IEOR E4602: Quantitative Risk Management Spring 2016

An Introduction to Copulas

These notes provide an introduction to modeling with copulas. Copulas are the mechanism which allows us to isolate the dependency structure in a multivariate distribution. In particular, we can construct any multivariate distribution by separately specifying the marginal distributions and the copula. Copula modeling has played an important role in finance in recent years and has been a source of controversy and debate both during and since the financial crisis of 2008 / 2009.

In these notes we discuss the main results including Sklar's Theorem and the Fréchet-Hoeffding bounds, and give several different examples of copulas including the Gaussian and t copulas. We discuss various measures of dependency including rank correlations and coefficient of tail dependence and also discuss various fallacies associated with the commonly used Pearson correlation coefficient. After discussing various methods for calibrating copulas we end with an application where we use the Gaussian copula model to price a simple stylized version of a *collateralized debt obligation* or CDO. CDO's were credited with playing a large role in the financial crisis – hence the infamy of the Gaussian copula model.

1 Introduction and Main Results

Copulas are functions that enable us to separate the marginal distributions from the dependency structure of a given multivariate distribution. They are useful for several reasons. First, they help to expose and understand the various fallacies associated with correlation. They also play an important role in pricing securities that depend on many underlying securities, e.g. equity basket options, collateralized debt obligations (CDO's), n^{th}-to-default options etc. Indeed the (in)famous Gaussian copula model was the model[1] of choice for pricing and hedging CDO's up to and even beyond the financial crisis.

There are some problems associated with the use of copulas, however. They are not always applied properly and are generally static in nature. Moreover, they are sometimes used in a "black-box" fashion and understanding the overall joint multivariate distribution can be difficult when it is constructed by separately specifying the marginals and copula. Nevertheless, an understanding of copulas is important in risk management. We begin with the definition of a copula.

Definition 1 *A d-dimensional copula, $C : [0,1]^d \; :\rightarrow \; [0,1]$ is a cumulative distribution function (CDF) with uniform marginals.*

We write $C(\mathbf{u}) = C(u_1, \ldots, u_d)$ for a generic copula and immediately have (why?) the following properties:

1. $C(u_1, \ldots, u_d)$ is non-decreasing in each component, u_i.
2. The i^{th} marginal distribution is obtained by setting $u_j = 1$ for $j \neq i$ and since it it is uniformly distributed
$$C(1, \ldots, 1, u_i, 1, \ldots, 1) \; = \; u_i.$$
3. For $a_i \leq b_i$, $P(U_1 \in [a_1, b_1], \ldots, U_d \in [a_d, b_d])$ must be non-negative. This implies the rectangle inequality
$$\sum_{i_1=1}^{2} \cdots \sum_{i_d=1}^{2} (-1)^{i_1+\cdots+i_d} \, C(u_{1,i_1}, \ldots, u_{d,i_d}) \; \geq \; 0$$
where $u_{j,1} = a_j$ and $u_{j,2} = b_j$.

[1] We will introduce the Gaussian copula model for pricing CDO's in Section 5 and we will return to it again later in the course when we discuss the important topic of model risk.

The reverse is also true in that any function that satisfies properties 1 to 3 is a copula. It is also easy to confirm that $C(1, u_1, \ldots, u_{d-1})$ is a $(d-1)$-dimensional copula and, more generally, that all k-dimensional marginals with $2 \leq k \leq d$ are copulas. We now recall the definition of the quantile function or generalized inverse: for a CDF, F, the generalized inverse, $F^{\leftarrow}$, is defined as

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

We then have the following well-known result:

Proposition 1 *If $U \sim U[0,1]$ and F_X is a CDF, then*

$$P\left(F^{\leftarrow}(U) \leq x\right) \;=\; F_X(x).$$

In the opposite direction, if X has a continuous CDF, F_X, then

$$F_X(X) \;\sim\; U[0,1].$$

Now let $\mathbf{X} = (X_1, \ldots, X_d)$ be a multivariate random vector with CDF F_X and with continuous and increasing marginals. Then by Proposition 1 it follows that the joint distribution of $F_{X_1}(X_1), \ldots, F_{X_d}(X_d)$ is a copula, C_X say. We can find an expression for C_X by noting that

$$\begin{aligned} C_X(u_1, \ldots, u_d) &= P\left(F_{X_1}(X_1) \leq u_1, \ldots, F_{X_d}(X_d) \leq u_d\right) \\ &= P\left(X_1 \leq F_{X_1}^{-1}(u_1), \ldots, X_d \leq F_{X_d}^{-1}(u_d)\right) \\ &= F_{\mathbf{X}}\left(F_{X_1}^{-1}(u_1), \ldots, F_{X_d}^{-1}(u_d)\right). \end{aligned} \tag{1}$$

If we now let $u_j := F_{X_j}(x_j)$ then (1) yields

$$F_{\mathbf{X}}\left(x_1, \ldots, x_d\right) = C_X\left(F_{X_1}(x_1), \ldots, F_{X_d}(x_d)\right)$$

This is one side of the famous Sklar's Theorem which we now state formally.

Theorem 2 (Sklar's Theorem 1959) *Consider a d-dimensional CDF, F, with marginals $F_1, \ldots, F_d$. Then there exists a copula, C, such that*

$$F(x_1, \ldots, x_d) \;=\; C\left(F_1(x_1), \ldots, F_d(x_d)\right) \tag{2}$$

for all $x_i \in [-\infty, \infty]$ and $i = 1, \ldots, d$.

If F_i is continuous for all $i = 1, \ldots, d$, then C is unique; otherwise C is uniquely determined only on Ran$(F_1) \times \cdots \times$ Ran(F_d) where Ran(F_i) denotes the range of the CDF, F_i.

In the opposite direction, consider a copula, C, and univariate CDF's, $F_1, \ldots, F_d$. Then F as defined in (2) is a multivariate CDF with marginals $F_1, \ldots, F_d$. □

Example 1 Let Y and Z be two IID random variables each with CDF, $F(\cdot)$. Let $X_1 \;:=\; \min(Y, Z)$ and $X_2 \;:=\; \max(Y, Z)$ with marginals $F_1(\cdot)$ and $F_2(\cdot)$, respectively. We then have

$$P(X_1 \leq x_1,\; X_2 \leq x_2) \;=\; 2\, F\left(\min\{x_1, x_2\}\right)\; F(x_2) \;-\; F\left(\min\{x_1, x_2\}\right)^2. \tag{3}$$

We can derive (3) by considering separately the two cases (i) $x_2 \leq x_1$ and (ii) $x_2 > x_1$.

We would like to compute the copula, $C(u_1, u_2)$, of (X_1, X_2). Towards this end we first note the two marginals satisfy

$$\begin{aligned} F_1(x) &= 2F(x) \;-\; F(x)^2 \\ F_2(x) &= F(x)^2. \end{aligned}$$

But Sklar's Theorem states that $C(\cdot, \cdot)$ satisfies $C(F_1(x_1),\; F_2(x_2)) \;=\; F(x_1, x_2)$ so if we connect the pieces(!) we will obtain

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

∎

When the Marginals Are Continuous

Suppose the marginal distributions, $F_1, \ldots, F_n$, are continuous. It can then be shown that

$$F_i\left(F_i^{\leftarrow}(y)\right) \;=\; y. \tag{4}$$

If we now evaluate (2) at $x_i = F_i^{\leftarrow}(u_i)$ and use (4) then we obtain the very useful characterization

$$C(\mathbf{u}) \;=\; F\left(F_1^{\leftarrow}(u_1), \ldots, F_d^{\leftarrow}(u_d)\right). \tag{5}$$

The Density and Conditional Distribution of a Copula

If the copula has a density, i.e. a PDF, then it is obtained in the usual manner as

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

When $d = 2$, we can plot $c(\mathbf{u})$ to gain some intuition regarding the copula. We will see such plots later in Section 2. If the copula has a density and is given in the form of (5) then we can write

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

where we have used $F_i^{\leftarrow} = F_i^{-1}$ since F_i is differentiable. It is also easy to obtain the conditional distribution of a copula. In particular, we have

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

This implies that the conditional CDF may be derived directly from the copula itself. The following result is one of the most important in the theory of copulas. It essentially states that if we apply a monotonic transformation to each component in $\mathbf{X} = (X_1, \ldots, X_d)$ then the copula of the resulting multivariate distribution remains the same.

Proposition 3 (Invariance Under Monotonic Transformations) *Suppose the random variables $X_1, \ldots, X_d$ have continuous marginals and copula, C_X. Let $T_i : \mathbb{R} \to \mathbb{R}$, for $i = 1, \ldots, d$ be strictly increasing functions. Then the dependence structure of the random variables*

$$Y_1 := T_1(X_1), \; \ldots, \; Y_d := T_d(X_d)$$

is also given by the copula C_X.

Sketch of proof when T_j's are continuous and $F_{X_j}^{-1}$'s exist:
First note that

$$\begin{aligned} F_Y(y_1, \ldots, y_d) \;&=\; P(T_1(X_1) \le y_1, \; \ldots \;, T_d(X_d) \le y_d) \\ &=\; P(X_1 \le T_1^{-1}(y_1), \; \ldots \;, X_d \le T_d^{-1}(y_d)) \\ &=\; F_X(T_1^{-1}(y_1), \; \ldots \;, T_d^{-1}(y_d)) \end{aligned} \tag{6}$$

so that (why?) $F_{Y_j}(y_j) = F_{X_j}(T_j^{-1}(y_j))$. This in turn implies

$$F_{Y_j}^{-1}(y_j) = T_j(F_{X_j}^{-1}(y_j)). \tag{7}$$

The proof now follows because

$$
\begin{aligned}
C_Y(u_1,\ldots,u_d) &= F_Y\left(F_{Y_1}^{-1}(u_1),\ \ldots\ ,F_{Y_d}^{-1}(u_d)\right) \text{ by (5)} \\
&= F_X\left(T_1^{-1}\left(F_{Y_1}^{-1}(u_1)\right),\ \ldots\ ,T_d^{-1}\left(F_{Y_d}^{-1}(u_d)\right)\right) \text{ by (6)} \\
&= F_X\left(F_{X_1}^{-1}(u_1),\ \ldots\ ,F_{X_d}^{-1}(u_d)\right) \text{ by (7)} \\
&= C_X(u_1,\ldots,u_d)
\end{aligned}
$$

and so $C_X = C_Y$. □

The following important result was derived independently by Fréchet and Hoeffding.

Theorem 4 (The Fréchet-Hoeffding Bounds) *Consider a copula $C(\mathbf{u}) = C(u_1,\ldots,u_d)$. Then*

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

Sketch of Proof: The first inequality follows from the observation

$$
\begin{aligned}
C(\mathbf{u}) &= P\left(\bigcap_{1\le i\le d}\{U_i \le u_i\}\right) \\
&= 1\ -\ P\left(\bigcup_{1\le i\le d}\{U_i > u_i\}\right) \\
&\ge 1\ - \sum_{i=1}^{d} P(U_i > u_i)\ =\ 1-d+\sum_{i=1}^{d} u_i.
\end{aligned}
$$

The second inequality follows since $\bigcap_{1\le i\le d}\{U_i \le u_i\}\ \subseteq \{U_i \le u_i\}$ for all i. □

The upper Fréchet-Hoeffding bound is tight for all d whereas the lower Fréchet-Hoeffding bound is tight only when $d = 2$. These bounds correspond to cases of extreme of dependency, i.e. comonotonicity and countermonotonicity. The **comonotonic copula** is given by

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

which is the Fréchet-Hoeffding upper bound. It corresponds to the case of *extreme positive dependence*. We have the following result.

Proposition 5 *Let $X_1,\ldots,X_d$ be random variables with continuous marginals and suppose $X_i = T_i(X_1)$ for $i = 2,\ldots,d$ where $T_2,\ldots,T_d$ are strictly increasing transformations. Then $X_1,\ldots,X_d$ have the comonotonic copula.*

Proof: Apply the *invariance under monotonic transformations* proposition and observe that the copula of $(X_1, X_1,\ldots,X_1)$ is the comonotonic copula. □

The **countermonotonic** copula is the 2-dimensional copula that is the Fréchet-Hoeffding lower bound. It satisfies

$$W(u_1,u_2)\ =\ \max\{u_1+u_2-1,\ 0\} \tag{8}$$

and corresponds to the case of *perfect negative dependence.*

Exercise 1 *Confirm that (8) is the joint distribution of $(U, 1-U)$ where $U \sim U(0,1)$. Why is the Fréchet-Hoeffding lower bound not a copula for $d > 2$?*

The **independence** copula satisfies

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

and it is easy to confirm using Sklar's Theorem that random variables are independent if and only if their copula is the independence copula.

---

2 The Gaussian, t and Other Parametric Copulas

We now discuss a number of other copulas, all of which are parametric. We begin with the very important Gaussian and t copulas.

2.1 The Gaussian Copula

Recall that when the marginal CDF's are continuous we have from (5) that $C(\mathbf{u}) \;=\; F\left(F_1^{\leftarrow}(u_1), \ldots, F_1^{\leftarrow}(u_d)\right)$. Now let $\mathbf{X} \sim \mathsf{MN}_d(\mathbf{0}, \mathbf{P})$, where P is the correlation matrix of $\mathbf{X}$. Then the corresponding Gaussian copula is defined as

$$C_P^{Gauss}(\mathbf{u}) \;:=\; \mathbf{\Phi}_{\boldsymbol{P}}\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\right) \tag{9}$$

where $\Phi(\cdot)$ is the standard univariate normal CDF and $\mathbf{\Phi}_{\boldsymbol{P}}(\cdot)$ denotes the joint CDF of $\mathbf{X}$.

Exercise 2 *Suppose* $\mathbf{Y} \sim MN_d(\mu, \Sigma)$ *with* $\mathrm{Corr}(\mathbf{Y}) = \mathbf{P}$. *Explain why* $\mathbf{Y}$ *has the same copula as* $\mathbf{X}$. *We can therefore conclude that a Gaussian copula is fully specified by a correlation matrix,* $\mathbf{P}$.

For $d = 2$, we obtain the countermonotonic, independence and comonotonic copulas in (9) when $\rho = -1,\ 0$, and 1, respectively. We can easily simulate the Gaussian copula via the following algorithm:

Simulating the Gaussian Copula

1. For an arbitrary covariance matrix, $\mathbf{\Sigma}$, let $\mathbf{P}$ be its corresponding correlation matrix.
2. Compute the Cholesky decomposition, $\mathbf{A}$, of $\mathbf{P}$ so that $\mathbf{P} = \mathbf{A}^T\mathbf{A}$.
3. Generate $\mathbf{Z} \sim \mathsf{MN}_d(\mathbf{0}, \mathbf{I}_d)$.
4. Set $\mathbf{X} = \mathbf{A}^T\,\mathbf{Z}$.
5. Return $\mathbf{U} = (\Phi(X_1), \ldots, \Phi(X_d))$.

The distribution of $\mathbf{U}$ is the Gaussian copula $C_P^{Gauss}(\mathbf{u})$ so that

$$\mathsf{Prob}(U_1 \le u_1, \ldots, U_d \le u_d) \;=\; \mathbf{\Phi}\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)\right)$$

where $\mathbf{u} = (u_1, \ldots, u_d)$. This is also (why?) the copula of $\mathbf{X}$.

We can use this algorithm to generate any random vector $\mathbf{Y} = (Y_1, \ldots, Y_d)$ with arbitrary marginals and Gaussian copula. To do this we first generate $\mathbf{U}$ using the above algorithm and then use the standard inverse transform method along each component to generate $Y_i = F_i^{-1}(U_i)$. That $\mathbf{Y}$ has the corresponding Gaussian copula follows by our result on the invariance of copulas to monotonic transformations.

2.2 The t Copula

Recall that $\mathbf{X} = (X_1, \ldots, X_d)$ has a multivariate t distribution with ν degrees of freedom (d.o.f.) if

$$\mathbf{X} \;=\; \frac{\mathbf{Z}}{\sqrt{\xi/\nu}}$$

where $\mathbf{Z} \sim \mathsf{MN}_d(\mathbf{0}, \mathbf{\Sigma})$ and $\xi \sim \chi^2_\nu$ independently of $\mathbf{Z}$. The d-dimensional t-copula is then defined as

$$C^t_{\nu,P}(\mathbf{u}) \;:=\; \mathbf{t}_{\nu,\mathbf{P}}\left(t_\nu^{-1}(u_1), \ldots, t_\nu^{-1}(u_d)\right) \tag{10}$$

where again $\mathbf{P}$ is a correlation matrix, $\mathbf{t}_{\nu,\mathbf{P}}$ is the joint CDF of $\mathbf{X} \sim \mathbf{t}_d(\nu, \mathbf{0}, \mathbf{P})$ and t_ν is the standard univariate CDF of a t-distribution with ν d.o.f. As with the Gaussian copula, we can easily simulate from the t copula:

Simulating the t Copula

1. For an arbitrary covariance matrix, $\mathbf{\Sigma}$, let $\mathbf{P}$ be its corresponding correlation matrix.
2. Generate $\mathbf{X} \sim \mathsf{MN}_d(\mathbf{0}, \mathbf{P})$.
3. Generate $\xi \sim \chi^2_\nu$ independent of $\mathbf{X}$.
4. Return $\mathbf{U} = \left(t_\nu(X_1/\sqrt{\xi/\nu}), \ldots, t_\nu(X_d/\sqrt{\xi/\nu})\right)$ where t_ν is the CDF of a univariate t distribution with ν degrees-of-freedom.

The distribution of $\mathbf{U}$ is the t copula, $C^t_{\nu,P}(\mathbf{u})$, and this is also (why?) the copula of $\mathbf{X}$. As before, we can use this algorithm to generate any random vector $\mathbf{Y} = (Y_1, \ldots, Y_d)$ with arbitrary marginals and a given t copula.We first generate $\mathbf{U}$ with the t copula as described above and then use the standard inverse transform method along each component to generate $Y_i = F_i^{-1}(U_i)$.

2.3 Other Parametric Copulas

The **bivariate Gumbel** copula is defined as

$$C^{Gu}_\theta(u_1, u_2) \;:=\; \exp\left(-\left((-\ln u_1)^\theta \;+\; (-\ln u_2)^\theta\right)^{\frac{1}{\theta}}\right) \tag{11}$$

where $\theta \in [1, \infty)$. When $\theta = 1$ in (11) we obtain the independence copula. When $\theta \to \infty$ the Gumbel copula converges to the comonotonicity copula. The Gumbel copula is an example of a copula with *tail dependence* (see Section 3) in just one corner.

Consider Figure 1 where we compare the bivariate normal and (bivariate) meta-Gumbel[2] distributions. In particular, we obtained the two scatterplots by simulating $5,000$ points from each bivariate distribution. In each case the marginal distributions were standard normal. The linear correlation is $\approx .7$ for both distributions but it should be clear from the scatterplots that the meta-Gumbel is much more likely to see large joint moves. We will return to this issue of tail dependence in Section 3.

In Figure 8.4 of Ruppert and Matteson (*SDAFE* 2015) we have plotted the bivariate Gumbel copula for various values of θ. As such earlier, we see that tail dependence only occurs in one corner of the distribution.

The **bivariate Clayton** copula is defined as

$$C^{Cl}_\theta(u_1, u_2) \;:=\; \left(u_1^{-\theta} \;+\; u_2^{-\theta} - 1\right)^{-1/\theta} \tag{12}$$

where $\theta \in [-1, \infty)\backslash\{0\}$.

[2]The meta-Gumbel distribution is the name of any distribution that has a Gumbel copula. In this example the marginals are standard normal. The "meta" terminology is generally used to denote the copula of a multivariate distribution, leaving the marginals to be otherwise specified.

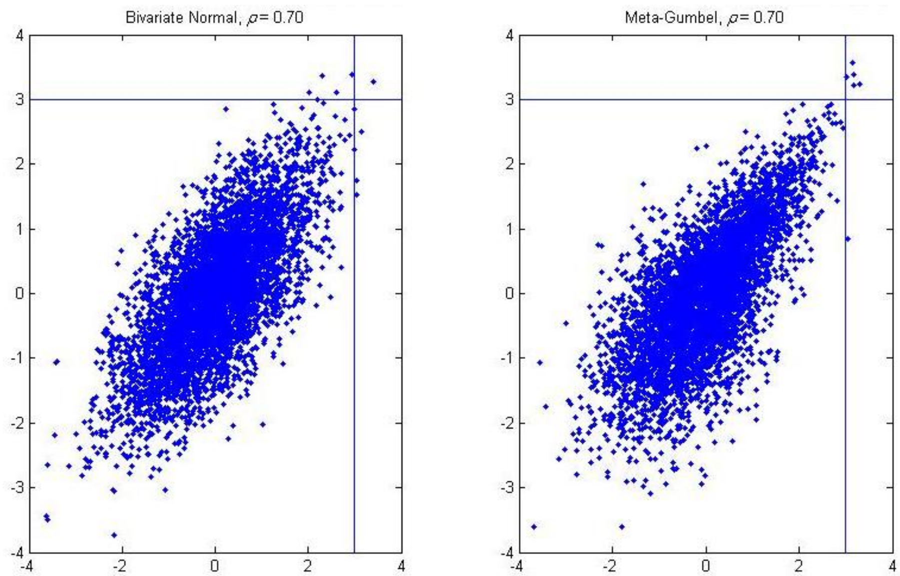


Figure 1: Bivariate Normal versus Bivariate Meta-Gumbel Copula

As $\theta \to 0$ in (12) we obtain the independence copula and as $\theta \to \infty$ it can checked that the Clayton copula converges to the comonotonic copula. When $\theta = -1$ we obtain the Fréchet-Hoeffding lower bound. The Clayton copula therefore traces the countermonotonic, independence and comonotonic copulas as θ moves from -1 through 0 to ∞. In Figure 8.3 of Ruppert and Matteson (*SDAFE*, 2015) we have plotted the bivariate Clayton copula for various values of θ.

The Clayton and Gumbel copulas belong to a more general family of copulas called the **Archimedean** copulas. They can be be generalized to d dimensions but their d-dimensional versions are **exchangeable**. This means that $C(u_1, \ldots, u_d)$ is unchanged if we permute $u_1, \ldots, u_d$ implying that all pairs have the same dependence structure. This has implications for modeling with Archimedean copulas!

There are many other examples of parametric copulas. Some of these are discussed in *MFE* by McNeil, Frey and Embrechts and *SDAFE* by Ruppert and Matteson.

---

3 Measures of Dependence

Understanding the dependence structure of copulas is vital to understanding their properties. There are three principal measures of dependence:

1. The usual **Pearson**, i.e. linear, correlation coefficient is only defined if second moments exist. It is invariant under positive linear transformations, but not under general strictly increasing transformations. Moreover, there are many fallacies associated with the Pearson correlation.

2. **Rank correlations** only depend on the unique copula of the joint distribution and are therefore (why?) invariant to strictly increasing transformations. Rank correlations can also be very useful for calibrating copulas to data. See Section 4.3.

3. **Coefficients of tail dependence** are a measure of dependence in the extremes of the distributions.

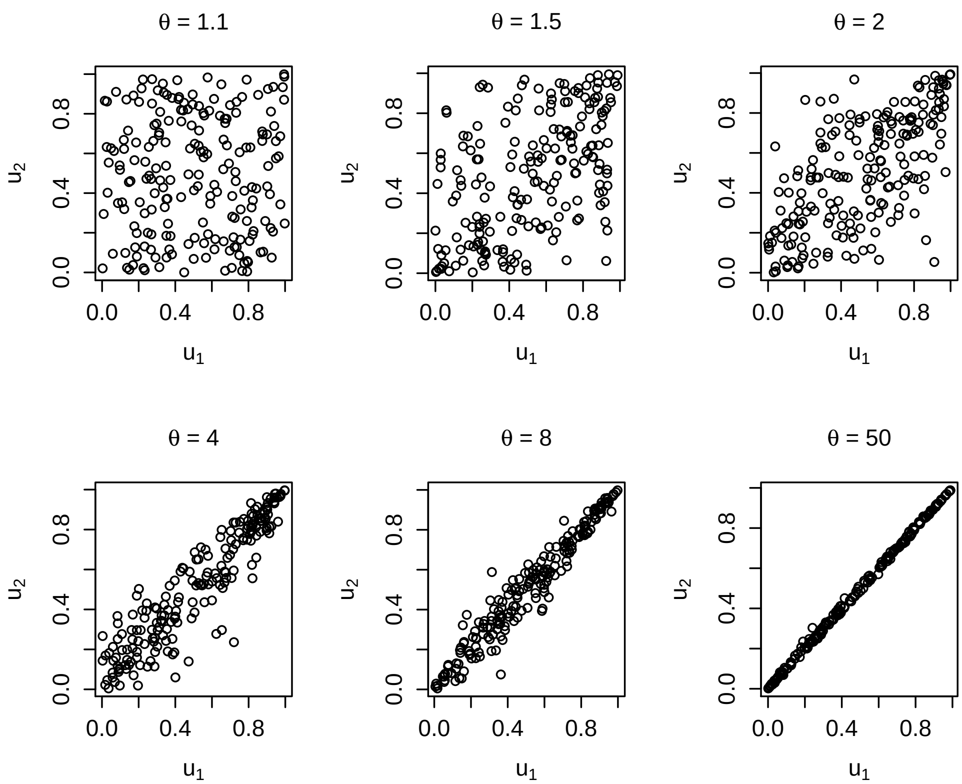


Figure 8.4 from Ruppert and Matteson: Bivariate random samples of size 200 from various Gumbel copulas.

Fallacies of the Pearson Correlation Coefficient

We begin by discussing some of the most common fallacies associated with the Pearson correlation coefficient. In particular, each of the following statements is false!

1. The marginal distributions and correlation matrix are enough to determine the joint distribution.
2. For given univariate distributions, F_1 and F_2, and any correlation value $\rho \in [-1, 1]$, it is always possible to construct a joint distribution F with margins F_1 and F_2 and correlation ρ.
3. The VaR of the sum of two risks is largest when the two risks have maximal correlation.

We have already provided (where?) a counter-example to the first statement earlier in these notes. We will focus on the second statement / fallacy but first we recall the following definition.

Definition 2 *We say two random variables, X_1 and X_2, are* of the same type *if there exist constants $a > 0$ and $b \in \mathbb{R}$ such that*

$$X_1 \sim aX_2 + b.$$

The following result addresses the second fallacy.

Theorem 6 *Let (X_1, X_2) be a random vector with finite-variance marginal CDF's F_1 and F_2, respectively, and an unspecified joint CDF. Assuming $\text{Var}(X_1) > 0$ and $\text{Var}(X_2) > 0$, then the following statements hold:*

1. *The attainable correlations form a closed interval $[\rho_{min}, \rho_{max}]$ with $\rho_{min} < 0 < \rho_{max}$.*

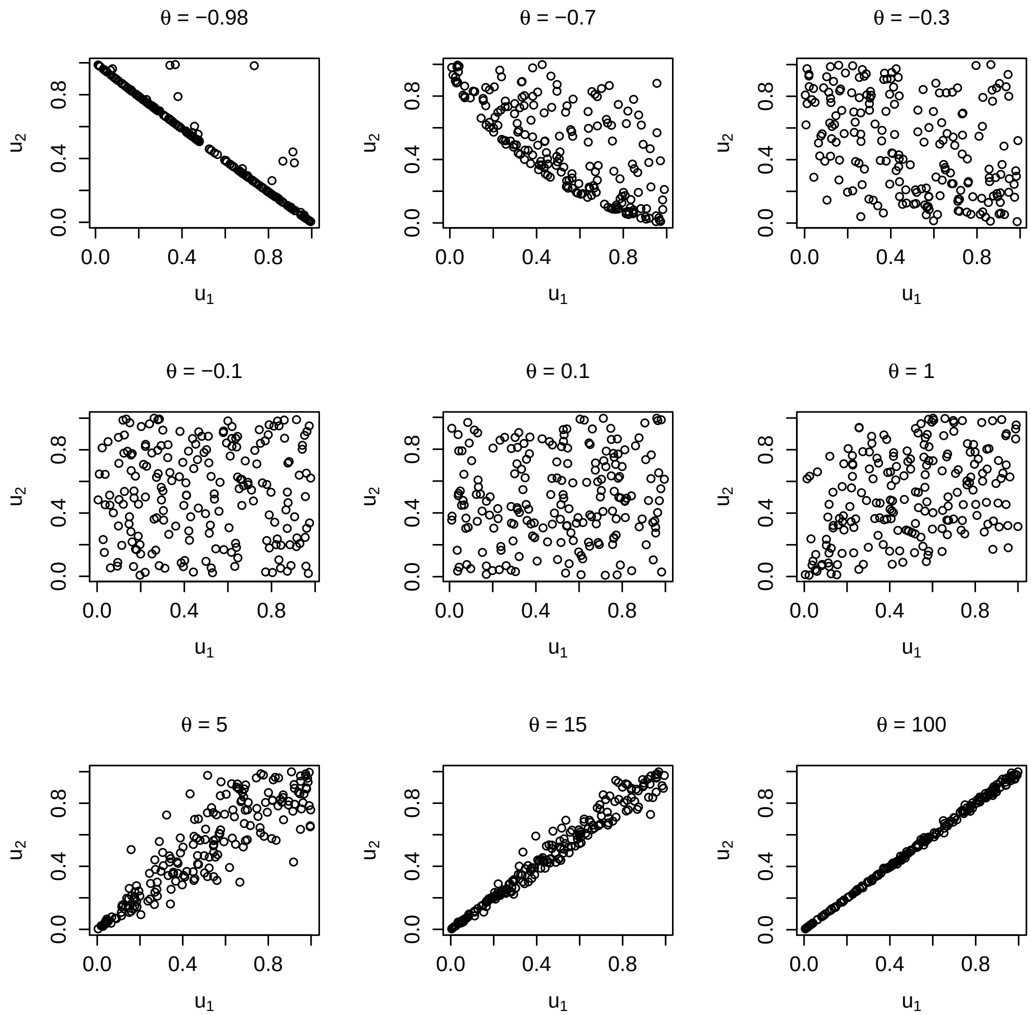


Figure 8.3 from Ruppert and Matteson: Bivariate random samples of size 200 from various Clayton copulas.

2. *The minimum correlation $\rho = \rho_{min}$ is attained if and only if X_1 and X_2 are countermonotonic. The maximum correlation $\rho = \rho_{max}$ is attained if and only if X_1 and X_2 are comonotonic.*

3. *$\rho_{min} = -1$ if and only if X_1 and $-X_2$ are of the same type. $\rho_{max} = 1$ if and only if X_1 and X_2 are of the same type.*

Proof: The proof is not very difficult; see *MFE* for details. □

3.1 Rank Correlations

There are two important rank correlation measures, namely Spearman's rho and Kendall's tau. We begin with the former.

Definition 3 *For random variables X_1 and X_2,* **Spearman's rho** *is defined as*

$$\rho_s(X_1, X_2) \ := \ \rho(F_1(X_1), F_2(X_2)).$$

Spearman's rho is therefore simply the linear correlation of the probability-transformed random variables. The Spearman's rho *matrix* is simply the matrix of pairwise Spearman's rho correlations, $\rho(F_i(X_i), F_j(X_j))$. It is (why?) a positive-definite matrix. If X_1 and X_2 have continuous marginals then it can be shown that

$$\rho_s(X_1, X_2) \;=\; 12 \int_0^1 \int_0^1 \left(C(u_1, u_2) \;-\; u_1 u_2\right)\, du_1\, du_2.$$

It is also possible to show that for a bivariate Gaussian copula

$$\rho_s(X_1, X_2) \;=\frac{6}{\pi}\;\arcsin\frac{\rho}{2} \;\simeq\; \rho$$

where ρ is the Pearson, i.e., linear, correlation coefficient.

Definition 4 *For random variables X_1 and X_2,* **Kendall's tau** *is defined as*

$$\rho_\tau(X_1, X_2) \;:=\; \mathrm{E}\left[\mathit{sign}\left((X_1 - \tilde{X}_1)\,(X_2 - \tilde{X}_2)\right)\right]$$

where $(\tilde{X}_1, \tilde{X}_2)$ is independent of (X_1, X_2) but has the same joint distribution as (X_1, X_2).

Note that Kendall's tau can be written as

$$\rho_\tau(X_1, X_2) \;=\; P\left((X_1 - \tilde{X}_1)\,(X_2 - \tilde{X}_2)\; > 0\right) \;-\; P\left((X_1 - \tilde{X}_1)\,(X_2 - \tilde{X}_2)\; < 0\right)$$

so if both probabilities are equal then $\rho_\tau(X_1, X_2) = 0$. If X_1 and X_2 have continuous marginals then it can be shown that

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

It may also be shown that for a bivariate Gaussian copula, or more generally, if $\mathbf{X} \sim \mathbf{E}_2(\mu, \mathbf{P}, \psi)$ and $P(\mathbf{X} = \mu) = 0$ then

$$\rho_\tau(X_1, X_2) \;=\frac{2}{\pi}\;\arcsin\rho \tag{13}$$

where $\rho = \mathbf{P}_{12} = \mathbf{P}_{21}$ is the Pearson correlation coefficient. We note that (13) can be very useful for estimating ρ with fat-tailed elliptical distributions. It generally provides much more robust estimates of ρ than the usual Pearson estimator. This can be seen from Figure 2 where we compare estimates of Pearson's correlation using the usual Pearson estimator versus using Kendall's τ and (13). The true underlying distribution was bivariate t with $\nu = 3$ degrees-of-freedom and true Pearson correlation $\rho = 0.5$. Each point in Figure 2 was an estimate constructed from a sample of $n = 60$ (simulated) data-points from the true distribution. We see the estimator based on Kendall's tau is clearly superior to the usual estimator.

Exercise 3 *Can you explain why Kendall's tau performs so well here?* Hint*: It may be easier to figure out why the usual Pearson estimator can sometimes perform so poorly.*

Properties of Spearman's Rho and Kendall's Tau

Spearman's rho and Kendall's tau are examples of rank correlations in that, when the marginals are continuous, they depend only on the bivariate copula and not on the marginals. They are therefore invariant in this case under strictly increasing transformations.

Spearman's rho and Kendall's tau both take values in $[-1, 1]$:

- They equal 0 for independent random variables. (It's possible, however, for dependent variables to also have a rank correlation of 0.)
- They take the value 1 when X_1 and X_2 are comonotonic.
- They take the value -1 when X_1 and X_2 are countermonotonic.

As discussed in Section 4.3, Spearman's rho and Kendall's tau can be very useful for calibrating copulas via method-of-moments type algorithms. And as we have seen, the second fallacy associated with the Pearson coefficient is clearly no longer an issue when we work with rank correlations.

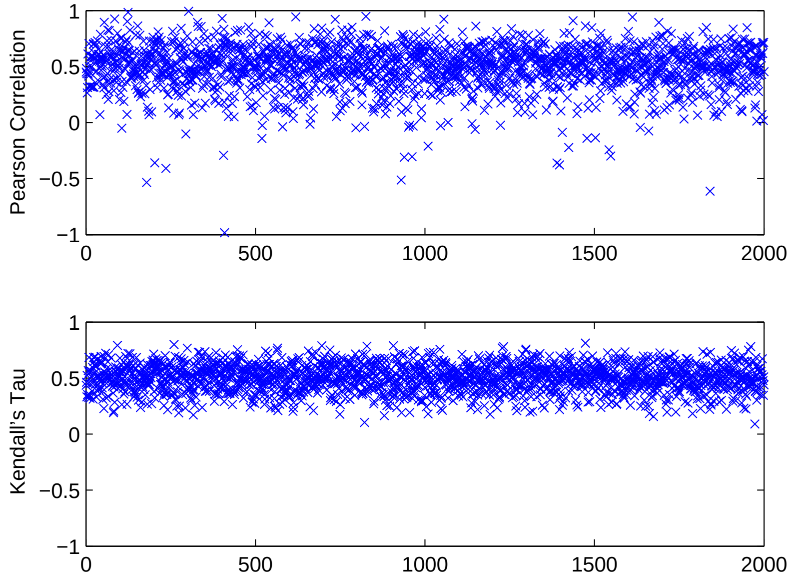


Figure 2: Estimating Pearson's correlation using the usual Pearson estimator versus using Kendall's τ. Underlying distribution was bivariate t with $\nu = 3$ degrees-of-freedom and true Pearson correlation $\rho = 0.5$.

3.2 Tail Dependence

We have the following definitions of lower and upper tail dependence.

Definition 5 *Let X_1 and X_2 denote two random variables with CDF's F_1 and F_2, respectively. Then the* **coefficient of upper tail dependence**, *λ_u, is given by*

$$\lambda_u := \lim_{q \nearrow 1} P\left(X_2 > F_2^{\leftarrow}(q) \mid X_1 > F_1^{\leftarrow}(q)\right)$$

provided that the limit exists. Similarly, the **coefficient of lower tail dependence**, *λ_l, is given by*

$$\lambda_l := \lim_{q \searrow 0} P\left(X_2 \leq F_2^{\leftarrow}(q) \mid X_1 \leq F_1^{\leftarrow}(q)\right)$$

provided again that the limit exists.

If $\lambda_u > 0$, then we say that X_1 and X_2 have *upper tail dependence* while if $\lambda_u = 0$ we say they are *asymptotically independent in the upper tail. Lower tail dependence* and *asymptotically independent in the lower tail* are similarly defined using λ_l. The upper and lower coefficients are identical for both the Gaussian and t copulas. Unless $\rho = \pm 1$ it can be shown that the Gaussian copula is asymptotically independent in both tails. This is not true of the t copula, however. In Figure 8.6 of Ruppert and Matteson (*SDAFE*, 2015) we have plotted $\lambda_l = \lambda_u$ as a function of ρ for various values of the d.o.f, ν. It is clear (and makes intuitive sense) that $\lambda_l = \lambda_u$ is decreasing in ν for any fixed value of ρ.

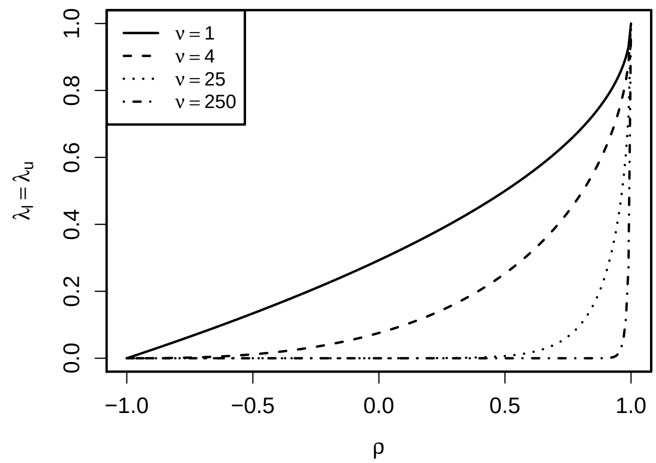


Figure 8.6 from Ruppert and Matteson: Coefficients of tail dependence for bivariate t-copulas as functions of ρ for $\nu = 1$, 4, 25 and 250.

---

4 Estimating Copulas

There are several related methods that can be used for estimating copulas:

1. Maximum likelihood estimation (MLE). It is often considered too difficult to apply as there too many parameters to estimate.

2. **Pseudo**-MLE of which there are two types: *parametric* pseudo-MLE and *semi-parametric* pseudo-MLE. Pseudo-MLE seems to be used most often in practice. The marginals are estimated via their *empirical CDFs* and the copula is then estimated via MLE.

3. **Moment-matching** methods are also sometimes used. They methods can also be used for finding starting points for (pseudo)-MLE.

We now discuss these methods in further detail.

4.1 Maximum Likelihood Estimation

Let $\mathbf{Y} = (Y_1 \dots Y_d)^\top$ be a random vector and suppose we have parametric models $F_{Y_1}(\cdot \mid \boldsymbol{\theta}_1), \dots, F_{Y_d}(\cdot \mid \boldsymbol{\theta}_d)$ for the marginal CDFs. We assume that we also have a parametric model $c_{\mathbf{Y}}(\cdot \mid \boldsymbol{\theta}_C)$ for the copula density of $\mathbf{Y}$. By differentiating (2) we see that the density of $\mathbf{Y}$ is given by

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{Y}}(y_1, \dots, y_d) = c_{\mathbf{Y}}\left(F_{Y_1}(y_1), \dots, F_{Y_d}(y_d)\right) \prod_{j=1}^{d} f_{Y_j}(y_j). \tag{14}$$

Suppose now that we are given an IID sample $\mathbf{Y}_{1:n} = (\mathbf{Y}_1, \ldots, \mathbf{Y}_n)$. We then obtain the log-likelihood as

$$\begin{aligned}
\log L(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d, \boldsymbol{\theta}_C) &= \log \prod_{i=1}^{n} f_{\mathbf{Y}}(\mathbf{y}_i) \\
&= \sum_{i=1}^{n} \Big(\log[c_{\mathbf{Y}}\left(F_{Y_1}(y_{i,1} \,|\, \boldsymbol{\theta}_1), \ldots, F_{Y_d}(y_{i,d} \,|\, \boldsymbol{\theta}_d) \,|\, \boldsymbol{\theta}_C\right)] \\
&\quad + \log\left(f_{Y_1}(y_{i,1} \,|\, \boldsymbol{\theta}_1)\right) + \cdots + \log\left(f_{Y_d}(y_{i,d} \,|\, \boldsymbol{\theta}_d)\right) \Big). \qquad (15)
\end{aligned}$$

The ML estimators $\widehat{\boldsymbol{\theta}}_1, \ldots, \widehat{\boldsymbol{\theta}}_d, \widehat{\boldsymbol{\theta}}_C$ are obtained by maximizing (15) with respect to $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_d, \boldsymbol{\theta}_C$. There are problems with this approach, however:

1. There are (tto) many parameters to estimate, especially for large values of d. As a result, performing the optimization can be difficult.
2. If any of the parametric univariate distributions $F_{Y_i}(\cdot \,|\, \boldsymbol{\theta}_i)$ are misspecified then this can cause biases in the estimation of both the univariate distributions and the copula.

4.2 Pseudo-Maximum Likelihood Estimation

The pseudo-MLE approach helps to resolve the problems associated with the MLE approach mentioned above. It has two steps:

1. First estimate the marginal CDFs to obtain $\widehat{F}_{Y_j}$ for $j = 1, \ldots, d$. We can do this using either:
   - The *empirical CDF* of $y_{1,j}, \ldots, y_{n,j}$ so that
   $$\widehat{F}_{Y_j}(y) = \frac{\sum_{i=1}^{n} 1_{\{y_{i,j} \leq y\}}}{n+1}.$$
   - A parametric model with $\widehat{\boldsymbol{\theta}}_j$ obtained using the usual univariate MLE approach.
2. Then estimate the copula parameters $\boldsymbol{\theta}_C$ by maximizing
$$\sum_{i=1}^{n} \log \left[c_{\mathbf{Y}} \left(\widehat{F}_{Y_1}(y_{i,1}), \ldots, \widehat{F}_{Y_d}(y_{i,d}) \,|\, \boldsymbol{\theta}_C \right) \right] \qquad (16)$$

Note that (16) is obtained directly from (15) by only including terms that depend on the (as yet) unestimated parameter vector $\boldsymbol{\theta}_C$ and setting the marginals at their fitted values obtained from Step 1. It is worth mentioning that even (16) may be difficult to maximize if d large. In that event it is important to have a good starting point for the optimization or to impose *additional structure* on $\boldsymbol{\theta}_C$.

4.3 Fitting Gaussian and t Copulas

A moment-matching approach for fitting either the Gaussian or t copulas can be obtained immediately from the following result.

Proposition 7 (Results 8.1 from Ruppert and Matteson)
Let $\mathbf{Y} = (Y_1 \ldots Y_d)^\top$ have a meta-Gaussian distribution with continuous univariate marginal distributions and copula $C_{\boldsymbol{\Omega}}^{Gauss}$ and let $\Omega_{i,j} = [\boldsymbol{\Omega}]_{i,j}$. Then

$$\rho_\tau(Y_i, Y_j) = \frac{2}{\pi} \arcsin\left(\Omega_{i,j}\right) \qquad (17)$$

and

$$\rho_S(Y_i, Y_j) = \frac{6}{\pi} \arcsin\left(\Omega_{i,j}/2\right) \approx \Omega_{i,j} \qquad (18)$$

If instead $\mathbf{Y}$ has a meta-t distribution with continuous univariate marginal distributions and copula $C_{\nu,\boldsymbol{\Omega}}^{t}$ then (17) still holds but (18) does not.

Exercise 4 *There are several ways to use this Proposition to fit meta Gaussian and t copulas. What are some of them?*

---

5 An Application: Pricing CDO's

We begin in subsection 5.1 with a simple[3] one-period stylized copula where each of the (corporate) bonds has identical characteristics. Then in subsection 5.2 we consider the more realistic case of multi-period CDO's with heterogeneous underlying securities.

5.1 A Simple Stylized 1-Period CDO

We want to find the expected losses in a simple 1-period CDO with the following characteristics:

- The maturity is 1 year.
- There are $N = 125$ bonds in the reference portfolio.
- Each bond pays a coupon of one unit after 1 year if it has not defaulted.
- The recovery rate on each defaulted bond is zero.
- There are 3 *tranches* of interest: the equity, mezzanine and senior tranches with attachment points 0-3 defaults, 4-6 defaults and 7-125 defaults, respectively.

We make the simple assumption that the probability, q, of defaulting within 1 year is identical across all bonds X_i is the normalized asset value of the i^{th} credit, i.e. bond, and we assume

$$X_i = \sqrt{\rho} M + \sqrt{1-\rho}\, Z_i \tag{19}$$

where $M, Z_1, \ldots, Z_N$ are IID normal random variables. Note that the correlation between each pair of asset values is identical. We assume also that the i^{th} credit defaults if $X_i \leq \bar{x}_i$. Since the probability of default, q, is identical across all bonds we must therefore have

$$\bar{x}_1 = \cdots \bar{x}_N = \Phi^{-1}(q). \tag{20}$$

It now follows from (19) and (20) that

$$\begin{aligned} \mathsf{P}(\text{Credit } i \text{ defaults} \,|\, M) &= \mathsf{P}(X_i \leq \bar{x}_i \,|\, M) \\ &= \mathsf{P}(\sqrt{\rho} M + \sqrt{1-\rho}\, Z_i \leq \Phi^{-1}(q) \,|\, M) \\ &= \mathsf{P}\left(Z_i \leq \frac{\Phi^{-1}(q) - \sqrt{\rho} M}{\sqrt{1-\rho}} \,\middle|\, M \right). \end{aligned}$$

Therefore conditional on M, the total number of defaults is Binomial(N, q_M) where

$$q_M := \Phi\left(\frac{\Phi^{-1}(q) - \sqrt{\rho} M}{\sqrt{1-\rho}} \right).$$

That is,

$$p(k \,|\, M) = \binom{N}{k} q_M^k (1 - q_M)^{N-k}. \tag{21}$$

The unconditional probabilities can be computed by integrating numerically the binomial probabilities with respect to M so that

$$\mathsf{P}(k \text{ defaults}) = \int_{-\infty}^{\infty} p(k \,|\, M) \phi(M)\, dM \tag{22}$$

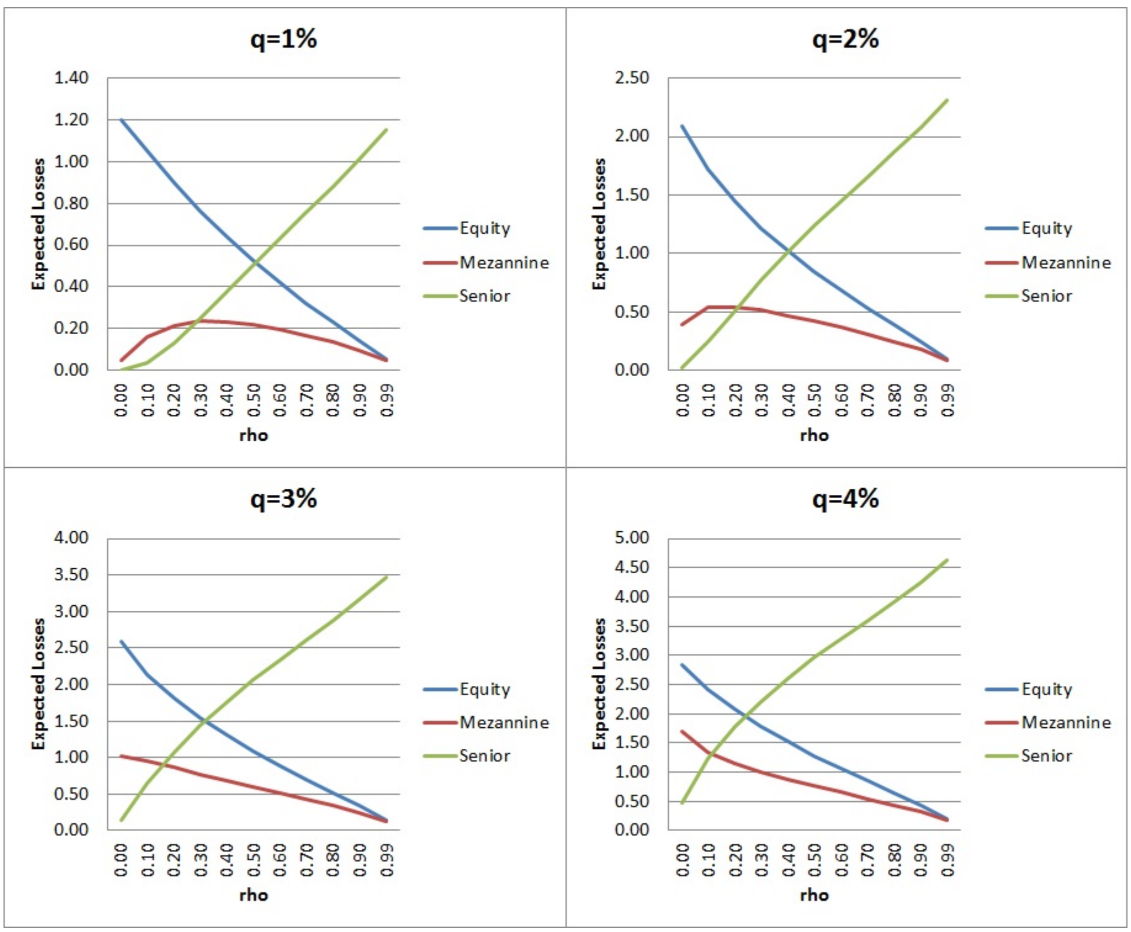
[3]The example is taken from "The Devil is in the Tails: Actuarial Mathematics and the Subprime Mortgage Crisis", by C. Donnelly and P. Embrechts in ASTIN Bulletin 40(1), 1-33.

where $\phi(\cdot)$ is the standard normal PDF. We can now compute the expected (risk-neutral) loss on each of the three tranches according to

$$\begin{aligned} \mathrm{E}_0^Q\,[\text{Equity tranche loss}] &= 3 \times \mathrm{P}(3 \text{ or more defaults}) + \sum_{k=1}^{2} k\,\mathrm{P}(k \text{ defaults}) \\ \mathrm{E}_0^Q\,[\text{Mezzanine tranche loss}] &= 3 \times \mathrm{P}(6 \text{ or more defaults}) + \sum_{k=1}^{2} k\,\mathrm{P}(k+3 \text{ defaults}) \\ \mathrm{E}_0^Q\,[\text{Senior tranche loss}] &= \sum_{k=1}^{119} k\,\mathrm{P}(k+6 \text{ defaults}). \end{aligned}$$

Results for various values of ρ and q are displayed in the figure below. Regardless of the individual default

Figure 3: **Expected Tranche Losses As a Function of q and ρ**



probability, q, and correlation, ρ, we see

$$\mathrm{E}_0^Q\,[\%\text{ Equity tranche loss}] \geq \mathrm{E}_0^Q\,[\%\text{ Mezzanine tranche loss}] \geq \mathrm{E}_0^Q\,[\%\text{ Senior tranche loss}]\,.$$

We also note that the expected losses in the equity tranche are always decreasing in ρ while mezzanine tranches are often relatively insensitive[4] to ρ. The expected losses in senior tranches (with upper attachment point of 100% or 125 units in our example) are always increasing in ρ.

[4]This has important implications when it comes to model calibration, an issue we will not pursue further here.

Exercise 5 *How does the total expected loss in the portfolio vary with ρ?*

Where Does the Gaussian Copula Appear?

Let $C_{X_1,\ldots,X_N}$ denote the copula of $\mathbf{X} := (X_1, \ldots, X_N)$ in (19). Then it should be clear that the copula of $\mathbf{X}$ is indeed the Gaussian copula C_P^{Gauss} where P is the correlation matrix with all off-diagonal elements equal to ρ. By definition of the Gaussian copula, $C_{X_1,\ldots,X_N}$ satisfies

$$C_{X_1,\ldots,X_N}(u_1,\ldots,u_N) = P(\Phi(X_1) \le u_1, \ldots, \Phi(X_N) \le u_N). \tag{23}$$

We can substitute for the X_i's in (23) using (19) and then condition on M to obtain

$$C_{X_1,\ldots,X_N}(u_1,\ldots,u_N) = \int_{-\infty}^{\infty} \prod_{i=1}^{N} \Phi\left(\frac{\Phi^{-1}(u_i) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right) \phi(M)\, dM \tag{24}$$

where $\phi(\cdot)$ is the standard Normal pdf. We can now use Sklar's Theorem (in particular (2)) and (24) to see that the joint probability of default satisfies

$$\begin{aligned} P(X_1 \le \bar{x}_1, \ldots, X_N \le \bar{x}_n) &= \int_{-\infty}^{\infty} \prod_{i=1}^{N} \Phi\left(\frac{\Phi^{-1}(\Phi(\bar{x}_i)) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right) \phi(M)\, dM \\ &= \int_{-\infty}^{\infty} \prod_{i=1}^{N} \Phi\left(\frac{\Phi^{-1}(q) - \sqrt{\rho}M}{\sqrt{1-\rho}}\right) \phi(M)\, dM. \end{aligned} \tag{25}$$

In this example, we assumed the default probabilities were identical and as a result we were able to use binomial probabilities as in (21). We therefore did not need to use (25) to compute the expected tranche losses and so the role of the Gaussian copula was a little obscured in this example. That will not be the case in Section 5.2 when the underlying bonds have different default probabilities.

Remark 1 *It should be clear from (21), (22) and (25) that the introduction of the $\bar{x}_i$'s was not necessary at all. They were merely used to allow for the economic interpretation of the X_i's as representing (normalized) asset values with default occurring if X_i fell below $\bar{x}_i$. But from a mathematical point of view this was not necessary and only the default probabilities, q, were required in conjunction with Gaussian copula assumption. We will therefore ignore the $\bar{x}_i$'s in Section 5.2.*

5.2 Multiperiod CDO's

In practice CDO's are multi-period securities and the underlying bonds are heterogeneous. More work is therefore required to explain their mechanics and analyze them. We assume that each of the N *credits* in the *reference portfolio* has a *notional amount* of A_i. This means that if the i^{th} credit defaults, then the portfolio incurs a loss of $A_i \times (1 - R_i)$ where R_i is the *recovery rate*, i.e., the percentage of the notional amount that is recovered upon default. We assume R_i is fixed and known. We also assume that the default of the i^{th} credit occurs according to an exponential[5] distribution with a constant arrival rate λ_i. Note that λ_i is easily estimated from either *credit-default-swap* (CDS) spreads or the prices of corporate bonds, all of which are observable in the market place. In particular, for any fixed time, t, we can compute $F_i(t)$, the risk-neutral probability that the i^{th} credit defaults before time t.

For $i = 1, \ldots, N$, we define

$$X_i \;=\; a_i M \;+\; \sqrt{1 - a_i^2}\, Z_i \tag{26}$$

where $M, Z_1, \ldots, Z_N$ are IID normal random variables. Each of the *factor loadings*, a_i, is assumed to lie in the interval $[0, 1]$. It is also clear that $\mathrm{Corr}(X_i, X_j) = a_i a_j$ and that the X_i's are multivariate normally distributed

[5]The exponential rate is commonly assumed but it is not an important assumption. What is important is that we can compute $F_i(t)$ for each credit.

with covariance matrix equal to the correlation matrix, $\mathbf{P}$, where $\mathbf{P}_{i,j} = a_i a_j$ for $i \neq j$. Let $F(t_1, \ldots, t_N)$ denote the joint distribution for the default times of the N credits in the portfolio. We then assume

$$F(t_1, \ldots, t_N) \;\; = \;\; \mathbf{\Phi}_P \left(\Phi^{-1}(F_1(t_1)), \ldots, \Phi^{-1}(F_n(t_N)) \right) \tag{27}$$

where $\mathbf{\Phi}_P(\cdot)$ denotes the multivariate normal CDF with mean vector $\mathbf{0}$ and correlation matrix, $\mathbf{P}$. We note that (27) amounts to assuming that the default times for the N credits have the same Gaussian copula as $\mathbf{X} = (X_1, \ldots, X_N)$. It is worth mentioning that this particular model is sometimes called the 1-*factor* gaussian copula model with the random variable, M, playing the role of the single factor. It is easy to generalize this to a multi-factor model by including additional factors in (26). The effect of these additional factors would be to allow for more general correlation matrices, P, in (27) but this is only achieved at the cost of having to compute multi-dimensional integrals in (28) below.

Computing the Portfolio Loss Distribution

In order to price credit derivatives with the Gaussian copula model of (27), we need to compute the portfolio loss distribution for a common fixed time $t = t_1 = \cdots = t_N$. Towards this end we let $q_i := F_i(t)$ denote the marginal (risk-neutral) probability of the i^{th} credit defaulting by this time. Noting (as before) that the default events for each of the N names are independent conditional on M, we obtain

$$F_i(t|M) \;=\; \Phi\left(\frac{\Phi^{-1}(q_i) - a_i M}{\sqrt{1 - a_i^2}} \right).$$

Now let $p^N(l, t)$ denote the risk-neutral probability that there are a total of l defaults in the portfolio before time t. Then we may write

$$p^N(l, t) \;=\; \int_{-\infty}^{\infty} p^N(l, t|M) \; \phi(M) \; dM. \tag{28}$$

While we will not go into the details[6], it is straightforward to calculate $p^N(l, t|M)$ using a simple iterative procedure. We can then perform a numerical integration on the right-hand-side of (28) to calculate $p(l, t)$. If we assume that the notional, A_i, and the recovery rate, R_i, are constant[7] across all credits, then the loss on any given credit will be either 0 or $A(1 - R)$. In particular, this implies that knowing the probability distribution of the number of defaults is equivalent to knowing the probability distribution of the total loss up to time t in the reference portfolio.

The Mechanics and Pricing of a Synthetic CDO Tranche

We now provide a more formal definition of a *tranche*. A tranche is defined by the *lower* and *upper attachment points*, L and U, respectively. Since the A_i's and R_i's are assumed constant across all i, the *tranche loss function*, $TL^{L,U}(l)$, for a fixed time, t, is a function of only the number of defaults, l, and is given by

$$TL_t^{L,U}(l) \;:=\; \max\left\{ \min\{lA(1 - R), \; U\} - L, \; 0 \right\}.$$

For a given number of defaults it tells us the loss suffered by the tranche. For example, if L and U are 3% and 7%, respectively, and the total portfolio loss is $lA(1 - R) = 5\%$, then the tranche loss is 2% of the total portfolio notional or 50% of the tranche notional.

When an investors *sells protection* on the tranche she is guaranteeing to reimburse any realized losses on the tranche to the *protection buyer*. In return for this guarantee the protection seller is paid a *premium* at regular intervals[8] until the contract expires. The *fair value* of the CDO tranche is defined to be that value of the premium for which the expected value of the *premium leg* is equal to the expected value of the *default leg*.

[6] Can you see how the iterative procedure might work?

[7] This is a common assumption although it is straightforward to relax it.

[8] Typically every three months. In some cases the protection buyer may also pay an upfront amount in addition to, or instead of, a regular premium. This typically occurs for *equity* tranches which have a lower attachment point of zero.

Clearly then the fair value of the CDO tranche will depend on the expected value of the tranche loss function. Indeed, for a fixed time, t, the expected tranche loss is given by

$$\mathrm{E}\left[TL_t^{L,U}\right] = \sum_{l=0}^{N} TL_t^{L,U}(l)\; p(l,t)$$

which we can compute using (28). We now compute the fair value of the premium and default legs.

Premium Leg: The premium leg represents the premium payments that are paid periodically by the protection buyer to the protection seller. They are paid at the end of each time interval and they are based upon the *remaining notional* in the tranche. Formally, the time $t = 0$ value of the premium leg, $P_0^{L,U}$, satisfies

$$PL_0^{L,U} = s \sum_{t=1}^{n} d_t \Delta_t \left((U-L) - \mathrm{E}_0\left[TL_t^{L,U}\right]\right) \tag{29}$$

where n is the number of periods in the contract, d_t is the risk-free discount factor for payment date t, s is the annualized *spread* or premium paid to the protection seller and Δ_t is the *accrual factor* for payment date t. For example, if we ignore day-count conventions, $\Delta_t = 1/4$ if payments take place quarterly and $\Delta_t = 1/2$ if payments take place semi-annually. Note that (29) is consistent with the statement that the premium paid at any time t is based only on the remaining notional in the tranche.

Default Leg: The default leg represents the cash flows paid to the protection buyer upon losses occurring in the tranche. Formally, the time $t = 0$ value of the default leg, $DL_0^{L,U}$, satisfies

$$DL_0^{L,U} = \sum_{t=1}^{n} d_t \left(\mathrm{E}_0\left[TL_t^{L,U}\right] - \mathrm{E}_0\left[TL_{t-1}^{L,U}\right]\right). \tag{30}$$

The fair premium, s^* say, is the value of s that equates the value of the default leg with the value of the premium leg. In particular, we have

$$s^* := \frac{DL_0^{L,U}}{\sum_{t=1}^{n} d_t \Delta_t \left((U-L) - \mathrm{E}_0\left[TL_t^{L,U}\right]\right)}.$$

As is the case with swaps and forwards, the fair value of the tranche to the protection buyer and seller at initiation is therefore zero.

Exercise 6 *Suppose the CDO contract also required the protection buyer to make a fixed payment, F say, to the protection seller at the initiation of the contract in addition to the periodic premium payments. How would you compute the fair premium, s^*, in that case?*

Remark 2 *As mentioned earlier, it is also possible to incorporate recovery values and notional values that vary with each credit in the portfolio. In this case it is straightforward to calculate the characteristic function of the total portfolio loss, again conditional on the factor, M. The fast Fourier transform can then be applied to calculate the conditional portfolio loss distribution. This in turn can be used to compute the unconditional portfolio loss distribution in a manner analogous to (28).*

5.3 Cash Versus Synthetic CDOs

The first CDOs to be traded were all *cash* CDOs where the reference portfolio was an actual physical portfolio and consisted of corporate bonds[9] that the CDO issuer usually kept on its balance sheet. Capital requirement regulations meant that these bonds required a substantial amount of capital to be set aside to cover any potential losses. In an effort to reduce these capital requirements, banks converted the portfolio into a series of tranches and sold most of these tranches to investors. By keeping the equity tranche for themselves the banks

[9]Or *loans* in the case of collateralized loan obligations (CLOs).

succeeded in keeping most of the economic risk of the portfolio and therefore the corresponding rewards. However, they also succeeded in dramatically reducing the amount of capital they needed to set aside. Hence the first CDO deals were motivated by *regulatory arbitrage* considerations.

It soon become clear, however, that there was an appetite in the market-place for these products. Hedge funds, for example, were keen to buy the riskier tranches whereas insurance companies and others sought the AAA-rated senior and super-senior tranches. This appetite and the explosion in the CDS market gave rise to *synthetic tranches*. In a synthetic tranche, the underlying reference portfolio is no longer a physical portfolio of corporate bonds or loans. It is instead a fictitious portfolio consisting of a number of credits with an associated notional amount for each credit. The mechanics of a synthetic tranche are precisely as we described above but they have at least two features that distinguish them from cash CDOs: (i) with a synthetic CDO it is no longer necessary to tranche the entire portfolio and sell the entire "deal". For example, a bank could sell protection on a 3%-7% tranche and never have to worry about selling the other pieces of the reference portfolio. This is not the case with cash CDOs. (ii) because the bank no longer owns the underlying bond portfolio, it is no longer hedged against adverse price movements. It therefore needs to dynamically hedge its synthetic tranche position and typically does so using the CDS markets.

5.4 Calibrating the Gaussian Copula

In practice, it is very common to calibrate synthetic tranches as follows. First we assume that all pairwise correlations, $\mathrm{Corr}(X_i, X_j)$, are identical. This is equivalent to taking $a_1 = \cdots = a_N = a$ in (26). We then obtain $\mathrm{Corr}(X_i, X_j) = a^2 := \rho$ for all i, j. In the case of the liquid CDO tranches whose prices are observable in the market-place, we then choose ρ so that the fair tranche spread in the model is equal to the quoted spread in the market place. We refer to this calibrated correlation, ρ_{imp} say, as the tranche *implied correlation*. If the model is correct, then we should obtain the same value of ρ_{imp} for every tranche. Unfortunately, this does not occur in practice. Indeed, in the case of mezzanine tranches it is possible that there is no value of ρ that fits the market price. It is also possible that there are multiple solutions.

The market has reacted to this problem by introducing the concept of *base correlation*. Base correlations are the implied correlations of equity tranches with increasing upper attachment points. Implied base correlations can always be computed and then bootstrapping techniques are employed to price the mezzanine tranches. Just as equity derivatives markets have an implied volatility surface, the CDO market has *implied base correlation curves*. These functions are generally increasing functions of the upper attachment point.

A distinguishing feature of CDOs and other credit derivatives such as n^{th}-to-default options is that they can be very sensitive to correlation assumptions. As a risk manager or investor in structured credit, it is very important to understand why equity, mezzanine and super senior tranches react as they do to changes in implied correlation.

Exercise 7 *Explain why the value of a long position in an equity tranche has a* positive *exposure to implied correlation. (A long position is equivalent to selling protection on the tranche.)*

Exercise 8 *Explain why the value of a long position in a super-senior tranche, i.e. a tranche with an upper attachment point of 100%, has a* negative *exposure to implied correlation.*

Because a mezzanine tranche falls between the equity and senior tranches its exposure to default correlation is not so clear-cut. Indeed a mezzanine tranche may have little or no exposure to implied correlation. This can have significant implications some of which we discuss below.

5.5 Risk Management and Weaknesses of the Gaussian Copula

The Gaussian copula model has many weaknesses and some of them are particularly severe. When discussing these weaknesses below it should be kept in mind that the Gaussian copula model has also been applied to other asset-backed securities and not just to CDO's. That said, some market participants understood these weaknesses well before the 2008 credit crisis and understood that the Gaussian copula was ultimately little more than a mechanism for quoting prices of CDO tranches via implied correlations. In the final reckoning, it is difficult to attach blame to any model. The blame that is attributed to the Gaussian copula surely lies with those who abused and misused these models and failed to understand the true extent of the market and liquidity risks associated with the CDOs and ABS structures more generally. Of course, it is also clear that in many circumstances there were no incentives in place at the institutional level to encourage users to properly understand and account for these risks.

We now describe some of the weaknesses of the Gaussian copula model (and the structured credit market more generally) as they pertain to the 2008 credit crisis.

1. The Gaussian copula model is completely *static*. There are no stochastic processes in the model and so it does not allow for the possibility that credit spreads, for example, change dynamically. The model therefore fails to account for many possible eventualities.

2. The model is not *transparent*. In particular, it is very difficult to interpret the implied correlation of a tranche. On a related point, the non-transparency of the model makes it very difficult to construct realistic scenarios that can be used to stress-test a portfolio of tranches. For example, a portfolio of ABS and MBS CDOs should be stress tested using macro-economic models and models of the housing market. But there is no clear way to do this using the Gaussian copula technology. As a result most of the stress-tests that were done in practice were based on stressing implied correlation and CDS spreads. But determining whether a stress of implied correlation is realistic in the context of the Gaussian copula model is very difficult.

3. Liquidity risk was clearly a key risk in these markets when so many institutions were holding enormous positions in these securities. Many participants obviously failed to consider this risk. Moreover once the crisis began, model risk contributed to the mounting liquidity problems as the market quickly realized that their models were utterly incapable of pricing these securities. The resulting price uncertainty, coupled with the non-transparency of the products themselves, led to massive bid-offer spreads which essentially shut the market down.

4. Investors in CDO tranches commonly hedged their aggregate credit exposure (as opposed to correlation exposure) by assuming that the credit spreads of the underlying names in the portfolio all moved together. This meant that they could hedge a long position in an equity $0\% - 3\%$ tranche, say, by taking a short position in a $3\% - 7\%$ mezzanine tranche. The notional of the position in the mezzanine tranche can be calculated (how exactly?) using the calibrated Gaussian copula model. While such a hedge did not protect against actual defaults occurring, investors typically viewed themselves as being hedged against movements in the underlying credit spreads. But this hedge can fail to protect you when a subset of individual credit spreads moves dramatically and when the remaining underlying spreads hardly move at all.

   This is what occurred in May 2005 when *Ford* and *General Motors* were suddenly downgraded by the ratings agencies. Their CDS spreads increased substantially and investors that had sold protection on the equity tranches of CDOs containing Ford and GM incurred substantial losses as a result. Moreover, some of these investors had hedged by taking short positions on mezzanine tranches of the same CDO's. However, the spreads on the mezzanine tranches barely moved and so investors who neglected to consider the possibility if idiosyncratic moves were not hedged at all and incurred substantial losses. Of course this is more a weakness of how the model was used rather than a weakness of the model itself.

5. A single parameter, ρ, is used to calibrate a model which has $O(N^2)$ pairwise correlations. This is clearly a very restrictive assumption.

6. In practice the model needed to be calibrated at very frequent intervals which of course only serves to highlight the inadequacy of the model. (This is unfortunately true of most financial models.)

On a related note, when people asserted (in defense of their risk management processes) that the US housing market had never fallen simultaneously across the entire country they were guilty of several mistakes: (i) They ignored the endogeneity in the system so that they did not account for the possibility of the huge structured credit market *exacerbating* a decline in the housing market. (ii) They also ignored that the world we live in today is very different from the world of even 20 years ago. It is far more globalized and interconnected and as a result, historical house price movements should have provided little comfort.

An interesting front-page article was published by the *Wall Street Journal* on Sept 12^{th} 2005 that discussed the role of the Gaussian copula as well as the market reaction to the sudden downgrading of *Ford* and *General Motors* in May 2005. It can be downloaded from

http://www.nowandfutures.com/download/credit_default_swaps_WSJ_news20050912.pdf

The article isn't entirely satisfactory[10] but it certainly proves that many market participants were well aware of the limitations of the Gaussian copula model long before the credit crisis began in 2008. It is also worth emphasizing again that many people were aware of the weaknesses and problems listed above. When they highlighted these problems, however, they were either ignored or discounted. In the absence of strong regulations this is likely to occur again as there will always be a large and powerful group of people who have a vested interest in "business-as-usual".

We will return later to the Gaussian copula model and the pricing of CDO's when we discuss model risk. The Gaussian copula model provides a rich source of examples of models that are poorly understood and used incorrectly!

[10]Even the most sophisticated business media sources seem incapable of getting everything right when it comes to quantitative matters. For those who require further evidence of this statement, do a search of the *Financial Times* and the *Monty Hall* problem.