Amath 546/Econ 589
Copulas

Eric Zivot

Updated: May 21, 2013

Reading

- FRF chapter 1
- QRM chapter 4, sections 5 and 6; chapter 5
- FMUND chapter 6
- SDAFE chapter 8

Introduction

- Capturing co-movement between financial asset returns with linear correlation has been the staple approach in modern finance since the birth of Harry Markowitz's portfolio theory.

- Linear correlation is the appropriate measure of dependence if asset returns follow a multivariate normal (or elliptical) distribution.

- However, the statistical analysis of the distribution of individual asset returns frequently finds fat-tails, skewness and other non-normal features. If the normal distribution is not adequate, then it is not clear how to appropriately measure the dependence between multiple asset returns.

- The theory of copulas provides a flexible methodology for the general modeling of multivariate dependence. The copula function methodology has become the most significant new technique to handle the co-movement between markets and risk factors in a flexible way.

Definitions and Basic Properties of Copulas

Let X be a random variable with distribution function (df) $F_X(x) = \mathsf{Pr}(X \leq x)$. The density function $f_X(x)$ is defined by

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} f_X(z)dz \\ f_X &= F'_X \end{aligned}$$

Let F_X^{-1} denote the quantile function

$$F_X(\alpha) = \mathsf{inf}\{x \mid F_X(x) \geq \alpha\}$$

for $\alpha \in (0, 1)$.

The following are useful results from probability theory:

- $F_X(x) \sim U(0,1)$, where $U(0,1)$ denotes a uniformly distributed random variable on $(0,1)$

- If $U \sim U(0,1)$ then $F_X^{-1}(U) \sim F_X$

The latter result gives a simple way to simulate observations from F_X provided F_X^{-1} is easy to calculate.

Let X and Y be random variables with marginal dfs (margins) F_X and F_Y, respectively, and joint df

$$F_{XY}(x, y) = \Pr(X \leq x, Y \leq y)$$

In general, the marginal dfs may be recovered from the joint df via

$$F_X(x) = F_{XY}(x, \infty),\ F_Y(y) = F_{XY}(\infty, y)$$

The joint density f_{XY} is defined by

$$f_{XY}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{XY}(x, y)$$

The random variables X and Y are independent if

$$F_{XY}(x, y) = F_X(x) F_Y(y)$$

for all values of x and y.

Copulas and Sklar's Theorem

A *bivariate copula* is a bivariate df C defined on $I^2 = [0,1] \times [0,1]$ with uniformly distributed margins. That is,

$$C(u,v) = \mathrm{Pr}(U \leq u, V \leq v)$$

where $U, V \sim U(0,1)$. As a result, it satisfies the following properties

- $C(u,0) = C(0,v) = 0$, $C(1,v) = v$, $C(u,1) = u$ for every $u, v \in [0,1]$

- $0 \leq C(u,v) \leq 1$

- For every $u_1 \leq u_2$, and $v_1 \leq v_2$ and $u_1, u_2, v_1, v_2 \in [0,1]$, the following inequality holds: $C(u_1,v_1) - C(u_2,v_1) - C(u_1,v_2) + C(u_2,v_2) \geq 0$

The idea of a copula is to separate a joint df F_{XY} into a part that describes the dependence between X and Y, and parts that only describe the marginal behavior. To see this, X and Y may be transformed into uniform random variables U and V via $U = F_X(X)$ and $V = F_Y(Y)$. Let the joint df of (U, V) be the copula C. Then, it follows that

$$\begin{aligned} F_{XY}(x, y) &= \mathsf{Pr}(X \le x, Y \le y) \\ &= \mathsf{Pr}(F_X(X) \le F_X(x), F_Y(Y) \le F_Y(y)) \\ &= C(F_X(x), F_Y(x)) = C(u, v) \end{aligned}$$

and so the joint df F_{XY} can be described by the margins F_X and F_Y and the copula C. The copula C captures the dependence structure between X and Y.

Sklar's Theorem

Let F_{XY} be a joint df with margins F_X and F_Y. Then there exists a copula C such that for all $x, y \in [-\infty, \infty]$

$$F_{XY}(x, y) = C(F_X(x), F_Y(y)) \tag{1}$$

If F_X and F_Y are continuous then C is unique. Otherwise, C is uniquely defined on Range $F_X \times$ Range F_Y. Conversely, if C is a copula and F_X and F_Y are univariate dfs, then F_{XY} defined in (1) is a joint df with margins F_X and F_Y.

Remarks

- Sklar's theorem (Sklar (1959)) above shows that the copula associated with a continuous df F_{XY} couples the margins F_X and F_Y with a dependence structure to uniquely create F_{XY}. As such, it is often stated that the copula of X and Y is the df C of $F_X(x)$ and $F_Y(y)$.

- The copula C of X and Y has the property that it is invariant to strictly increasing transformations of the margins F_X and F_Y. That is if T_X and T_Y are strictly increasing functions then $T_X(X)$ and $T_Y(Y))$ have the same copula as X and Y. This property of copulas is useful for defining measures of dependence.

Examples of Simple Copulas

If X and Y are independent then their copula satisfies

$$C(u, v) = u \cdot v$$

This copula is called the *independent copula* or *product copula*. Its form follows from the definition of independence.

Suppose that X and Y are perfectly positively dependent or *co-monotonic*. This occurs if

$$Y = T(X)$$

and T is a strictly increasing transformation. Then the copula for X and Y satisfies

$$C(u, v) = \min(u, v)$$

Notice that this is df for the pair (U, U) where $U \sim U(0, 1)$.

Finally, suppose that X and Y are perfectly negatively dependent or *counter-monotonic*. This occurs if

$$Y = T(X)$$

and T is a strictly decreasing transformation. Then the copula for X and Y satisfies

$$C(u, v) = \max(u + v - 1, 0)$$

The above is the df for the pair $(U, 1 - U)$.

The copulas for co-monotonic and counter-monotonic random variables form the so-called Fréchet bounds for any copula $C(u, v)$:

$$\max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v)$$

Copula Density

The copula density is defined by

$$c(u,v) = \frac{\partial^2}{\partial u \partial v} C(u,v)$$

Let F_{XY} be a joint df with margins F_X and F_Y. Then, using the chain-rule, the joint density of X and Y may be recovered using

$$\begin{aligned} f_{XY}(x,y) &= \frac{\partial^2}{\partial x \partial y} F_{XY}(x,y) \\ &= \frac{\partial^2}{\partial u \partial v} C(F_X(x), F_Y(y)) \frac{\partial F_X}{\partial x} \frac{\partial F_Y}{\partial y} \\ &= c(F_X(x), F_Y(y)) \cdot f_X(x) f_Y(y) \end{aligned}$$

The above result shows that it is always possible to specify a bivariate density by specifying the marginal densities and a copula density.

Dependence Measures and Copulas

For two random variables X and Y, four desirable properties of a general, single number measure of dependence $\delta(X, Y)$ are:

1. $\delta(X, Y) = \delta(Y, X)$

2. $-1 \leq \delta(X, Y) \leq 1$

3. $\delta(X, Y) = 1$ if X and Y are co-monotonic; $\delta(X, Y) = -1$ if X and Y are counter-monotonic

4. If T is strictly monotonic, then

$$\delta(T(X), Y) = \begin{cases} \delta(X, Y) & T \text{ increasing} \\ -\delta(X, Y) & T \text{ decreasing} \end{cases}$$

Remark

The usual (Pearson) linear correlation only satisfies the first two properties. The rank correlations Spearman's rho and Kendall's tau satisfy all four properties.

Pearson's Linear Correlation

The Pearson correlation coefficient

$$\rho = \frac{cov(X,Y)}{\sqrt{var(X)var(Y)}}$$

gives a scalar summary of the linear dependence between X and Y.

If $Y = a + bX$ then $\rho = \pm 1$.

If X and Y are independent then $\rho = 0$.

The following are shortcomings of Pearson's linear correlation:

- ρ requires that both $var(X)$ and $var(Y)$ exist.
- $\rho = 0$ does not imply independence. Only if X and Y are bivariate normal does $\rho = 0$ imply independence.
- ρ is not invariant under nonlinear strictly increasing transformations
- marginal distributions and correlation do not determine the joint distribution. This is only true for the bivariate normal distribution.
- For given marginal distributions F_X and F_Y, $\rho \in [\rho_{\text{min}}, \rho_{\text{max}}]$ and it may be the case that $\rho_{\text{min}} > -1$ and $\rho_{\text{max}} < 1$.

Concordance Measures

Suppose the random variables X and Y represent financial returns or payoffs. It is often the case that both X and Y take either large or small values together, while it is seldom the case that X takes a large value and, at the same time, Y takes a small value (or vice-versa). The concept of *concordance* is used to measure this type of association.

- Concordance measures have the useful property of being invariant to increasing transformations of X and Y.

- Concordance measures may be expressed as a function of the copula between X and Y.

- Since the linear correlation ρ is not invariant to increasing transformations of X and Y, it is does not measure concordance.

- Two common measures of concordance are Kendall's tau statistic and Spearman's rho statistic.

Kendall's tau statistic

Let F be a continuous bivariate cdf, and let (X_1, Y_1) and (X_2, Y_2) be two independent pairs of random variables from this distribution. The vectors (X_1, Y_1) and (X_2, Y_2) are said to be *concordant* if $X_1 > X_2$ whenever $Y_1 > Y_2$, and $X_1 < X_2$ whenever $Y_1 < Y_2$; and they are said to be *discordant* in the opposite case.

Kendall's tau statistic for the distribution F is a measure of concordance and is defined as

$$\begin{aligned} \tau &= \Pr\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - \Pr\{(X_1 - X_2)(Y_1 - Y_2) < 0\} \\ &= E\left[\text{sign}\left\{(X_1 - X_2)(Y_1 - Y_2)\right\}\right] \end{aligned}$$

Relationship between Copula and Kendall's tau

If C is the copula associated with F, then it can be shown that

$$\tau = 4E[C(U,V)] = 4\int\int_{I^2} C dC - 1 = 4\int\int_{I^2} C(u,v)c(u,v)dudv - 1$$

where $c(u,v)$ is the copula density.

The empirical estimate of τ for a sample of size n is the number of the sample's concordant pairs minus the number of discordant pairs divided by the total number of pairs $\binom{n}{2}$:

$$\hat{\tau} = \frac{1}{\binom{n}{2}} \sum_{1 \le i \le j \le n} \text{sign}\left(\left(x_i - x_j\right)\left(y_i - y_j\right)\right)$$

Spearman's rho statistic

For a pair of random variables (X, Y) with joint df F and marginal distributions F_X and F_Y, *Spearman's rho statistic*, ρ_S, is defined as the (Pearson) correlation between $F_X(X)$ and $F_Y(Y)$. It is a measure of *rank correlation* in terms of the integral transforms of X and Y.

For a copula associated with X and Y, it can be shown that

$$\rho_S = cor(F_X(X), F_Y(Y)) = 12 \int\int_{I^2} C(u, v) du dv - 3$$

For a sample of size n, ρ_S may be estimated using

$$\hat{\rho}_S = \frac{12}{n\left(n^2 - 1\right)} \sum_{i=1}^{n} \left(\text{rank}(x_i) - \frac{n+1}{2}\right)\left(\text{rank}(y_i) - \frac{n+1}{2}\right)$$

Tail dependence measures

Tail dependence measures are used to capture dependence in the joint tail of bivariate distributions.

The coefficient of upper tail dependence is defined as

$$\lambda_u(X, Y) = \lim_{q \to 1} \Pr(Y > VaR_q(Y) | X > VaR_q(X))$$

where $VaR_q(X)$ and $VaR_q(Y)$ denote the $100 \cdot q$th percent quantiles of X and Y, respectively. Loosely speaking, $\lambda_u(X, Y)$ measures the probability that Y is above a high quantile given that X is above a high quantile.

Similarly, the coefficient of lower tail dependence is

$$\lambda_l(X, Y) = \lim_{q \to 0} \Pr(Y \leq VaR_q(Y) | X \leq VaR_q(X))$$

and measures the probability that Y is below a low quantile given that X is below a low quantile.

It can be shown that the coefficients of tail dependence are functions of the copula C given by

$$\begin{aligned}\lambda_u &= \lim_{q\to 1} \frac{1-2q+C(q,q)}{1-q} \\ \lambda_l &= \lim_{q\to 0} \frac{C(q,q)}{q}\end{aligned}$$

If $\lambda_u \in (0,1]$, then there is upper tail dependence; if $\lambda_u = 0$ then there is independence in the upper tail. Similarly, if $\lambda_l \in (0,1]$, then there is lower tail dependence; if $\lambda_l = 0$ then there is independence in the lower tail.

Elliptical Copulas

Let X be an n-dimensional random vector, $\mu \in \mathbb{R}^n$ and $\mathbf{\Sigma}$ a $n \times n$ covariance matrix. If X has an elliptical distribution then its density is of the form

$$f(x) = |\mathbf{\Sigma}|^{-1/2} g\left((x-\mu)'\mathbf{\Sigma}^{-1}(x-\mu)\right)$$

for some scalar non-negative function $g(\cdot)$.Note: the contours of equal density form ellipsoids in $\mathbb{R}^n$.

Let F be the multivariate CDF of an elliptical distribution. Let F_i be the CDF of the ith margin and F_i^{-1} be its quantile function $i = 1, \ldots, n$. The elliptical copula determined by F is

$$C(u_1, \ldots, u_n) = F\left[F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\right]$$

The two most common elliptical copulas are the normal (Gaussian) and the Student t.

Normal (Gaussian) Copula (See FMUND chapter 6)

One of the most frequently used copulas for financial modeling is the copula of a standard bivariate normal distribution with correlation parameter ρ defined by

$$\begin{aligned} C(u, v; \rho) &= \Phi_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \\ &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{ -\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)} \right\} dx dy \end{aligned}$$

where $\Phi^{-1}(\cdot)$ is the quantile function of the standard normal distribution, and Φ_ρ is the joint cumulative distribution function of a standard bivariate normal distribution with correlation coefficient ρ.

The density of the normal copula is given by

$$
\begin{aligned}
c(u, v; \rho) &= \frac{1}{|R|^{1/2}} \exp\left(-\frac{1}{2}\psi'(R^{-1} - I_2)\psi\right) \\
\psi &= (\Phi^{-1}(u), \Phi^{-1}(v))'
\end{aligned}
$$

and R is the correlation matrix between u and v with correlation coefficient ρ.

Remarks

- Bivariate distributions whose dependence is captured by the Gaussian copula are called *meta-Gaussian distributions*.

- From Sklar's theorem, the *normal copula* generates the bivariate standard normal distribution if and only if the margins are standard normal. For any other margins, the normal copula does not generate a bivariate standard normal distribution.

- For the normal copula, Kendall's tau and Spearman's rho are given by

$$\begin{aligned} \tau &= \frac{2}{\pi} \arcsin \rho \\ \rho_S &= \frac{6}{\pi} \arcsin \frac{\rho}{2} \end{aligned}$$

- Except for the case $\rho = 1$, the normal copula does not display either lower or upper tail dependence:

$$\lambda_L = \lambda_U = \begin{cases} 0 & \text{for } \rho < 1 \\ 1 & \text{for } \rho = 1 \end{cases}$$

Student t Copula (See FMUND chapter 6)

The Student t copula with correlation parameter ρ and degrees of freedom parameter δ is defined by

$$
\begin{aligned}
C(u, v; \rho, \delta) &= t_{\rho,\delta}(t_\delta^{-1}(u), t_\delta^{-1}(v)) \\
&= \int_{-\infty}^{t_\delta^{-1}(u)} \int_{-\infty}^{t_\delta^{-1}(v)} \frac{\Gamma\left(\frac{\delta+2}{2}\right)}{\Gamma\left(\frac{\delta}{2}\right) \pi\delta\sqrt{1-\rho^2}} \left(1 + \frac{\psi' R^{-1} \psi}{n}\right)^{-\frac{\delta+2}{2}} d\psi \\
\psi &= (t_\delta^{-1}(u), t_\delta^{-1}(v))'
\end{aligned}
$$

where t_δ^{-1} denotes the quantile function of the Student t with δ degrees of freedom.

The density of the Student t copula is

$$c(u, v; \rho, \delta) = \frac{1}{\sqrt{|R|}} \frac{\Gamma\left(\frac{\delta+2}{2}\right)\Gamma\left(\frac{\delta}{2}\right)\left(1 + \frac{1}{\delta}\psi' R^{-1}\psi\right)^{-\frac{\delta+2}{2}}}{\Gamma\left(\frac{\delta+2}{2}\right)^2 \Pi_{i=1}^{2}(1 + \frac{1}{\delta}\psi_i^2)^{-\frac{\delta+2}{2}}}$$

and Kentall's tau is given by

$$\tau = \frac{2}{\pi}\arcsin(\rho)$$

Note: The Student t copula exhibits both upper and lower tail dependence.

Archimedean Copulas

Archimedean copulas are copulas that may be written in the form

$$C(u, v) = \phi^{-1}\left[\phi(u) + \phi(v)\right]$$

for a function $\phi : I \longrightarrow \mathbb{R}^{+}$ that is continuous, strictly decreasing, convex and satisfies $\phi(0) = \infty$ and $\phi(1) = 0$.

The function ϕ is called the *Archimedean generator*, and ϕ^{-1} is its inverse function.

The density of an Archimedean copula may be determined using

$$c(u, v) = \frac{-\phi''\left(C(u, v)\right)\phi'(u)\phi'(v)}{\left(\phi'(C(u, v))\right)^{3}}$$

where ϕ' and ϕ'' denote the first and second derivatives of ϕ, respectively.

For an Archimedian copula, Kendall's tau may be computed using

$$\tau = 4 \int_I \frac{\phi(v)}{\phi'(v)} dv + 1$$

Gumbel copula

The *Gumbel copula* with parameter δ is given by:

$$C(u, v) = \exp\left\{-[(-\ln(u))^{\delta} + (-\ln(v))^{\delta}]^{1/\delta}\right\}, \ \delta \geq 1$$

and has generator function $\phi(t) = (-\ln t)^{\delta}$.

The parameter δ controls the strength of dependence. When $\delta = 1$, there is no dependence; when $\delta = +\infty$ there is perfect dependence.

It can be shown that Kendall's tau is given by

$$\tau = 1 - \delta^{-1}$$

Further, the Gumbel copula exhibits upper tail dependency with

$$\lambda_U = 2 - 2^{1/\delta}$$

Kimeldorf-Sampson (Clayton) copula

The *Kimeldorf and Sampson copula* or *Clayton copula* has the following form:

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

where $0 < \delta < \infty$, and the generator function is

$$\phi(t) = t^{-\delta} - 1$$

The parameter δ controls the strength of dependence. When $\delta = 0$, there is no dependence; when $\delta = +\infty$ there is perfect dependence.

Kendall's tau is given by

$$\tau = \frac{\delta}{\delta + 2}$$

and it exhibits only lower tail dependency

$$\lambda_L = 2^{-1/\delta}$$

Nonparametric Copula

Deheuvels (1978) proposed the following non-parametric estimate of a copula C. Let $u_{(1)} \leq u_{(2)} \leq \cdots \leq u_{(n)}$ and $v_{(1)} \leq v_{(2)} \leq \cdots \leq v_{(n)}$ be the order statistics of the univariate samples from a copula C. The *empirical copula* $\hat{C}$ is defined at the points $\left(\frac{i}{n}, \frac{j}{n}\right)$ by

$$\hat{C}\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n} \sum_{k=1}^{n} 1_{\{u_k \leq u_{(i)}, v_k \leq v_{(j)}\}}, \quad i, j = 1, 2, \ldots, n.$$

Deheuvels proved that $\hat{C}$ converges uniformly to C as the sample size tends to infinity. The *empirical copula frequency* $\hat{c}$ is given by

$$\hat{c}\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} \frac{1}{n} & \text{if } (u_{(i)}, v_{(j)}) \text{ is an element of the sample} \\ 0 & \text{otherwise} \end{cases}$$

Estimates of Spearman's rho and Kendall's tau for a sample of size n may be computed from the empirical copula using

$$\begin{aligned}
\hat{\rho}_S &= \frac{12}{n^2-1}\sum_{j=1}^{n}\sum_{i=1}^{n}\left[\hat{C}\left(\frac{i}{n},\frac{j}{n}\right)-\frac{i}{n}\cdot\frac{j}{n}\right] \\
\hat{\tau} &= \frac{2n}{n-1}\sum_{j=2}^{n}\sum_{i=2}^{n}\left[\hat{C}\left(\frac{i}{n},\frac{j}{n}\right)\hat{C}\left(\frac{i-1}{n},\frac{j-1}{n}\right)\right. \\
&\quad \left. -\hat{C}\left(\frac{i}{n},\frac{j-1}{n}\right)\hat{C}\left(\frac{i-1}{n},\frac{j}{n}\right)\right]
\end{aligned}$$

The tail index parameters may be inferred from the empirical copula by plotting

$$\hat{\lambda}_U(q) = \frac{1 - 2q + \hat{C}(q,q)}{1-q}$$

$$\hat{\lambda}_L(q) = \frac{\hat{C}(q,q)}{q}$$

as functions of q and visually observing convergence as $q \longrightarrow 1$ and $q \longrightarrow 0$, respectively.

Goodness-of-Fit Tests for Copulas

- The empirical copula can be used to evaluate the fit of a particular parametric copula.

- The hypothesis to be tested is

$$
\begin{aligned}
H_0 &: C_\theta \text{ is the true copula} \\
H_1 &: C_\theta \text{ is not the true copula}
\end{aligned}
$$

  where C_θ represents a parametric copula

- The test statistics is a Cramer-von Mises statistic of the form

$$S_n = \sum_{i=1}^{n} \left(\hat{C}\left(\frac{i}{n}, \frac{j}{n}\right) - C_{\hat{\theta}}\left(\frac{i}{n}, \frac{j}{n}\right) \right)$$

  where $C_{\hat{\theta}}$ represents a fitted parametric copula

- Reject $H_0 : C_\theta$ is the true copula if S_n is large

- P-value for the test can be computed by a bootstrap procedure

Creating Arbitrary Bivariate Distributions

From Sklar's theorem, any bivariate CDF $F_{XY}(x, y)$ can be constructed by specifying

- the margins $F_X(x)$ and $F_Y(y)$
- a copula $C(u, v)$

Then

$$F_{XY}(x, y) = C(F_X(x), F_Y(x))$$

Maximum Likelihood Estimation

- Let $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ denote a random sample from a bivariate distribution F with marginal distributions F_X and F_Y (with density functions f_X and f_Y) and copula C with density c.

- The joint density of (x_i, y_i) may be represented as

$$f(x_i, y_i; \boldsymbol{\eta}) = c(F_X(x_i; \boldsymbol{\alpha}_x), F_Y(y_i; \boldsymbol{\alpha}_y); \boldsymbol{\theta}) f_X(x_i; \boldsymbol{\alpha}_x) f_Y(x_i; \boldsymbol{\alpha}_y)$$

  where $\boldsymbol{\alpha}_x$ are the parameters for the marginal distribution F_X, $\boldsymbol{\alpha}_y$ are the parameters for the the marginal distribution F_Y, $\boldsymbol{\theta}$ are the parameters for the the copula density c, and $\boldsymbol{\eta} = (\boldsymbol{\alpha}'_{\mathbf{x}}, \boldsymbol{\alpha}'_y, \boldsymbol{\theta}')'$ are the parameters of the joint density.

- The *exact log-likelihood* function is then

$$l(\boldsymbol{\eta}; \mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} \left(\begin{array}{c} \ln c(F_X(x_i; \boldsymbol{\alpha}_x), F_Y(y_i; \boldsymbol{\alpha}_y); \boldsymbol{\theta}) \\ + \ln f_X(x_i; \boldsymbol{\alpha}_x) + \ln f_Y(x_i; \boldsymbol{\alpha}_y) \end{array} \right)$$

- The exact maximum likelihood estimator (MLE) is defined as

$$\hat{\boldsymbol{\eta}}_{MLE} = \arg\max_{\boldsymbol{\eta}} l(\boldsymbol{\eta}; \mathbf{x}, \mathbf{y})$$

- Exact MLE estimates the marginal distribution parameters $\boldsymbol{\alpha}_x$ and $\boldsymbol{\alpha}_y$ jointly with the copula parameters $\boldsymbol{\theta}$. This can be computationaly difficult for high dimensional models.

Inference Functions for Margins Estimation

Instead of maximizing the exact likelihood as a function of $\boldsymbol{\eta}$, the copula parameters $\boldsymbol{\theta}$ may be estimated using a two-stage procedure.

- First, the marginal distributions F_X and F_Y are estimated. This could be done using parametric models (e.g. normal, Student-t, skew-t etc. distributions), the empirical CDF, or a combination of empirical CDFs with parametric CDFs.

- Next, given estimates $\hat{F}_X$ and $\hat{F}_Y$, form a pseudo-sample of observations from the copula

$$(\hat{u}_i, \hat{v}_i) = (\hat{F}_X(x_i), \hat{F}_Y(y_i)),\ i = 1, \ldots, n$$

- Then, for a specified parametric copula $C(u, v; \boldsymbol{\theta})$ with density $c(u, v; \boldsymbol{\theta})$ and unknown copula parameters $\boldsymbol{\theta}$, the log-likelihood

$$l(\boldsymbol{\theta} : \hat{\mathbf{u}}, \hat{\mathbf{v}}) = \sum_{i=1}^{n} \ln c(\hat{u}_{i,} \hat{v}_i; \boldsymbol{\theta})$$

  is maximized using standard numerical methods.

This two-step method, due to Joe and Xu (1996), is called the *inference functions for margins* (IFM) method and the resulting estimator of $\boldsymbol{\theta}$ is called the *IFM estimator* (IFME).

Under standard regularity conditions, the IFME is consistent and asymptotically normally distributed. In particular Joe (1997) shows that the IFME often nearly as efficient as the exact MLE.