs this case is obtained for $U_2 = 1 - U_1$.

The related copula equals for $1 - u_2 < u_1$

$$C(u_1, u_2) = P(U_1 \leq u_1, 1 - U_1 \leq u_2) \tag{7}$$
$$= P(U_1 \leq u_1, 1 - u_2 \leq U_1) = u_1 + u_2 - 1 \tag{8}$$

and zero otherwise. Compare also Fig. 1 for a graph of the two copulas (6) and (8).

These considerations can be extended to the multidimensional case. However, whereas a comonotonic copula exists in any dimension d, there is no countermonotonicity copula in the case of dimensions greater than two[4]. To explain this, consider rvs X_1, X_2, X_3. We are free to choose X_1 and X_2 countermonotonic as well as X_1 and X_3. However, this already gives some restriction on the relation between X_2 and X_3. In particular, if X_1 decreases, both X_2 and X_3 have to increase, so they can not be countermonotonic again. On the other hand, even if a countermonotonic copula does not exist, the bound still holds.

Theorem 2.5. (Fréchet-Hoeffding bounds) *Consider a copula* $C(\mathbf{u}) = C(u_1, \ldots, u_d)$. *Then*

$$\max\Big\{\sum_{i=1}^{d} u_i + 1 - d, 0\Big\} \leq C(\mathbf{u}) \leq \min\{u_1, \ldots, u_d\}.$$

3 Important copulas

It is time to give the most important examples of explicit copulas. Some of them have already been mentioned, or were implicitly given in the examples.

3.1 Perfect dependence and independence

First of all, the *independence copula*

$$\Pi(\mathbf{u}) = \prod_{i=1}^{d} u_i$$

relates to the case of no dependence. As a direct consequence of Sklar's theorem we obtain that rvs are independent if and only if their copula is the independence copula. The related copula density is simply constant.

As already mentioned, the Fréchet-Hoeffding bounds are also related to copulas, where the *comonotononicity copula* or the *Fréchet-Hoeffding upper bound* is given by

$$M(\mathbf{u}) = \min\{u_1, \ldots, u_d\}.$$

Comonotonicity refers to the following case of perfect positive dependence. Consider strictly increasing transformations $T_2, \ldots, T_d$ and and set $X_i = T_i(X_1)$ for $i = 2, \ldots, d$. Using Proposition 2.4 we see that these random variables indeed have the comonotonicity copula.

The other extreme is given by countermonotonicity and as already mentioned, the related bound is only attained in the two-dimensional case. The *countermonotonicity copula* or *Fréchet-Hoeffding lower bound* reads

$$W(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\}.$$

[4]See McNeil, Frey, and Embrechts (2005), Example 5.21, for a counter-example.

This case refers to perfect negative dependence in the sense that $X_2 = T(X_1)$ with a strictly decreasing function T.

These two copulas do not have copula densities as they both admit a kink and therefore are not differentiable. It may be recalled, that conditionally on U_1, U_2 is perfectly determined, for example $U_2 = U_1$ in the comonotonicity case. Hence the distribution has mass only on the diagonal $u_1 = u_2$ and this is the reason why this copula can not be described by a density. Similar, in the countermonotonicity case there is mass only on $\{u_1 = 1 - u_2\}$.

3.2 Copulas derived from distributions

Quite exchangeably we used the expression copula from rvs $X_1, \ldots, X_n$ and copula according to its precise definition. This suggests that typical multivariate distributions describe important dependence structures. The copulas derived therefrom shall be introduced in this section. The multivariate normal distribution will lead to the Gaussian copula while the multivariate Student t-distribution leads to the t-copula.

First, consider two normally distributed rvs X_1 and X_2 which are also jointly normal. It is well known that their correlation

$$\mathrm{Corr}(X_1, X_2) := \frac{\mathrm{Cov}(X_1, X_2)}{\sqrt{\mathrm{Var}(X_1) \cdot \mathrm{Var}(X_2)}} \tag{9}$$

fully describes the dependence structure. This remains true in the whole family of elliptical distributions[5], while it is totally wrong outside this family. Unfortunately, this is a typical pitfall which frequently occurs when measuring correlation and deducing results from the dependence structure outside this family (see, for example, Embrechts et al. (2002)). We will clarify this in the following section, when we explore different measures of dependence.

Following relation (5), we obtain the two-dimensional *Gaussian copula*

$$C_\rho^{Ga}(u_1, u_2) = \Phi_{\mathbf{\Sigma}}\big(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\big),$$

where $\mathbf{\Sigma}$ is the 2×2 matrix with 1 on the diagonal and ρ otherwise, explicitly stated in (10) below. Φ denotes the cdf of a standard normal distribution while $\Phi_{\mathbf{\Sigma}}$ is the cdf for a bivariate normal distribution with zero mean and covariance matrix $\mathbf{\Sigma}$. Note that this representation is equivalent to

$$\int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(- \frac{s_1^2 - 2\rho s_1 s_2 + s_2^2}{2(1-\rho^2)} \right) ds_1\, ds_2.$$

For normal and elliptical distributions independence is equivalent to zero correlation. Hence for $\rho = 0$, the Gaussian copula equals the independence copula. On the other side, if $\rho = 1$ we obtain the comonotonicity copula, while for $\rho = -1$ the countermonotonicity copula is obtained. The intuition of positive or negative dependence recurs in the form of positive/negative linear dependence in this example. Thus the Gaussian copula interpolates between these three fundamental dependency structures via one simple parameter, the correlation ρ. A copula with this feature is named *comprehensive*.

Note that the covariance matrix $\mathbf{\Sigma}$, used in the above example, is not arbitrary. In fact it is a *correlation matrix*, which is obtained from an arbitrary covariance matrix by scaling each component to variance 1, which we illustrate in the following example. According to Proposition 2.4 this does not change the copula.

[5] Elliptical distributions are considered in more detail in the section on "measures of dependence".

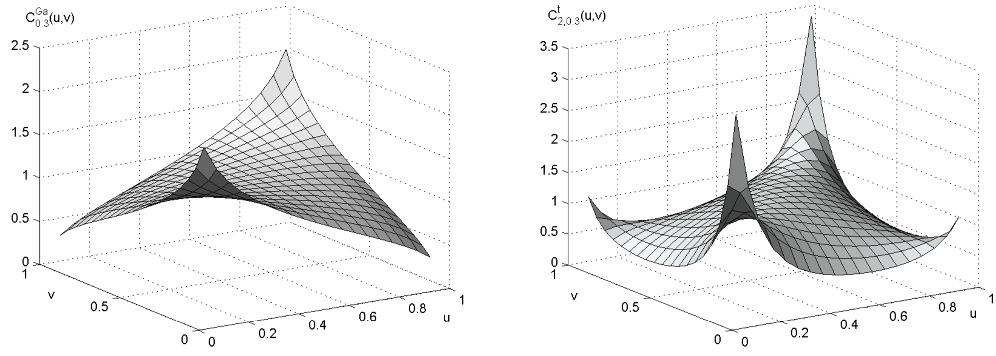

Figure 2: The copula densities of a Gaussian copula (left) and a Student t-copula (right). Both copulas have correlation coefficient $\rho = 0.3$ and the t-copula has 2 degrees of freedom.

Example 3.1. The covariance matrix

$$\tilde{\mathbf{\Sigma}} = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}$$

leads to the correlation matrix $\mathbf{\Sigma}$, already mentioned above,

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}. \tag{10}$$

In the multivariate case the *Gaussian copula* for a *correlation matrix* $\mathbf{\Sigma}$ is given by

$$C^{Ga}_{\mathbf{\Sigma}}(\mathbf{u}) = \Phi_{\mathbf{\Sigma}}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)).$$

A plot of the bivariate Gaussian copula density with correlation 0.3 is given in the left part of Figure 2.

In statistics, the Student t-distribution arises in the context of tests on the mean of normally distributed random variables when the variance is not known but has to be estimated. Quite natural, the t-distribution is a so-called mixture of a normal distribution. A mixture occurs if a parameter of the main distribution is itself random. A rv η from a t-distribution with ν degrees of freedom can always be represented as

$$\eta = \frac{X_1}{\sqrt{\xi/\nu}} = \frac{\sqrt{\nu} \cdot X_1}{\sqrt{Y_1^2 + \cdots + Y_\nu^2}},$$

where X_1 as well as $\mathbf{Y} = Y_1, \ldots, Y_\nu$ are standard normal, while ξ has a χ^2_ν-distribution, X_1 and $\mathbf{Y}$ being independent. The χ^2 distribution stems from sums of squared normal rvs, which was used for the second equality.

The *multivariate t-distribution* in d dimensions with ν degrees of freedom is obtained from $\mathbf{X} = X_1, \ldots, X_d \sim \mathcal{N}(0, \mathbf{\Sigma})$ via

$$(\eta_1, \ldots, \eta_d) = \Big(\frac{X_1}{\sqrt{\xi/\nu}}, \ldots, \frac{X_d}{\sqrt{\xi/\nu}}\Big), \tag{11}$$

where ξ is again χ^2_ν, independent from $\mathbf{X}$. If $d = 2$ and if the correlation of X_1 and X_2 is 1 (or -1) we again have comonotonicity (countermonotonicity, resp.) while it is important to notice that for $\rho = 0$ we do not have independence, as some dependence is introduced via ξ.

Formally stated the *t-copula* or the *Student copula* is given by

$$C^{\,t}_{\nu,\mathbf{\Sigma}}(\mathbf{u}) = t_{\nu,\mathbf{\Sigma}}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)), \tag{12}$$

where $\mathbf{\Sigma}$ is a correlation matrix, t_ν is the cdf of the one dimensional t_ν distribution and $t_{\nu,\mathbf{\Sigma}}$ is the cdf of the multivariate $t_{\nu,\mathbf{\Sigma}}$ distribution. The correlation matrix is obtained from an arbitrary covariance matrix by scaling each component to variance 1, compare Example 3.1. In the bivariate case one uses the simpler notation $C^t_{\nu,\rho}$ referring directly to the correlation analogously to the Gaussian copula.

The bivariate t-copula density for 2 degrees of freedom and correlation 0.3 is plotted in the right part of Figure 2. Observe that the behaviour at the four corners is different from the Gaussian copula while in the center they are quite similar. This means, although having the same correlation, the extreme cases (which correspond to the four corner points) are much more pronounced under the t-copula. Particularly in applications in finance, where the rvs under consideration describe losses of a portfolio, values in the $(0,0)$ corner correspond to big losses in both entities of the portfolio. The fact that the t-copula is able to model such extreme cases is due to the *tail-dependence* which we will examine more closely soon.

Furthermore, the t-copula also shows peaks at the $(0,1)$ and $(1,0)$ corners which stem from its mixing nature. Suppose that the mixing variable ξ has a low value. Then the peaks in theses respective corners stem from a positive value in X_1 and a negative one in X_2. For the independent case, this is as likely as two values with the same sign, so in this case the density rises up at all four corners symmetrically. However, introducing some correlation changes this probability. In Figure 2 the correlation was 0.3, so it is more likely to have values with the same sign, which is perfectly in line with the observation that their peaks are less pronounced as the peaks at $(0,0)$ and $(1,1)$.

3.3 Copulas given explicitly

In contrast to the copulas derived from multivariate distributions, there are a number of copulas which can be stated directly and have a quite simple form. They typically fall in the class of Archimedean copulas, and we give a short introduction to Archimedean copulas in the bivariate case. It is also possible to define Archimedean copulas in the multivariate case and we refer to McNeil, Frey, and Embrechts (2005), Section 5.4.2 for this issue. Further examples of Archimedean copulas may be found in Nelsen (1999), in particular on may consider Table 4.1 therein.

After this we shortly discuss the bivariate Marshall-Olkin copula. For the multivariate extension we refer to Embrechts, Lindskog, and McNeil (2003).

3.3.1 Archimedean copulas

The bivariate *Gumbel copula* or *Gumbel-Hougaard copula* is given in the following form:

$$C^{\,Gu}_{\theta}(u_1, u_2) = \exp\Big[- \big((-\ln u_1)^\theta + (-\ln u_2)^\theta\big)^{\frac{1}{\theta}} \Big],$$

where $\theta \in [1, \infty)$. For $\theta = 1$ we discover the independence copula, while for the $\theta \to \infty$ the Gumbel copula tends to the comonotonicity copula so that the Gumbel copula interpolates between independence and perfect positive dependence. The Gumbel copula comes as an example of a simple copula which has tail dependence in one corner which we explore in more detail in the following section.

A second example is the *Clayton copula*, given by

$$C_\theta^{Cl}(u_1,u_2) = \Big(\max\{u_1^{-\theta} + u_2^{-\theta} - 1, 0\}\Big)^{-\frac{1}{\theta}},$$

where $\theta \in [-1,\infty)\setminus\{0\}$. For the limits $\theta \to 0$ we obtain the independence copula, while for $\theta \to \infty$ the Clayton copula arrives at the comonotonicity copula. For $\theta = -1$ we obtain the Fréchet-Hoeffding lower bound. Thus, as the Gumbel copula, the Clayton copula interpolates between certain dependency structures, namely countermonotonicity, independence and comonotonicity. Both these copulas belong to the family of Archimedean copulas, which we will introduce now.

The above examples suggest that it might be possible to generate quite a number of copulas from interpolating between certain copulas in a clever way. Indeed, the family of *Archimedean copulas* is a useful tool to generate copulas.

Revisiting the two examples above more closely, one realizes that the copula itself was always of the form

$$C(u_1,u_2) = \phi^{-1}\Big(\phi(u_1) + \phi(u_2)\Big),$$

where ϕ was a decreasing function mapping $[0,1]$ into $[0,\infty]$. Indeed, with

$$\begin{aligned}\phi_{Gu}(u) &= (-\ln u)^\theta,\\ \phi_{Cl}(u) &= \frac{1}{\theta}(u^{-\theta} - 1)\end{aligned}$$

we obtain the Gumbel and the Clayton copula, respectively[6]. The function ϕ is called the *generator* of the copula.

Theorem 3.2. *Consider a continuous and strictly decreasing function $\phi : [0,1] :\longrightarrow [0,\infty]$ with $\phi(1) = 0$. Then*

$$C(u_1,u_2) = \begin{cases}\phi^{-1}(\phi(u_1)+\phi(u_2)) & \text{if } \phi(u_1)+\phi(u_2) \le \phi(0),\\ 0 & \text{otherwise}\end{cases}$$

is a copula, if and only if ϕ is convex.

A proof can be found in Nelsen (1999), p. 111.

If $\phi(0) = \infty$ the generator is said to be *strict* and $C(u_1,u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$. For example, the generator of the Clayton copula is strict only if $\theta > 0$. For $\theta > 0$ we obtain the representation

$$\big(u_1^{-\theta} + u_2^{-\theta} - 1\big)^{-\frac{1}{\theta}}.$$

Using this approach, one is able to generate quite a number of copulas. For example, the generator $\ln(e^{-\theta} - 1) - \ln(e^{-\theta u} - 1)$ leads to the *Frank copula* given by

$$C_\theta^{Fr}(u_1,u_2) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u_1} - 1)\cdot(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right),$$

for $\theta \in \mathbb{R}\setminus\{0\}$. Another example is the *generalized Clayton copula* obtained from the generator $\theta^{-\delta}(u^{-\theta} - 1)^\delta$:

$$C_{\theta,\delta}^{Cl}(u_1,u_2) = \Big([(u_1^{-\theta} - 1)^\delta + (u_2^{-\theta} - 1)^\delta]^{\frac{1}{\delta}} + 1\Big)^{-\frac{1}{\theta}},$$

[6] Note, that for generating the Clayton copula it would be sufficient to use $(u^{-\theta} - 1)$ instead of $\frac{1}{\theta}(u^{-\theta} - 1)$ as generator. However, for negative θ this function is increasing, and Theorem 1.7 would not apply.

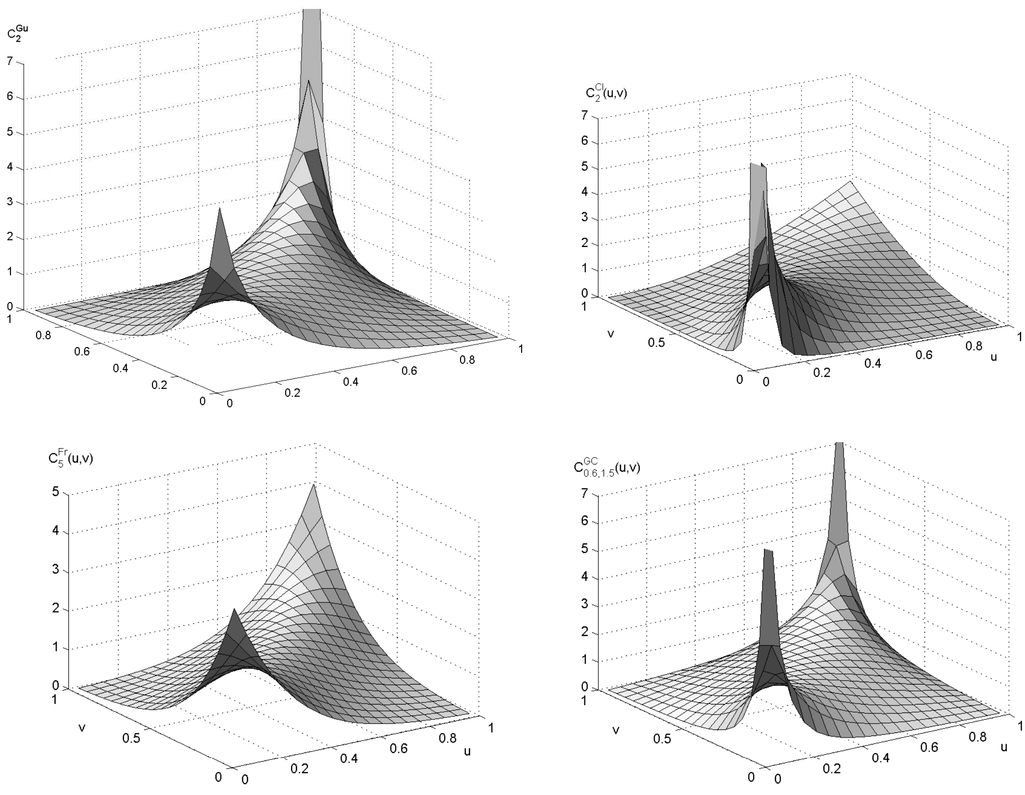

Figure 3: Densities of the Gumbel (upper left, $\theta = 2$), Clayton (upper right, $\theta = 2$), Frank (lower left, $\theta = 2$) and generalized Clayton (lower right, $\theta = \delta = 2$) copulas. Note that all copulas except for the Frank copula have been cut at a level of 7.

with $\theta > 0$ and $\delta \geq 1$. Note that for $\delta = 1$ the standard Clayton copula is attained. Furthermore, it is not possible to generalize this copula to higher dimensions in this case.

To illustrate the differences of the presented copulas, we show their densities in Figure 3. It may be spotted at first sight, that the presented copulas have different behaviour at the left lower and right upper (the points $(0, 0)$ and $(1, 1)$, respectively) corners. In a qualitative manner, the Gumbel copula shows an extremely uprising peak at $(1, 1)$ while a less pronounced behaviour at $(0, 0)$. As we will see later, a mathematical analysis of the so-called tail behaviour reflects this precisely by saying that the Gumbel copula has upper tail dependence. For the Clayton copula the situation is reverse, while for the Frank copula no tail dependence will show up. It may be glimpsed that the standard Clayton copula differs quite dramatically from the generalized one in the behaviour at the corners. The reason for this is that the generalized Clayton copula shows tail behaviour at both corners in contrast to the standard one. We will discuss this later.

An additional observation is that the t-copula is the only one which shows peaks (small ones, though) at the $(0, 1)$ and $(1, 0)$ corners (see the right part of Figure 2). As was already mentioned, this is due to the mixing nature of the t-distribution.

3.3.2 Marshall-Olkin copulas

As a final example of copulas we discuss the *Marshall-Olkin copulas*. For a motivation, consider two components which are subject to certain shocks which lead to failure of either one of them or both components. The shocks occur at times assumed to be independent and exponentially distributed with parameters λ_1, λ_2 and λ_{12}, respectively. Here the subindex 1 or 2 refers to the

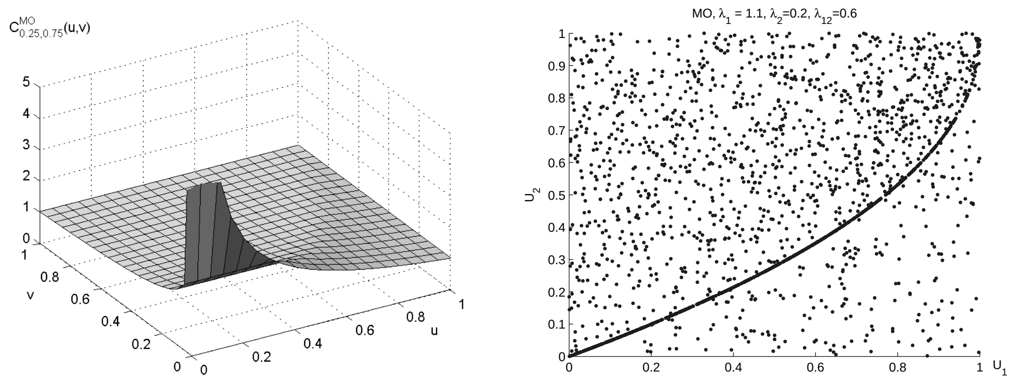

Figure 4: The Marshall-Olkin copula. Left: Continous part of the density, cut at 5. Right: Simulation of 2000 rvs with a Marshall-Olkin copula. The parameters are $\alpha_1 = 0.25$ and $\alpha_2 = 0.75$ or $\lambda_1 = 1.1$, $\lambda_2 = 0.2$ and $\lambda_{12} = 0.6$, respectively.

shock effecting a single component only, while the shock affecting both entities is referenced by the subindex 12. Denote the realized shock times by Z_1, Z_2 and Z_{12}. Then we obtain for the probability that the components live longer than x_1 and x_2, respectively,

$$P(Z_1 > x_1)\, P(Z_2 > x_2)\, P(Z_{12} > \max\{x_1, x_2\}).$$

The related copula can be computed with small effort and equals

$$C_{\alpha_1,\alpha_2}(u_1, u_2) = \min\{u_2 \cdot u_1^{1-\alpha_1}, u_1 \cdot u_2^{1-\alpha_2}\}, \tag{13}$$

where we set $\alpha_i = \lambda_{12}/(\lambda_i + \lambda_{12})$, $i = 1, 2$.

A copula of the form (13) with $\alpha_i \in [0, 1]$ is called *bivariate Marshall-Olkin copula* or *generalized Cuadras-Augé copula*. In contrast to the copulas presented up to now, this copula has both an absolutely continuous and a singular component. For the absolutely continuous component we can compute the density and obtain

$$\frac{\partial^2}{\partial u_1 \partial u_2} C_{\alpha_1,\alpha_2}(u_1, u_2) = \begin{cases} u_1^{-\alpha_1} & \text{if } u_1^{\alpha_1} > u_2^{\alpha_2}, \\ u_2^{-\alpha_2} & \text{if } u_1^{\alpha_1} < u_2^{\alpha_2}. \end{cases}$$

The continuous part of the density is shown in the left part of Figure 4. The singular component distributes mass only on the curve $u_1 = u_2^{\alpha_2/\alpha_1}$. More precisely, one can show[7] that for uniform U_1 and U_2 with this copula

$$P(U_1 = U_2^{\alpha_2/\alpha_1}) = \frac{\alpha_1 \alpha_2}{\alpha_1 + \alpha_2 - \alpha_1 \alpha_2}.$$

This effect becomes visible from the simulation data presented in Figure 4 (right). The discontinuous part of the copula results in data points frequently lying on the curve $u_1 = u_2^{\alpha_2/\alpha_1}$ directly.

For more details on Marshall-Olkin copulas, in particular the multivariate ones, we refer to Embrechts, Lindskog, and McNeil (2003) and Nelsen (1999).

[7]See Nelsen (1999), Section 3.1.1.

4 Measures of dependence

Measures of dependence are common instruments to summarize a complicated dependence structure in a single number (in the bivariate case). There are three important concepts: The classical one is the linear correlation as defined in Equation (9). Unfortunately, correlation is only a suitable measure in a special class of distributions, i.e. elliptical distributions. This class includes the normal distribution and mixtures of normal distributions. It is well known, that outside this class correlation leads to a number of fallacies. The two other dependence measures to be considered are rank correlation and the coefficients of tail dependence. Both measures are general enough to give sensible measures for any dependence structure. As already mentioned above, tail dependence considers dependence in the extremes which is an important concept in many financial applications.

4.1 Linear correlation

Linear correlation is a well studied concept. It is a dependence measure which is, however, useful only for *elliptical distributions*. An elliptical distribution is obtained by an affine transformation

$$\mathbf{X} = \mu + A\mathbf{Y}$$

of a spherical distribution $\mathbf{Y}$ with $\mu \in \mathbb{R}^d, A \in \mathbb{R}^{d\times d}$. By definition $\mathbf{Y}$ has a spherical distribution, if and only if the characteristic function can be represented as

$$\mathbb{E}(\exp(i\,\mathbf{t}'\mathbf{Y})) = \psi\left(t_1^2 + \cdots + t_d^2\right),$$

with some function $\psi : \mathbb{R} \mapsto \mathbb{R}$. The normal distribution plays quite an important role in spherical distributions, so in most cases spherical distributions are mixtures of normal distributions[8].

It seems most important to mention certain pitfalls which occur when correlation is considered outside the class of elliptical distributions, compare Embrechts et al. (2002) or McNeil, Frey, and Embrechts (2005), Chapter 5.2.1.

1. A correlation of 0 is equivalent to independence for normal distributions. However, already for Student t distributed rvs, this is no longer true, compare Equation (11) and the discussion thereafter.
2. Correlation is invariant under *linear* transformations, but not under *general transformations*. For example, two log-normal rvs have a different correlation than the underlying normal rvs.
3. On the other side, one could guess, that for any marginal distribution and given correlation ρ it would be possible to construct a joint distribution. However, whereas this is possible for the class of elliptical distributions, this is not true in general. For example, in the case of log-normal marginals, the interval of attainable correlations becomes smaller with increasing volatility. To illustrate this, consider two normal rvs X_1 and X_2, both with zero mean and variance $\sigma^2 > 0$. Denote the correlation of X_1 and X_2 by ρ. The correlation of the two log-normal rvs $Y_i = \exp(X_i), i = 1, 2$ equals

$$\mathrm{Corr}(Y_1, Y_2) = \frac{e^{\rho\sigma^2} - 1}{\sqrt{\left(e^{\sigma^2} - 1\right)\left(e^{\sigma^2} - 1\right)}}.$$

[8]Compare McNeil, Frey, and Embrechts (2005), Theorem 3.25. Section 3.3 therein gives a short introduction to elliptical distributions.

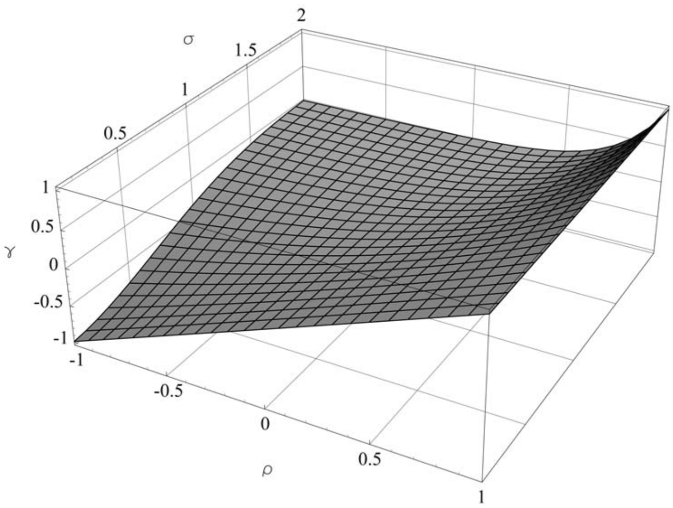

Figure 5: The graph shows $\rho = \mathrm{Corr}(Y_1, Y_2)$ where $Y_i = \exp(\sigma X_i)$ with $X_i \sim \mathcal{N}(0, \sigma^2)$ and $\mathrm{Corr}(X_1, X_2) = \rho$. Note that the smallest attained correlation is increasing with σ, so for $\sigma = 1$ we have that $\mathrm{Corr}(Y_1, Y_2) \geq -0.368$ and for $\sigma = 2$ even $\mathrm{Corr}(Y_1, Y_2) \geq -0.018$.

For $\rho = 1$ we always obtain $\mathrm{Corr}(Y_1, Y_2) = 1$. However, the smallest correlation is always larger than -1. For example, if $\sigma = 1$ we obtain $\mathrm{Corr}(Y_1, Y_2) \in [-0.368, 1]$. In Figure 5 a plot of the attained correlation as function of σ and ρ is given and it shows that the interval of attained correlation becomes smaller with increasing σ.

On the other side, this observation implies that it is in general wrong to deduce a small degree of dependency from a small correlation. Even perfectly related rvs can have zero correlation: Consider $X_1 \sim \mathcal{N}(0, 1)$ and $X_2 = X_1^2$. Then

$$\mathrm{Cov}(X_1, X_2) = E\Big(X_1 \cdot (X_1^2 - 1)\Big) = E(X_1^3) - E(X_1) = 0.$$

Having covariance 0 implies of course zero correlation, while on the other side the observation of X_1 immediately yields full knowledge of X_2.

4.2 Rank correlation

It is a well-known concept in nonparametric statistics to concentrate on the ranks of given data rather than on the data itself. This led to important correlation estimators, *Kendall's tau* and *Spearman's rho*, which we present directly in a form related to copulas. Considering ranks leads to scale invariant estimates, which in turn is very pleasant when working with copulas. Therefore rank correlations give a possible way of fitting copulas to data.

The motivation of *Spearman's rho* is very much along the lines of the quantile transform given in Equation 3. It should be noted that by applying the empirical cdf to the data itself the ranks of the data points are obtained, up to a multiplicative factor. The idea is simply to take the correlation of these transformed variables, which is equivalent to the correlation of the ranks itself. Considering a formulation suitable to copulas one arrives at the following definition.

Definition 4.1. Consider two rvs X_1 and X_2 with marginals F_1 and F_2, respectively. We define Spearman's rho by

$$\rho_S := \mathrm{Corr}\,\big(F_1(X_1), F_2(X_2)\big),$$

correlation being defined in (9). In the multivariate case, the Spearman's rho matrix is

$$\rho_S(\mathbf{X}) := \mathrm{Corr}\,\big(F_1(X_1), \ldots, F_d(X_d)\big).$$

In the multivariate case, Corr refers to the correlation matrix, in other words the entries of $\rho_S(\mathbf{X})$ are given by

$$\rho_S(\mathbf{X})_{ij} = \text{Corr}\big(F_i(X_i), F_j(X_j)\big).$$

In the bivariate case Spearman's rho was simply defined as the correlation of the uniform rvs $F_1(X_1)$ and $F_2(X_2)$.

To motivate *Kendall's tau*, we consider two random variables X_1 and X_2. For a comparison we take two additional rvs $\tilde{X}_1$ and $\tilde{X}_2$ into account, both being independent of the first two rvs but with the same joint distribution. Now we plot two points, obtained from these rvs, in a graph, namely (x_1, x_2) and $(\tilde{x}_1, \tilde{x}_2)$, and connect them by a line. In the case of positive dependence we would expect that the line is increasing and otherwise that the line is decreasing. Considering $(X_1 - \tilde{X}_1) \cdot (X_2 - \tilde{X}_2)$, a positive sign refers to the increasing case, while a negative sign would turn up in the decreasing case. Taking expectations leads to the following definition.

Definition 4.2. Consider two rvs X_1 and X_2 and denote by $\tilde{X}_1, \tilde{X}_2$ rvs with the same joint distribution, but independent of X_1, X_2. Then we define Kendall's tau by

$$\rho_\tau(X_1, X_2) := \mathbb{E}\Big[\operatorname{sign}\Big((X_1 - \tilde{X}_1)\cdot(X_2 - \tilde{X}_2)\Big)\Big]. \tag{14}$$

For a d-dimensional rv $\mathbf{X}$ and an independent copy $\tilde{\mathbf{X}}$ we define Kendall's tau by

$$\rho_\tau(\mathbf{X}) := \text{Cov}\big[\operatorname{sign}\big(\mathbf{X} - \tilde{\mathbf{X}}\big)\big].$$

Recall, that by an independent copy of $\mathbf{X}$ we mean a rv with the same distribution as $\mathbf{X}$, being independent of $\mathbf{X}$. Alternatively, the expectation in (14) can be written as

$$\begin{aligned}\rho_\tau(X_1, X_2) = P\big((X_1 - \tilde{X}_1)\cdot(X_2 - \tilde{X}_2) > 0\big) \\ - P\big((X_1 - \tilde{X}_1)\cdot(X_2 - \tilde{X}_2) < 0\big).\end{aligned}$$

Thus if both probabilities are the same, which intuitively means that we expect upward slopes with the same probability as downward slopes, we obtain $\rho_\tau = 0$. On the other side, if Kendall's tau is positive, there is a higher probability of upward slopes and similarly for a negative value we would expect rather downward sloping outcomes.

Both measures have a lot of properties in common. Being obviously measures with values in $[-1, 1]$, they take the value 0 for independent variables (while there might also be non-independent rvs with zero rank correlation) and return 1 (-1) for the comonotonic (countermonotonic) case. Moreover, they can directly be derived from the unique copula C that describes the dependence between X_1 and X_2[9]:

$$\rho_S(X_1, X_2) = 12 \int_0^1 \int_0^1 \big(C(u_1, u_2) - u_1 u_2\big)\, du_1\, du_2$$

$$\rho_\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2)\, dC(u_1, u_2) - 1.$$

Summarizing, using rank correlations helps to solve point 3 of the above remarked pitfalls. Indeed, for given marginals any rank correlation in $[-1, 1]$ may be attained.

[9]Compare McNeil, Frey, and Embrechts (2005), Proposition 5.29.

The rank correlations derived from a bivariate Gaussian copula are given by[10]

$$\rho_S(X_1, X_2) = \frac{6}{\pi} \arcsin \frac{\rho}{2}, \qquad \rho_\tau(X_1, X_2) = \frac{2}{\pi} \arcsin \rho.$$

In this case Spearman's rho is very close to the actual correlation.

For other interesting examples and certain bounds which interrelate those two measures we refer to Nelsen (1999), Sections 5.1.1 – 5.1.3.

4.3 Tail dependence

Finally, we introduce the *tail dependence* of copulas, where we distinguish between *upper* and *lower* tail dependence. It may be recalled, that the copula densities of the Gaussian and the Student t-copula showed different behaviour at both the left lower and right upper corner (compare Figure 2). The right upper corner refers to upper tail dependence, while the lower left corner refers to lower tail dependence. The aim of this section is to consider this behaviour in more detail.

Consider two uniform rvs U_1 and U_2 with copula C. *Upper tail dependence* means intuitively, that with large values of U_1 also large values of U_2 are expected. More precisely, the probability that U_1 exceeds a given threshold q, given that U_2 has exceeded the same level q for $q \to 1$ is considered. If this is smaller than of order q, then the rvs have no tail dependence, like for example in the independent case. Otherwise they have tail dependence.

Definition 4.3. For rvs X_1 and X_2 with cdfs $F_i, i = 1, 2$ we define the *coefficient of upper tail dependence* by

$$\lambda_u := \lim_{q \nearrow 1} P\big(X_2 > F_2^{\leftarrow}(q) \big| X_1 > F_1^{\leftarrow}(q)\big),$$

provided that the limit exists and $\lambda_u \in [0, 1]$. The *coefficient of lower tail dependence* is defined analogously by

$$\lambda_l := \lim_{q \searrow 0} P\big(X_2 \leq F_2^{\leftarrow}(q) \big| X_1 \leq F_1^{\leftarrow}(q)\big).$$

If $\lambda_u > 0$, we say that X_1 and X_2 have upper tail dependence, while for $\lambda_u = 0$ we say that they are *asymptotically independent* in the upper tail and analogously for λ_l.

For continuous cdfs, quite simple expressions are obtained for the coefficients using Bayes' rule, namely

$$\begin{aligned} \lambda_l &= \lim_{q \searrow 0} \frac{P\big(X_2 \leq F_2^{\leftarrow}(q), X_1 \leq F_1^{\leftarrow}(q)\big)}{P(X_1 \leq F_1^{\leftarrow}(q))} \\ &= \lim_{q \searrow 0} \frac{C(q,q)}{q} \end{aligned}$$

and, similarly,

$$\lambda_u = 2 + \lim_{q \searrow 0} \frac{C(1-q, 1-q) - 1}{q}. \tag{15}$$

Thus, basically, we need to look at the slope of the copula when approaching $(0, 0)$ or $(1, 1)$. If the slope is greater than 1 (which refers to the independent case), the copula exhibits tail dependence. Of course, the greater the slope is, the higher the tail dependence.

[10] Compare McNeil, Frey, and Embrechts (2005), Theorem 5.36. They also give a discussion which consideres these measures for general elliptic distributions.

Example 4.4. The tail coefficients for Gumbel and Clayton copulas are easily computed. Consider for example the Clayton copula. According to Figure 3 there is no upper tail dependence. The coefficient of lower tail dependence equals

$$\begin{aligned}\frac{(2q^{-\theta}-1)^{-1/\theta}}{q} &= \left(2-q^{\theta}\right)^{-1/\theta} \\ &\to 2^{-1/\theta} = \lambda_l.\end{aligned}$$

Thus, for $\theta > 0$, the Clayton copula has lower tail dependence. Furthermore, for $\theta \to \infty$ the coefficient converges two 1. This is because the Clayton copula tends to the comonotonicity copula as θ goes to infinity.

It is little more complicated to show that for the Gumbel copula $\lambda_u = 2 - 2^{1/\theta}$, thus the Gumbel copula exhibits upper tail dependence for $\theta > 1$.

This example reflects the impressions from Figure 3, which already suggested that the Gumbel copula has upper tail dependence, while the Clayton copula has lower tail dependence.

As the Gaussian copula as well as the Student t-copula are not given as explicitly as the Clayton or Gumbel copula, computing the tail dependence is more complicated. In McNeil, Frey, and Embrechts (2005), Section 5.3.1, this is done for both and we revisit shortly their results. First, the Gaussian copula is asymptotically independent in both upper and lower tails. This means, no matter what high correlation we choose, there will be no tail dependence from a Gaussian copula. However, for the Student t-distribution the matters are different. For the bivariate t-copula $C^t_{\nu,\rho}$, the coefficient of tail dependence equals

$$\lambda_l = \lambda_u = 2\,t_{\nu+1}\Bigg(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\Bigg),$$

provided that $\rho > -1$. Thus, in this case the t-copula has upper and lower tail dependence. Note that even for zero correlation this copula shows tail dependence.

However, it is not typical that normal mixtures exhibit tail dependence. This property obviously depends on the distribution of the mixing variable, and if the mixing distribution does not have power tails, the property of asymptotical tail independence is inherited from the normal distribution.

5 Simulating from copulas

This section provides the essential steps necessary to implement Monte Carlo simulations of pseudo rvs which have a certain copula C. Of course, also ordinary rvs can be generated with these methods. In contrast to these, pseudo rvs do not aim at mimicking randomness as they are computed from deterministic algorithms. Besides increasing the precision of the simulations the outcome from the simulations become reproducible.

We always give an algorithm which allows to simulate one single vector $\mathbf{U} = (U_1, \ldots, U_d)$ which has uniform marginals and the desired copula. Arbitrary marginals can then be obtained using the quantile transformation from equation (2). Repeating the algorithm n times one obtains a sample of n independent identically distributed (iid) multivariate pseudo rvs admitting the desired copulas. For example, the simulations which led to the right part of Figure 4 were achieved in this way.

5.1 Gaussian and t-copula

If the copula is derived from a multivariate distribution, like for the Gaussian and the t-copula, it turns out to be particularly easy to simulate such pseudo rvs. The first step is always to simulate pseudo rvs according to the underlying multivariate distribution and the second step is to transform these results to uniform marginals using (2). This leads to the following algorithms.

Algorithm 5.1. **(Multivariate Gaussian Copula)**

1. For an arbitrary covariance matrix $\tilde{\mathbf{\Sigma}}$ obtain the correlation matrix $\mathbf{\Sigma}$ as outlined in Example 3.1
2. Perform a Cholesky-decomposition[11] $\mathbf{\Sigma} = \mathbf{A}'\mathbf{A}$
3. Generate iid standard normal pseudo rvs $\tilde{X}_1, \ldots, \tilde{X}_d$
4. Compute $(X_1, \ldots, X_d)' = \mathbf{X} = \mathbf{A}\tilde{\mathbf{X}}$ from $\tilde{\mathbf{X}} = (\tilde{X}_1, \ldots, \tilde{X}_d)'$
5. Return $U_i = \Phi(X_i)$, $i = 1, \ldots, d$ where Φ is the standard normal cumulative distribution function

Algorithm 5.2. **(Multivariate t-Copula)**

1. For an arbitrary covariance matrix $\tilde{\mathbf{\Sigma}}$ obtain the correlation matrix $\mathbf{\Sigma}$ as outlined in Example 3.1
2. Generate multivariate normal $\mathbf{X}$ with covariance matrix $\mathbf{\Sigma}$ as above
3. Generate independent $\xi \sim \chi^2_\nu$ for example via $\xi = \sum_{i=1}^{\nu} Y_i^2$, where Y_i are iid $\mathcal{N}(0, 1)$
4. Return $U_i = t_\nu(X_i/\sqrt{\xi/\nu})$, $i = 1, \ldots, d$ where t_ν is the cumulative distribution function of a univariate t-distribution with ν degrees of freedom

5.2 Archimedean copulas

The considered copulas – Gumbel, Clayton and Frank copulas – fall into the class of so-called *Laplace transform Archimedean copulas* (or LT-Archimedean copulas). For this class, the inverse of the generator ϕ has a nice representation as a Laplace transform of some function G. The simulation algorithm uses that such pseudo rvs may be generated easily.

To consider this approach in more detail, consider a cumulative distribution function G and denote its Laplace transform by

$$\hat{G}(t) := \int_0^\infty e^{-tx} dG(x), \quad t \geq 0.$$

We set $\hat{G}(\infty) := 0$ and realize that $\hat{G}$ is a continuous and strictly decreasing function, thus may serve well as a candidate for ϕ^{-1}. Indeed, generate a pseudo rv V with cdf G and iid standard uniform pseudo rvs $X_1, \ldots, X_d$ (also independent of V). Set

$$U_i := \hat{G}\Big(-\frac{\ln X_i}{V}\Big),$$

then the vector $\mathbf{U}$ has the desired Archimedean copula dependence structure with generator $\phi = \hat{G}^{-1}$. A proof is given in McNeil, Frey, and Embrechts (2005), Proposition 5.46.

Thus, we have the following algorithm

[11] The Cholesky decomposition is implemented in a lot of packages. Alternatively, you may consider Press et al (1992).

Algorithm 5.3. **(Laplace transform Archimedean copula)**

1. Generate a pseudo rv V with cdf G
   a) For a Clayton copula, V is gamma distributed, $\mathrm{Ga}(1/\theta, 1)$, and $\hat{G}(t) = (1+t)^{-1/\theta}$
   b) For a Gumbel copula V is stable distributed, $\mathrm{St}(1/\theta, 1, \gamma, 0)$ with $\gamma = (\cos(\pi/2/\theta))^\theta$ and $\hat{G}(t) = \exp(-t^{1/\theta})$
   c) For a Frank copula, V is discrete with $P(V = k) = (1 - e^{-\theta})^k/(k\theta)$ for $k = 1, 2, \ldots$
2. Generate iid uniform pseudo rvs $X_1, \ldots, X_d$
3. Return $U_i = \hat{G}\Big(-\frac{\ln X_i}{V}\Big), i = 1, \ldots, d$

Note that for this algorithm it is necessary to simulate gamma, stable or discrete distributions. We refer to McNeil, Frey, and Embrechts (2005) or Schoutens (2003) for a detailed outline how this may be done.

5.3 Marshall-Olkin copula

The Marshall-Olkin copula stems from a concrete model assumption which can easily be used to simulate pseudo rvs. It may be recalled that rvs X_1 and X_2 with a Marshall-Olkin copula are obtained from the independent, exponentially distributed rvs Z_1, Z_2 and Z_{12} with intensities λ_1, λ_2 and λ_{12}, respectively, by

$$X_i = \min\{Z_i, Z_{12}\}, \quad i = 1, 2.$$

The marginals of the X_i are easily computed and one obtains that

$$F_i(x) := P(X_i \le x) = 1 - \exp\big(-(\lambda_i + \lambda_{12})x\big).$$

Finally, we obtain uniform distributed rvs letting $U_i = F_i(X_i)$ and U_1, U_2 have the desired copula. We summarize the steps in the following algorithm.

Algorithm 5.4. **(Marshall-Olkin copula)**

1. Generate independent pseudo rvs $Z_i \sim \mathrm{Exp}(\lambda_i)$, $i = 1, \ldots, 3$
2. Return $U_i = 1 - \exp\big(-(\lambda_i + \lambda_{12})\max\{Z_i, Z_3\}\big)$, $i = 1, 2$

If the parametrization via α_1, α_2 as in (13) is used, one chooses for some λ_{12} the parameters λ_1 and λ_2, such that

$$\alpha_i = \frac{\lambda_{12}}{\lambda_i + \lambda_3}, i = 1, 2.$$

6 Conclusion and a word of caution

The literature on copulas is growing fast. For a more detailed exposition of copulas with different applications in view we refer to the excellent textbook McNeil, Frey, and Embrechts (2005) as well as to the article Embrechts, Lindskog, and McNeil (2003). The book of Cherubini, Luciano, and Vecchiato (2004) gives additional examples but offers less theoretical detail. Using extreme value theory and copulas, the book of Malevergne and Sornette (2006) analyses extreme financial risks. For an in-depth study of copulas the book of Nelsen (1999) is the standard reference, which only deals with two-dimensional copulas as this is sufficient for mathematical purposes[12]. Interesting remarks of the history and the development of copulas may be found in Fisher (1995).

[12] The construction of multivariate copulas from bivariate ones is covered in Section 3.5 in Nelsen (1999).

A word of caution is due at this place. Copulas are a very general tool to describe dependence structures, and have been successfully applied in many cases. However, the immense generality is also the drawback of copulas. For a discussion we refer to Mikosch (2006) and the commentary of Genest and Rémillard (2006). Some points mentioned therein shall be emphasized here.

- In general, it is quite difficult to estimate copulas from data. Of course, this stems in particular from the generality of copulas. In many multivariate distributions it is quite well understood how to estimate the dependence structure, so a sensible statistical estimation of a copula will consider a particular parametrization, i.e. stemming from a multivariate family of distributions, and estimation of the parameters will typically be much easier.
- As we have seen, some copulas stem from multivariate distributions, as for example the Gaussian and the t-copula. Using that the copula marginals and the dependence structure can be disentangled, the dependence structure can be handled independently from the marginals. Obviously the Gaussian copula can be applied to t-marginals and vice versa. This results in new multivariate distributions with different behaviour. This or similar procedures may often be seen in applications of copulas. However, the use of this procedure is questionable and the outcomes should be handled with care.
- On the other side, various copulas which do not stem from multivariate distributions, like Archimedean copulas for example, have their form mainly because of mathematical tractability. Therefore their applicability as natural models for dependence should be verified in each case.
- Finally, copulas refer to a static concept of dependence, while many applications, especially in Finance, refer to time series and a dynamic concept of dependence is needed.

Despite of this critics, it is remarkable that the use of copulas has greatly improved the modelling of dependencies in practice. For example, in contrast to linear correlation, the use of copulas avoids typical pitfalls and therefore leads to a mathematically consistent modelling of dependence. Also the variety of possible dependence structures applied increased significantly.

It is obvious that a structural modelling of the dependency, for example via factors, is always preferable to a simple application of an arbitrarily chosen copula. When the time for a deeper analysis is short, copulas may still serve as a quick fix for approximating the underlying relations.

Needless to say, despite of these cautious words the application of copulas has been a great success to a number of fields, and especially in finance they are a frequently used concept. The complicated interaction observed, e.g. in credit risk, has lead to a number of approaches where copulas are used to model the dependence structure and calibrations to market data seem to be quite successful. Furthermore, they serve as an excellent tool for stress testing portfolios or other products in finance and insurance as they allow to interpolate between extreme cases of dependence. This in turn makes the copula an effective element of risk management.

This chapter may serve as a short introduction to the concept of copulas. Needless to say that many important topics have not been covered and the interested reader may find additional material from the bibliography and the following chapters of this book.

References

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

Embrechts, P., F. Lindskog, and A. McNeil (2003). Modelling dependence with copulas and applications to risk management. In S. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance*, pp. 331–385. Elsevier.

Malevergne, Y. and D. Sornette (2006). *Extreme Financial Risks*. Springer Verlag. Berlin Heidelberg New York.

McNeil, A., R. Frey, and P. Embrechts (2005). *Quantitative Risk Management: Concepts, Techniques and Tools*. Princeton University Press.

Nelsen, R. B. (1999). *An introduction to copulas*, Volume 139 of *Lecture Notes in Statistics*. Springer Verlag. Berlin Heidelberg New York.

Schoutens, W. (2003). *Lévy Processes in Finance*. Wiley Series in Probability and Statistics. John Wiley & Sons, England.

Sklar, A. (1959). Fonctions de répartition à n dimensions e leurs marges. *Publications de l'Institut de Statistique de l'Univiversité de Paris 8*, 229 – 231.

Name	Copula	Paramter range
Independence or product copula	$\Pi(\mathbf{u}) = \prod_{i=1}^{d} u_i$	
Comonotonicity copula or Fréchet-Hoeffding upper bound	$M(\mathbf{u}) = \min\{u_1, \ldots, u_d\}$	
Countermonotonicity copula or Fréchet-Hoeffding lower bound	$W(u_1, u_2) = \max\{u_1 + u_2 - 1, 0\}$	
Gaussian copula	$C_{\mathbf{\Sigma}}^{Ga}(\mathbf{u}) = \Phi_{\mathbf{\Sigma}}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d))$	
t- or Student copula	$C_{\nu,\mathbf{\Sigma}}^{t}(\mathbf{u}) = t_{\nu,\mathbf{\Sigma}}(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d))$	
Gumbel copula or Gumbel-Hougaard copula	$C_\theta^{Gu}(u_1, u_2) = \exp\Big[-\big((-\ln u_1)^\theta + (-\ln u_2)^\theta\big)^{\frac{1}{\theta}}\Big]$	$\theta \in [1, \infty)$
Clayton copula	$C_\theta^{Cl}(u_1, u_2) = \Big(\max\{u_1^{-\theta} + u_2^{-\theta} - 1, 0\}\Big)^{-\frac{1}{\theta}}$	$\theta \in [-1, \infty) \setminus \{0\}$
Generalized Clayton copula	$C_{\theta,\delta}^{Cl}(u_1, u_2) = \Big([(u_1^{-\theta} - 1)^\delta + (u_2^{-\theta} - 1)^\delta]^{\frac{1}{\delta}} + 1\Big)^{-\frac{1}{\theta}}$	$\theta > 0, \delta \geq 1$
Frank copula	$C_\theta^{Fr}(u_1, u_2) = -\frac{1}{\theta}\ln\left(1 + \frac{(e^{-\theta u_1} - 1)\cdot(e^{-\theta u_2} - 1)}{e^{-\theta} - 1}\right)$	$\theta \in \mathbb{R} \setminus \{0\}$
Marshall-Olkin copula or generalized Cuadras-Augé copula	$C_{\alpha_1,\alpha_2}(u_1, u_2) = \min\{u_2 \cdot u_1^{1-\alpha_1}, u_1 \cdot u_2^{1-\alpha_2}\}$	$\alpha_1, \alpha_2 \in [0, 1]$

Table 1: List of Copulas. Note that for the Gumbel, Clayton, Frank and the Marshall-Olkin copula the bivariate versions are stated.