Copulas

- A copula is the joint distribution of random variables $U_1, U_2, \dots, U_p$, each of which is marginally uniformly distributed as $U(0, 1)$.
- The term copula is also used for the joint cumulative distribution function of such a distribution,

$$C(u_1, u_2, \dots, u_p) = P(U_1 \le u_1, U_2 \le u_2, \dots, U_p \le u_p).$$

Examples

- If $U_1, U_2, \ldots, U_p$ are *independent*,

$$C(u_1, u_2, \ldots, u_p) = u_1 \times u_2 \times \cdots \times u_p.$$

- If they are completely dependent ($U_1 = U_2 = \cdots = U_p$ with probability 1),

$$C(u_1, u_2, \ldots, u_p) = \min(u_1, u_2, \ldots, u_p)$$

Sklar's Theorem

- Copulas are important because of Sklar's Theorem:
  - For any random variables $X_1, X_2, \ldots, X_p$ with joint c.d.f.

    $$F(x_1, x_2, \ldots, x_p) = P\left(X_1 \le x_1, X_2 \le x_2, \ldots, X_p \le x_p\right)$$

    and marginal c.d.f.s

    $$F_j(x) = P\left(X_j \le x\right), \quad j = 1, 2, \ldots, p,$$

    there exists a copula such that

    $$F\left(x_1, x_2, \ldots, x_p\right) = C\left[F_1\left(x_1\right), F_2\left(x_2\right), \ldots, F_p\left(x_p\right)\right].$$

  - If each $F_j(x)$ is continuous, C is unique.

- That is, we can describe the joint distribution of $X_1, X_2, \ldots, X_p$ by the marginal distributions $F_j(x)$ and the copula C.
- The copula (Latin: link) links the marginal distributions together to form the joint distribution.
- From a modeling perspective, Sklar's Theorem allows us to separate the modeling of the marginal distributions $F_j(x)$ from the dependence structure, which is expressed in C.

- The proof is simple in the case that all $F_j(x)$ are continuous, because in this case each has an inverse function $F_j^{-1}(\cdot)$ such that
$$F_j\left[F_j^{-1}(u)\right] = u, \quad \text{for all } 0 \leq u \leq 1.$$
- If $U_j = F_j(X_j)$, then $U_j \sim U(0,1)$:
$$\begin{aligned} P(U_j \leq u) &= P[F_j(X_j) \leq u] = P\left[X_j \leq F_j^{-1}(u)\right] \\ &= F_j\left[F_j^{-1}(u)\right] = u. \end{aligned}$$
- Write C for the c.d.f. of this copula; then
$$\begin{aligned} F(x_1, x_2, \ldots, x_p) &= P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_p \leq x_p) \\ &= P[U_1 \leq F_1(x_1), U_2 \leq F_2(x_2), \ldots, U_p \leq F_p(x_p \\ &= C[F_1(x_1), F_2(x_2), \ldots, F_p(x_p)]. \end{aligned}$$

Copula Density

- When $F(\cdot)$ and $C(\cdot)$ are differentiable, the equation

$$F\left(x_1, x_2, \ldots, x_p\right) = C\left[F_1\left(x_1\right), F_2\left(x_2\right), \ldots, F_p\left(x_p\right)\right]$$

for the joint cumulative distribution function implies that the joint probability density function (pdf) satisfies

$$\frac{f\left(x_1, x_2, \ldots, x_p\right)}{f_1\left(x_1\right) f_2\left(x_2\right) \ldots f_p\left(x_p\right)} = c\left[F_1\left(x_1\right), F_2\left(x_2\right), \ldots, F_p\left(x_p\right)\right]$$

where $c(\cdot)$ is the pdf of the copula distribution:

$$c\left(u_1, u_2, \ldots, u_p\right) = \frac{\partial^p}{\partial u_1 \partial u_2 \ldots \partial u_p} C\left(u_1, u_2, \ldots, u_p\right).$$

- That is, the copula pdf is the ratio of the joint pdf to what it would have been under independence.
- So we can also interpret the copula as the adjustment that we need to make to convert the independence pdf into the joint pdf.

Classes of Copulas

- An Archimedean copula has the form

$$C(u_1, u_2, \ldots, u_p) = \psi\left(\psi^{-1}(u_1) + \psi^{-1}(u_2) + \cdots + \psi^{-1}(u_p)\right)$$

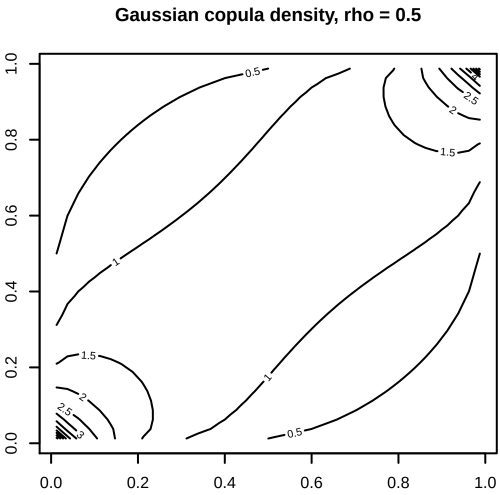
  for an appropriate *generator* $\psi(\cdot)$.
- The Clayton, Frank, Gumbel, and Joe copulas are Archimedean.
- Each has a single parameter that controls the degree of dependence.

- A copula can be extracted from any joint distribution $F(\cdot)$:

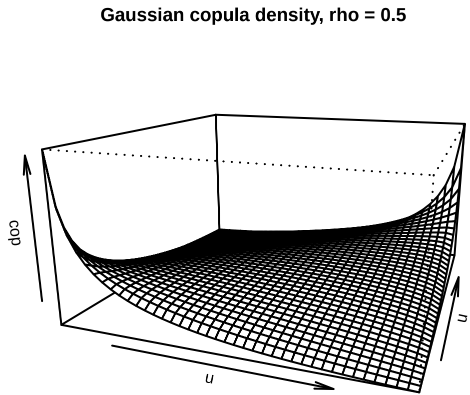
$$C(u_1, u_2, \ldots, u_p) = F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots F_p^{-1}(u_p)\right)$$

  - Unlike the Pearson (or *linear*) correlation, the copula is invariant under monotone increasing transformations of the random variables.
- If $F(\cdot)$ is a multivariate normal distribution $N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $C(\cdot)$ is a *Gaussian* copula.
- If the location or scale of the distribution is changed, the copula does not change, so conventionally $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{R}$, a correlation matrix.

Gaussian copula for $p = 2, \rho = 0.5$

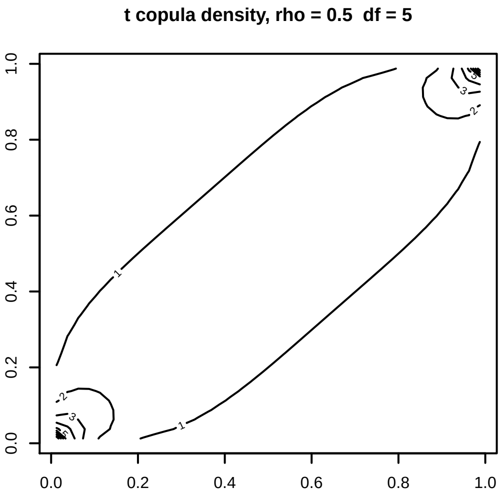


Gaussian copula for $p = 2, \rho = 0.5$

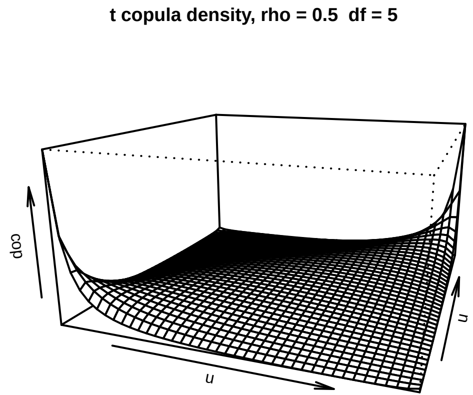


- If $F(\cdot)$ is a multivariate t-distribution $t_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, then $C(\cdot)$ is a t-copula.
- Again, conventionally $\boldsymbol{\mu} = \mathbf{0}$ and $\boldsymbol{\Sigma} = \mathbf{R}$, a correlation matrix.
- But note that when $\mathbf{X} \sim t_\nu(\boldsymbol{\mu}, \boldsymbol{\Sigma})$:
  - if $\nu > 1$ then
$$E(\mathbf{X}) = \boldsymbol{\mu};$$
  - if $\nu > 2$ then
$$\text{Var}(\mathbf{X}) = \left(\frac{\nu}{\nu - 2}\right) \boldsymbol{\Sigma}.$$
- $\mathbf{Y} = \sqrt{\frac{\nu-2}{\nu}}\mathbf{X}$ has the *standardized* multivariate t-distribution.

t-copula for $p = 2, \nu = 5, \rho = 0.5$

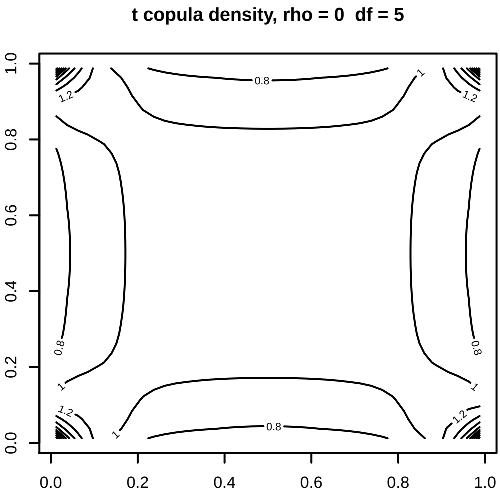


t-copula for $p = 2, \nu = 5, \rho = 0.5$

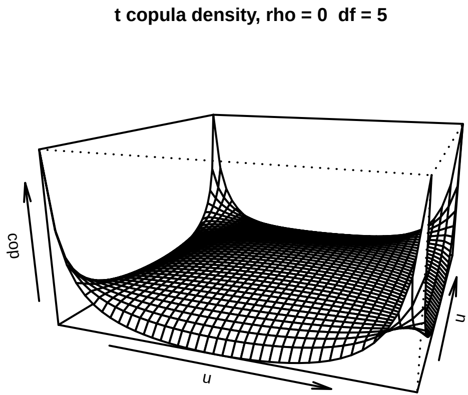


- When $\mathbf{R} = \mathbf{I}$, the multivariate normal distribution is that of independent standard normal variables, and the copula has a constant density.
- But the t-copula still shows dependence.

t-copula for $p = 2, \nu = 5, \rho = 0$



t-copula for $p = 2, \nu = 5, \rho = 0$



Application to Portfolio Risk

- Consider a portfolio of p corporate bonds.
- Ignore interest rate risk, and write the value of bond i at time t when its rating is ρ as $B_i(t, \rho)$.
- If the rating of bond i at time t is $\rho_i(t)$, the portfolio value is

$$V(t) = \sum_i B_i\left[t, \rho_i(t)\right].$$

- The distribution of the change in value from t to $t + \delta t$ can be calculated from the *joint* conditional distribution of

$$\rho_1(t + \delta t), \rho_2(t + \delta t), \ldots, \rho_p(t + \delta t).$$

- Historical data give plausible values for the *marginal* transition probabilities

$$P_{\rho,\rho'} = P\left[\rho_i(t+\delta t) = \rho' | \rho_i(t) = \rho\right].$$

- We can write $\rho_i(t+\delta t)$ as a function of a uniform variable U_i: $\rho_i(t+\delta t) = \rho'$, where ρ' satisfies

$$P\left[\rho_i(t+\delta t) < \rho' | \rho_i(t)\right] < U_i \leq P\left[\rho_i(t+\delta t) \leq \rho' | \rho_i(t)\right].$$

- That is, we break the interval $(0,1)$ into subintervals with endpoints at the cumulative probabilities $P\left[\rho_i(t+\delta t) \leq \rho' | \rho_i(t)\right]$, and find the subinterval containing U_i.

- To model the *joint* transition probabilities, we need a joint distribution for $U_1, U_2, \ldots, U_p$.
- Independence is not plausible: if the bonds were all at the same initial rating, independence implies a multinomial distribution for the counts in the various ratings at a later time, and historical data show considerable over-dispersion relative to the multinomial.
- We therefore need a *copula*; the question is, which one?

The Gaussian Copula

- Modeling and simulation of dependent transitions is usually described in terms of Gaussian random variables instead of uniform random variables.
- Write $Z_i = \Phi^{-1}(U_i)$, so $Z_i \sim N(0, 1)$, and break the whole real line $(-\infty, \infty)$ into subintervals at the Gaussian quantiles corresponding to the cumulative probabilities, $\Phi^{-1}\left\{P\left[\rho_i(t + \delta t) \leq \rho' | \rho_i(t)\right]\right\}$.
- Then Z_i falls into one of the new subintervals iff U_i falls into the corresponding subinterval of $(0, 1)$, and we can determine the new rating as a function of Z_i by finding the subinterval containing Z_i.

- The question of what copula to use now becomes: how to model the joint distribution of $Z_1, Z_2, \ldots, Z_p$, given that each of them is marginally $N(0, 1)$.
- The obvious choice is the multivariate normal distribution $N_p(\mathbf{0}, \mathbf{R})$ with zero mean vector and with dispersion matrix equal to some correlation matrix $\mathbf{R}$.
- This is equivalent to using the Gaussian copula defined by

$$U_i = \Phi\left(Z_i\right), \quad i = 1, 2, \ldots, p,$$

where the Z's have this multivariate normal distribution.

Choice of R

- The Gaussian copula for $p = 2$ has only the single parameter ρ.
- When $p > 2$, the $\frac{1}{2}p(p-1)$ off-diagonal correlations must be specified, constrained by the non-negative definiteness of $\mathbf{R}$.
- When the variables correspond to entities with appropriate structure, a factor model may be used, such as in the Moody's capital model:

$$Z_i = \lambda_1 \zeta_G + \lambda_2 \zeta_{R_i} + \lambda_3 \zeta_{I_i} + \lambda_4 \zeta_{F,i}$$

  where $\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 = 1$, and the ζs are independent standard normal variables.
- So each Z_i is $N(0,1)$, and the correlation of Z_i and Z_j depends on which ζs they share.

- Specifically, for $i \neq j$,

$$\mathrm{Corr}(Z_i, Z_j) = \lambda_1^2 + \lambda_2^2 \mathbb{1}_{\{R_i = R_j\}} + \lambda_3^2 \mathbb{1}_{\{I_i = I_j\}},$$

  where the first term ("global") is present for all i and j, and the second and third are present for entities in the same region R or industry I, respectively.

- The resulting correlation matrix $\mathbf{R}$ is necessarily non-negative definite.

- In the absence of such structure, general parametric models may be used.
- An *exchangeable* correlation matrix has

$$\text{Corr}(Z_i, Z_j) = \rho, \quad i \neq j.$$

- The resulting correlation matrix $\mathbf{R}$ is non-negative definite for $-1/p \leq \rho \leq 1$.
- For $\rho \geq 0$, it is the same as the Moody's model with $\lambda_2 = \lambda_3 = 0$.
- An exchangeable correlation matrix may be a reasonable choice when the dependence of all pairs is expected to be the same.

- Sometimes the variables have a natural ordering, such as by time, and nearby variables are expected to be more dependent than widely separated variables.
- The simplest correlation structure of this form is that of the first order autoregressive model, AR(1):

$$\mathrm{Corr}(Z_i, Z_j) = \phi^{|i-j|}$$

  for some $\phi \in [-1, 1]$.
- Like the exchangeable correlation matrix, the AR(1) correlation matrix is characterized by a single parameter, ϕ.

Tail Dependence

- In risk management we are concerned about the tail of the distribution of losses, and often large losses in a portfolio are caused by simultaneous large moves in several components, so one interesting aspect of any copula is the probability it gives to simultaneous extremes in several dimensions.
- The lower tail dependence of X_i and X_j is defined as

$$\lambda_l = \lim_{u \to 0} P\left[X_i \le F_i^{-1}(u) | X_j \le F_j^{-1}(u)\right].$$

- It depends only on the copula, and equals

$$\lim_{u \to 0} \frac{1}{u} C_{i,j}(u, u).$$

- Tail dependence is symmetric: the tail dependence of X_i and X_j is the same as that of X_j and X_i.
- Upper tail dependence λ_u is defined similarly.
- The Gaussian copula has the curious property that its tail dependence is 0, regardless of the correlation matrix.
- The t-copula is similarly derived from the multivariate t-distribution, and has positive tail dependence even when X_i and X_j are uncorrelated:

$$\lambda_l = \lambda_u = 2t_{\nu+1}\left(-\sqrt{(\nu+1)\left(\frac{1-\rho_{i,j}}{1+\rho_{i,j}}\right)}\right)$$

where $t_{\nu+1}(\cdot)$ is the cumulative distribution function of the univariate t-distribution with $\nu+1$ degrees of freedom.

Table : Tail dependence in t-distributions

	ρ				
	-0.5	0	0.5	0.9	1
$\nu = 2$	0.058	0.182	0.391	0.718	1
4	0.012	0.076	0.253	0.630	1
10	0.000	0.007	0.082	0.463	1
∞	0	0	0	0	1

Tail Correlation

- In some cases, the quantity of interest is the *correlation* of extreme tail events.
- Write

$$I_i = \mathbb{1}_{\{X_i \leq F_i^{-1}(u)\}}.$$

- For instance, I_i could be the indicator of default for the i^{th} bond in a portfolio.
- Then

$$\begin{aligned} E(I_i I_j) &= P(I_i = 1 \cap I_j = 1) \\ &= P\left[X_i \leq F_i^{-1}(u) \cap X_j \leq F_j^{-1}(u)\right] \\ &= C_{i,j}(u, u). \end{aligned}$$

- Then

$$\mathrm{Corr}(I_i, I_j) = \frac{C_{i,j}(u,u) - u^2}{u(1-u)}.$$

- So the *lower tail correlation* of X_i and X_j,

$$\lim_{u \to 0} \mathrm{Corr}(I_i, I_j) = \lim_{u \to 0} \frac{1}{u} C_{i,j}(u,u),$$

  is the same as the lower tail dependence.
- Similarly for the upper tail correlation.

Application to CDO Pricing

- In a CDO (Collateralized Debt Obligation), cash flows from a pool of credit-sensitive instruments are allocated to investors in various *tranches*, in order of seniority.
- In a *cash* CDO, the instruments are bonds, and the investors make initial investments to fund the purchase of the bonds.

- In a *synthetic* CDO, the instruments are CDSs (Credit Default Swaps).
  - A CDS duplicates the cash flows of a loan + bond purchase, but without money or bonds changing hands.
  - Since no assets are actually purchased, initial investment is not required (unfunded synthetic CDO).
  - However, in most cases, investors in the lower tranches are required to invest their notional amounts in the vehicle, since they are at most risk of making payments under the pool of CDSs (partially funded synthetic CDO).

- In either type of CDO, an investor in a particular tranche faces a stream of cash flows, say C_j at times t_j.
- Those cash flows have maximum value when no name in the pool defaults, but may be reduced if defaults have occurred.
- Let τ_i be the time of default of name i. Then C_j depends on which names have defaulted by time t_j:

$$C_j = C_j\left(\mathbb{1}_{\{\tau_1 \le t_j\}}, \mathbb{1}_{\{\tau_2 \le t_j\}}, \dots, \mathbb{1}_{\{\tau_N \le t_j\}}\right).$$

- The value of the tranche is therefore

$$\mathrm{E}\left(\sum_j e^{-rt_j} C_j\right),$$

where the expected value is with respect to the risk-neutral measure, and r is the risk-free interest rate, assumed to be constant.

- The marginal risk-neutral distribution of τ_i can be inferred from the bond market, if this name has issued bonds with the relevant maturities.
- To construct the full joint risk-neutral distribution, we therefore need a copula.

- The market convention is to use the Gaussian copula with the correlation matrix

$$\mathbf{R} = \begin{pmatrix} 1 & \rho & \rho & \dots & \rho \\ \rho & 1 & \rho & \dots & \rho \\ \rho & \rho & 1 & \dots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \dots & 1 \end{pmatrix}.$$

- This is known as a single-factor copula, since the underlying Z's can be constructed as

$$Z_i = \zeta^G \sqrt{\rho} + \zeta_i^{\text{name}} \sqrt{1-\rho},$$

where $\zeta^G \sim N(0,1)$ is common to all Z's, $\zeta_i^{\text{name}} \sim N(0,1)$ is specific to name i, and all ζ's on the right hand side are independent of each other.

- Since the copula involves only a single parameter, ρ, the resulting risk-neutral distribution can be calibrated to the market price of a single tranche.
- A *first-loss* tranche is the first to take losses, up to a *detachment point*, expressed as a percentage of the pool. A correlation calibrated to a first-loss tranche is called a *base* correlation.
- The more senior tranches are characterized by lower and upper detachment (or attachment) points. A correlation calibrated to a non-first-loss tranche is called a *compound* correlation.

Unfortunately:

- Base correlations are found to depend on the detachment point: no one correlation gives prices that match the market price for all detachment points. Instead, they almost always increase with the detachment point.
- Compound correlations may not be unique: for different values of ρ, the expected value matches the market price.
- Compound correlations may not even exist: for no value $-1 < \rho < 1$ does the expected value match the market price.

- Current market practice is to view a non-first-loss tranche as the combination of a long position in the first-loss tranche for the upper detachment point with a short position in the first-loss tranche for the lower detachment (or attachment) point.
- Its value is then the difference of the values for the two first-loss tranches.
- That is, two different risk-neutral distributions are used to find the value of a single tranche!