Copulas and Measures of Dependence

Uttara Naik-Nimbalkar

December 28, 2014

• Measures for determining the relationship between two variables: **the Pearson's correlation coefficient** , Kendalls tau and Spearmans rho.

• Insurance covering several lives, like last-survivor annuities for a married couple have to be fairly priced. The data indicate a very high positive dependence in the time of deaths between coupled lives or associated individuals. The dependence could be because of conditions such as common disaster, common life style, or the brokenheart syndrome.

• Financial instruments based on more than one asset: model dependence between the asset prices or returns ($\frac{S_{i+1}-S_i}{S_i}$).

• In Finance: data display dependence among extreme values and inferences based on multivariate tail probabilities are needed.

Multivariate Gaussian distributions are unsuitable as they do not have tail dependence. It has been widely observed that market crashes or financial crises often occur in different markets and countries at about the same time period even when the correlation among those markets is fairly low.

• Patients in a clinical trial may usually drop out from treatment when the prognosis is poor and hence their dropout and survival times may be dependent.

Linear Correlation

Definition: The Pearson (linear) correlation coefficient between two random variables X and Y is

$$\rho = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}.$$

Problems with the Pearson correlation coefficient

• The Pearson correlation is a measure for linear dependence only.

• Extreme events are frequently observed in security prices or returns; variance of returns in such securities tends to be infinite, that is,the linear correlation between these securities is undefined.

- The linear correlation is not invariant under nonlinear strictly increasing transformations, implying that returns might be uncorrelated whereas prices are correlated or vice versa.
- Linear correlation only measures the degree of dependence but does not clearly discover the structure of dependence.

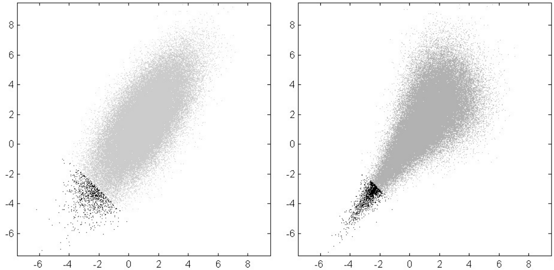


Figure: Simulated values with equal margins and equal estimated correlation coefficient $\hat{\rho} = 0.701$ but different overall dependence structure.

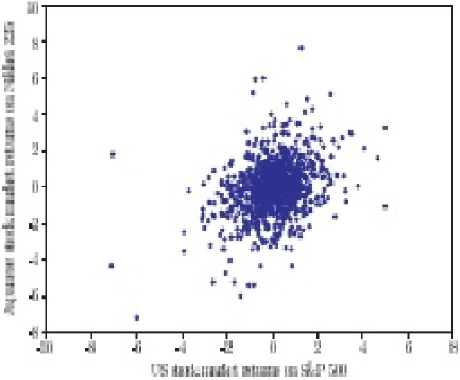


Figure: US versus Japanese Stock Market

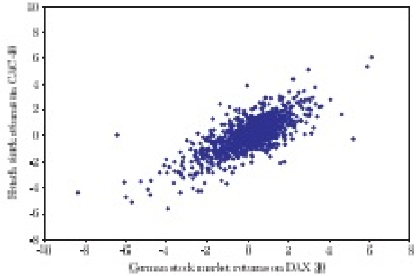


Figure: German versus French Stock Market

• The linear correlation coefficient does not completely determine the joint distribution.

• It is in general not possible to construct a joint distribution of the margins with arbitrary associated correlation coefficient.

• Given that the marginal cdf F_Y can be obtained by application of a non-linear deterministic function f on X; , the linear correlation associated with $(f(X), X)$ is in general lower than 1.

Let X $N(0,1)$ and $Y = X^2$. Then

$Cov(X,Y) = E[X^3] - E[X]E[X^2] = 0$. That is the correlation is 0.

However given X, one can predict Y.

Consider the bivariate distribution with uniform margins. For $(u_1, u_2) \in [0,1] \times [0,1]$,

$$C(u_1, u_2) = u_1 u_2 + \alpha[u_1(u_1 - 1)(2u_1 - 1)][u_2(u_2 - 1)(2u_2 - 1)]$$

with $\alpha \in [-1, 2]$.

If the margins F_1 and F_2 are continous and symmetric, the Pearson correlation is zero, but for $\alpha \neq 0$ the random variables are not independent.

Let U_1 and U_2 be two $U(0,1)$ r.v.s with joint distribution

$$C(u_1,u_2) = \begin{cases} u_1 & 0 \le u_1 \le u_2/2 \le 1/2 \\ u_2/2, & 0 \le u_2/2 \le u_1 \le 1 - u_2/2 \\ u_1 + u_2 - 1, & 1/2 \le 1 - u_2/2 \le u_1 \le 1 \end{cases},$$

$Cov(U_1, U_2) = 0$, but $P[U_2 = 1 - |2U_1 - 1|] = 1$. That is, the two r.v.s are *uncorrelated* but one can be perfectly predicted from the other.

Copula Function

(Bivariate distribution function with uniform margins. A d-variate distribution with uniform margins.) A (two-dimensional) copula $C(u, v)$ is a function from $\mathcal{I}^2 = [0, 1]^2$ to $\mathcal{I} = [0, 1]$ satisfying

- $C(u, 0) = 0 = C(0, v)$.
- $C(u, 1) = u,\ C(1, v) = v$
- C(u,v) is 2-increasing

  (That is, if $0 \le u_1 \le u_2 \le 1$ and $0 \le v_1 \le v_2 \le 1$ then $C(u_2, v_2) - C(u_2, v_1) - C(u_1, v_2) + C(u_1, v_1) \ge 0$.)

Sklar's theorem (Sklar 1959)

Let H be a bivariate distribution function with margins F and G.

Then $\exists$ a copula C such that for all $x, y \in [-\infty, \infty]$,

$$H(x, y) = C(F(x), G(y)). \tag{1}$$

IF F and G are continuous, then C is unique; otherwise C is uniquely determined on $Range(F) \times Range(G)$.

Conversely, if C is a copula and F and G are distribution functions, then the function H defined in (01) is a bivariate distribution function with margins F and G.

Example:

Gumbel's bivariate d.f. with exponential margins is:

$H(x, y) = 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)}$.

To obtain the corresponding copula, replace x by $F^{-1}(u)$ and y by $G^{-1}(v)$ in $H(x, y)$, where F and G are exponential d.f. with mean=1. ($F(x) = 1 - exp(-x) = u$ and $F^{-1}(u) = -ln(1 - u)$.)

The resulting copula is:

$$C(u, v) = u + v - 1 + (1 - u)(1 - v)e^{-\theta \ ln(1-u)ln(1-v)}.$$

Bivariate d.f.s with specified marginals F_1 and G_1 can be obtained by substituting $u = F_1(x)$ and $v = G_1(x)$ in $C(u, v)$.

Properties

- Let X and Y be continuous r.v.s and let C_{XY} denote the copula corresponding to the joint distribution of (X, Y). Then C_{XY} is invariant under strictly increasing transformations of X and Y.
- **Survival copula C^*:** Consider

$$C^*(u, v) = u + v - 1 + C(1 - u, 1 - v).$$

Then

$$P[X > x, Y > y] = C^*(\bar{F}(x), \bar{G}(y)).$$

The function C^* is called the survival copula of (X, Y).

- **The Fréchet-Hoeffding Bounds**:

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

  $C_L(u,v)$ and $C_U(u,v)$ are copulas.

The r.v.s X and Y are independent iff $C(u,v) = uv$. The r.v. Y is almost surely an increasing function of X iff $C(u,v) = min\{u,v\}$ and Y is almost surely a decreasing function of X iff $C(u,v) = max\{u+v-1,0\}$

- **Lipschitz Continuity**: For every $(u_1, v_1), (u_2, v_2) \in \mathcal{I}^2$,

$$|C(u_2, v_2) - C(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|.$$

  Thus copula is uniformly continuous on $\mathcal{I}^2$.

- **Partial derivatives**: For any v in $\mathcal{I}$, the partial derivative $\partial C(u, v)/\partial u$ exists for almost all u, and for such v and u,

$$0 \leq \frac{\partial}{\partial u} C(u, v) \leq 1.$$

  Further $\partial C(u, v)/\partial u$ as a function of v is defined and nondecreasing almost everywhere on $\mathcal{I}$.

  The above result holds for all first order partial derivatives.

- **Partial derivatives and conditional expectations**:

$$E[I_{[X\le x]}|Y) = P(X \le x|Y) = \frac{\partial}{\partial v}C(u,v)|_{u=F(x),v=F(Y)}, \quad a.s.,$$

$$E[I_{[Y\le y]}|X) = P(Y \le y|X) = \frac{\partial}{\partial u}C(u,v)|_{u=F(X),v=F(y)}, \quad a.s.$$

.

The copula of a distribution is invariant under strictly increasing transformations of the marginals. • If $(X_1, X_2, \cdots, X_d)$ is a random vector with continuous margins and copula C and if $T_1, \cdots, T_d)$ are strictly increasing functions then $()T_1(X_1), T_2(X_2), \cdots T_d(X_d))$ has the same copula C.

Relation with dependence measures (notions of positive dependence)

Let (X_1, Y_1) and (X_2, Y_2) be independent vectors (of continuous r.v.s.) with joint distribution functions H_1 and H_2 and copulas C_1 and C_2 respectively. Further, X_1 and X_2 have the same distribution F and Y_1 and Y_2 have the common distribution G.

Concordance function:

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0].$$

$$\text{Then } Q = Q(C_1, C_2) = 4 \int \int_{\mathcal{I}^2} C_2(u, v) dC_1(u, v) - 1,$$

i.e., Q depends only on the copulas.

Intutively, a pair of r.v.s are concordant if large values of one variable correspond to large values of the other and small values of one with small values of the other.

Concordance measures Kendall's tau $\tau_{X,Y}$:

$$\tau_{X,Y} = 4 \int\int_{\mathcal{I}^2} C(u,v)dC(u,v) = Q(C,C)$$

Spearman's rho $\rho_{X,Y}$:

$$\rho_{XY} = 3Q(C,\Pi) = 12 \int\int_{\mathcal{I}^2} C(u,v)dudv - 3,$$

where $\Pi(u,v) = uv$,

$$\rho_{X,Y} = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]),$$

where (X_1, Y_1) and (X_2, Y_3) have the same margins but one vector has distribution function H, while the components of the other are independent.

Gini's measure of association:

$$\gamma_{X,Y} = Q(C, C_L) + Q(C, C_U) = 4[\int_0^1 C(u, 1-u)du - \int_0^1 (u - C(u, u))du].$$

where $M(u, v) = min(u, v)$ and $W(u, v) = max(u + v - 1, 0)$.

Quadrant Dependence: The r.v.s X and Y are said to be positively quadrant dependent (PQD) if

$$P[X \leq x, Y \leq y] \geq P[X \leq x]P[Y \leq y], \quad \text{for all } (x, y) \in R^2,$$

which is equivalent to

$$C(u, v) \geq uv, \quad \text{for all } (u, v) \in \mathcal{I}^2.$$

(Similarly NQD copulas are defined).

Tail Monotonicity: Y is left tail decreasing in X ($LTD(Y|X)$) if $P[Y \leq y|X \leq x]$ is a nonincreasing function of x for all y.

equivalently: ($LTD(Y|X)$) iff for each $v \in \mathcal{I}$, $C(u, v)/u$ is nonincreasing in u, **or** ($LTD(Y|X)$) iff for each $v \in \mathcal{I}$, $\partial C(u, v)/\partial u \leq C(u, v)/u$ for almost all u.

Y is right tail increasing in X ($RTI(Y|X)$) if $P[Y > y|X > x]$ is a nondecreasing function of x for all y.

equivalently: ($RTI(Y|X)$) iff for each $v \in \mathcal{I}$, $[v - C(u, v)]/(1 - u)$ is nonincreasing in u, **or** ($RTI(Y|X)$) iff for each $v \in \mathcal{I}$, $\partial C(u, v)/\partial u \geq [v - C(u, v)]/(1 - u)$ for almost all u.

Let X and Y represent lifetimes of components 1 and 2 respectively. Then $(LTD(Y|X))$ says that the probability that component 2 has a short lifetime decreases as the lifetime of component 1 increases.

(The negative dependence concepts LTI and RTD can be defined in an analogous manner.)

Stochastic Monotonicity: Y is said to be stochastically increasing in X ($SI(Y|X)$) if $P[Y > y|X = x]$ is a nondecreasing function of x for all y.

equivalently: ($SI(Y|X)$) iff for each $v \in \mathcal{I}$, and for almost all u, $\partial C(u, v)/\partial u$ is nondecreasing in u.

Tail Dependence

Upper tail dependence parameter λ_U:

$$\lambda_U = \lim_{t \to 1^-} P[Y > G^{-1}(t) | X > F^{-1}(t)] = 2 - \lim_{t \to 1^-} \frac{1 - C(t,t)}{1-t},$$

if the limit exists. It is the limit (if it exists) as T approaches 1 of the probability that Y is greater than the $100t - th$ percentile of G given that X is greater than the $100t - th$ percentile of F. The r.v.s are said to be upper tail dependent if $\lambda_U \neq 0$, and are said to have no upper tail dependence if $\lambda_U = 0$.

Similarly a lower tail dependence parameter λ_L is defined and in terms of the copula,

$\lambda_L = \lim_{t \to 0^+} \frac{C(t,t)}{t}$.

A measure of association δ_{XY} (δ_C) between two r.v.s X and Y is a measure of dependence if it satisfies

1. δ_{XY} is defined for every pair of r.v.s X and Y,
2. $\delta_{XY} = \delta_{YX}$,
3. $0 \leq \delta_{XY} \leq 1$,
4. $\delta_{XY} = 0$ **iff** X **and** Y **are independent,**
5. $\delta_{XY} = 1$ iff X and Y are monotone functions of each other,
6. $\delta_{\alpha(X)\beta(Y)} = \delta_{XY}$ for strictly monotone functions α and β,
7. If a sequence $\{C_n\}$ of copulas converges point wise to C, then $\lim_{n\to\infty} \delta_{C_n} = \delta_C$.

$L_p, 1 \leq p < \infty$ **distance between the copulas** C **and** $\prod$**:**

$(k_p \int\int_{\mathcal{I}^2} |C(u,v) - uv|^p dudv)^{1/p}$, where k_p is such that the above expression equals 1 when $C = C_U$ or C_L.

Schweizer and Wolff's σ : (p=1)

$\sigma_{X,Y} = \sigma_C = 12 \int\int_{\mathcal{I}^2} |C(u,v) - uv| dudv$

Hoeffding's dependence index $\Phi^2_{X,Y}$**:** (p = 2)

$\Phi_{X,Y} = \Phi_C = (90 \int\int_{\mathcal{I}^2} |C(u,v) - uv|^2 dudv)^{1/2}$.

L_∞ **distance:** $\Lambda_{X,Y} = \Lambda_C = 4 sup_{u,v \in \mathcal{I}} |C(u,v) - uv|$.

It can be shown that the L_∞ distance does not satisfy property (v) of Definition 2. Further, with respect to this distance, the independence copula can be approximated arbitrarily well by copulas corresponding to mutually completely dependent (predictable) r.v.s. This can not then distinguish between ‘degrees’ of dependence.

Some Families of Copulas:

Elliptical Copulas

Copula functions obtained from multivariate elliptical distributions are known as elipitical copulas. An important example being the normal (Gaussian) copula. The bivariate normal copula is given by:

$$C_\rho(u_1, u_2) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{1-\rho^2}} exp\left(\frac{-(s^2 - 2\rho st + t^2)}{2(1-\rho^2)}\right) dsdt,$$

where $\rho \in (-1, 1)$ and Φ^{-1} denotes the inverse of the univariate normal distribution. A d-variate normal copula is

$$C(u_1, \cdots, u_d) = N_\Sigma(\Phi^{-1}(u_1), \cdots, \Phi^{-1}(u_d)),$$

where Σ is a symmetric, positive definite matrix with diagonal entries $=$

Archimedean Copulas: A d-dimensional copula C is said to belong to the Archimedean family if it admits a representation

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

for all $(u_1, \cdots, u_d)$ in $[0,1]^d$ and some function ϕ called the *Archimedean (additive) generator*. The function $\phi : [0,1] \to [0, \infty)$ should satisfy (i) ϕ is continuous and strictly decreasing function, (ii) Its inverse ϕ^{-1} is differentiable up to order $d-2$ in $(0, \infty)$ and the derivatives satisfy $(-1)^k(\phi^{-1})^{(k)}(x) \geq 0$, for $k = 0, 1, ..., d-2$ and every $x > 0$. (iii) $(-1)^{d-2}(\phi^{-1})^{(d-2)}$ is decreasing and convex in $(0, \infty)$.

Examples of Archimedean copulas:

- Gumbel-Hougaard family:$\phi(t) = (-log\ \ t)^{\theta}$ and the copula $C_{\theta}(u_1, ..., u_d) = exp\left(-\left(\sum_{i=1}^{d}(-log(u_i))^{\theta}\right)^{1/\theta}\right),\ \ theta \geq 1.$ Independence for $theta = 1$. Positive dependence.

  (Gumbel(1960), Hougaard (1986)).

- Clayton family (or Mardia-Takahasi-Clayton family):

  $\phi(t) = (t^{-\theta} - 1)/\theta$ and the copula $C_{\theta}(u_1, ..., u_d) = max\left\{\left(\sum_{i=1}^{d} u_i^{-\theta} - (d-1)\right)^{-1/\theta}, 0\right\},\ \ \theta \geq \frac{-1}{d-1}, \theta \neq 0.$ Independence as the limiting case $\theta \to 0$. Negative

  dependence for $\theta \in [-1, 0)$.(Clayton (1978), Oakes (1982).)

- Frank copula: $\phi(t) = -log\frac{e^{-\theta t}-1}{e^{-\theta}-1}$ and the copula $C_\theta(u_1, ..., u_d) = -\frac{1}{\theta}log\left(1 + \frac{\prod_{i=1}^{d}(e^{-\theta u_i}-1)}{(e^{-}\theta-1)^{d-1}}\right),\ \theta > 0.$ Independence as the limiting case $\theta \to 0$. (Frank (1979)). This is the only Archimedean family in which the survival copula is same as the copula. (This equality also holds for the elliptical copulas).

- Proportional Hazards Frailty models: introduce dependence between survival times $X_1, ..., X_d$ by using an unobserved random variable W, called the frailty. Given the frailty W, the survival times are assumed to be independent. The corresponding survival copula is an Archimedean survival copula with a generator corresponding to the inverse of the Laplace transform of the distribution of the frailty variable. These model exhibit a PQD behavior only.

Let $u_C^1 = u$, and $u_C^{n+1} = C(u, u_C^n)$.

Theorem 1: Let C be an Archimedean copula. Then for any $u,\ v \in \mathcal{I}$, there exists a positive integer n such that $u_C^n < v$.

Construction

- Several methods of constructing bivariate copulas are given in Nelsen (2006). Construction of copulas with some known information about them, such as: support; sections; diagonals.
- Joe(1997, chapter 4) considers a construction method that uses (conditional) pair copulas only, (PCC method)
- Bedford and Cooke (2002) propose graphical PCC methods involving a sequence of nested trees called regular vines.
- Czado (2010): survey of the methods.

The joint density

$$f(x_1 \cdots, x_d) = \prod_{k=2}^{d} f(x_k | x_1, \cdots, x_{k-1}) * f(x_1).$$

Joint density for d=2 in terms of a copula density c_{12} and the marginals can be written as

$$f(x_1, x_2) = c_{12}(F_1(x_1), F_2(x_2)) * f_1(x_1) * f_2(x_2),$$

and the conditional density

$$f(x_1 | x_2) = c_{12}(F_1(x_1), F_2(x_2)) * f_1(x_1).$$

Thus

$$f(x_k|x_1,\cdots,x_{k-1}) = c_{1,k|2,\cdots,k-1} * f(x_k|x_2,\cdots,x_{k-1})$$
$$= [\prod_{s=1}^{k-2} c_{s,k|s+1,\cdots,k-1}] * c_{k-1,k} * f_k(x_k).$$

and

$$f(x_1\cdots,x_d) = [\prod_{k=2}^{d}\prod_{s=1}^{k-2} c_{s,k|s+1,\cdots,k-1}] * [\prod_{k=2}^{d} c_{k-1,k}] * [\prod_{k=1}^{d} f_k(x_k)],$$

where $c_{i,j|i_1,\cdots,i_k} = c_{i,j|i_1,\cdots,i_k}(F(x_i|x_{i_1},\cdots x_{i_k}), F(x_j|x_{i_1},\cdots x_{i_k}))$. In case of conditional independence, the corresponding pair copulas are product copulas with density equal to one.

Example: A D-vine for d=4. D-vines form a subclass of regular vines.

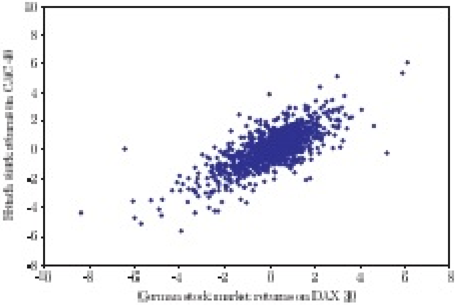
$$T_1: \quad [1] \overset{12}{\leftrightarrow} [2] \overset{23}{\leftrightarrow} [3] \overset{34}{\leftrightarrow} [4]$$

$$T_2: \quad [12] \overset{13|2}{\leftrightarrow} [23] \overset{24|3}{\leftrightarrow} [34]$$

$$T_3: \quad [13|2] \overset{14|23}{\leftrightarrow} [24|3]$$

The corresponding joint density of the D-vine 4-variate distribution is given by

$$f(x_1, \cdots, x_4) = [\prod_{k=1}^{4} f(x_k)] * c_{12} * c_{23} * c_{34} * c_{13|2} * c_{24|3} * c_{14|23},$$


8
6
4
2
0
-2
-4
-6
-8
-8
-6
-4
-2
0
2
4
6
8

The expressions for pair copulas do involve conditional cdf's, which may be constructed recursively from conditional cdf's with a lower dimensional conditioning set and the relation between them and partial derivatives of the corresponding copula

Applications

Measures of Risk in Finance

- Regulations of financial institutions lead to requirements on the capital. These are based on certain risk measures.
- Value-at-risk (VaR_α) is a bound such that the loss over a specified time horizon is less than this bound with probability equal to a given confidence coefficient (1- a).
- VaR_α is the a -quantile of the distribution function of the returns.

The Expected Shortfall measure (ES_α) : Let the returns be denoted by the random variable R. then

$$ES_\alpha = E[R|R \leq VaR_\alpha].$$

- To obtain VaR for an asset estimate the a-quantile of the distribution of the returns on an asset.
- Past data on returns on a single asset, such as a stock or some index fund, can be used to estimate the VaR for a single asset.
- VaR is to be estimated for a portfolio of assets. Thus a model for the joint distribution of the returns on all the assets in the portfolio is required.

Consider a portfolio of two assets. Let X and Y be their continuous returns (over a common time horizon) with distribution functions F_1 and F_2 , respectively. Let λ be the proportion of X in the portfolio. The portfolio return Z is,

$$Z = \lambda X + (1 - \lambda)Y,$$

with the corresponding distribution function $F_Z(z) = Pr[Z \leq z]$.

$$VaR_\alpha \text{ is defined as the } \alpha - th \text{ quantile of } F_Z(z)$$

and

$$ES_\alpha = E[Z|Z \leq VaR_\alpha].$$

If C is the copula of the joint disttribution function of (X, Y), then

$$F_Z(z) = \int_0^1 C_{,2}(v^*, v) dv,$$

where $C_{,2}(u, v) = \frac{\partial C(x,y)}{\partial y}|_{x=u,y=v}$.

and $v^* = F_1\left(\frac{z}{\lambda} - \frac{1-\lambda}{\lambda} F_2^{-1}(v)\right)$.

To obtain VaR_α solve $F_B(z) = \alpha$ for z.

Simulate V_{ijk} from Uniform $(0, 1)$, $k = 1, \cdots, K$, for large K, and $i, j = 1, \cdots, m$. Solve numerically the following equation for z,

$$\alpha = \sum_{i=1}^{m} \sum_{j=1}^{m} C_n \left(\frac{i}{m}, \frac{j}{m} \right) \left\{ \frac{1}{K} \sum_{k=1}^{K} \left[\frac{j - mV_{ijk}}{V_{ijk}(1 - V_{ijk})} P_{m,i}(V^*_{ijk}) P_{m,j}(V_{ijk}) \right] \right\},$$

where $V^*_{ijk} = F_1 \left(\frac{z}{\lambda} - \frac{1-\lambda}{\lambda} F_2^{-1}(V_{ijk}) \right)$.

$$ES_\alpha =$$

$$\frac{1}{\alpha}\int_0^1\int_0^1(\lambda F_1^{-1}(u)+(1-\lambda)F_2^{-1}(v))I[A]\ c(u,v)dudv,$$

where $A = [\lambda F_1^{-1}(u)+(1-\lambda)F_2^{-1}(v) \leq Var_\alpha]$ and $I[A]$ denotes the indicator function of the set A.

Let $\mathbf{X} = (X_1, \cdots, X_d)$ be a random vector with survival function $S_{\mathbf{X}}(\underline{x}) = Pr[X_1 > x_1, \cdots, X_d > x_d]$. Let $\pi^i_{\mathbf{X}}(\underline{x}) = \int_{x_i}^{\infty} S_{\mathbf{X}}(x_1, \cdots x_{i-1}, t, x_{i+1} \cdots, x_d)$. Then the 'stop-loss transform vector' is defined as $\pi_{\mathbf{X}}(\underline{x}) = (\pi^1_{\mathbf{X}}(\underline{x}), \cdots, \pi^d_{\mathbf{X}}(\underline{x}))$. The conditional value-at-risk vector for the confidence level α is defined as:

$$CVaR_{\alpha}[X] = (CVaR^1_{\alpha}[X], \cdots, CVaR^d_{\alpha}[X]),$$

where

$$CVaR^i_{\alpha}[X] = E[X_i | X_j > VaR_{\alpha}[X_j] = VaR_{\alpha}[X_i] + \frac{\pi^i_{\mathbf{X}}[VaR_{\alpha}[X]]}{S_{\mathbf{X}}[VaR_{\alpha}[X]]}.$$

Statistical Inference procedures: (Joe(1997) for general multivariate as well as copula based models. A more recent survey of estimation methods for copula models is given in Chorós et al.(2010).)

The joint density f in terms of the copula density c and the marginal p.d.f. s $f_i's$:

$$f(x_1, \cdots, x_d) = c_\theta(F_1(x_1, \alpha_1), \cdots, F_d(x_d, \alpha_d)) \prod_{i=1}^{d} f(x_i, \alpha_i).$$

If $(x_{1,(j)}, \cdots, x_{d,(j)})$ $j = 1, \cdots, n$ is a random sample from f. Then the log-likelihood L is:

$$L = \sum_{j=1}^{n} log\ c_\theta(F_1(x_{1,(j)}, \alpha_1), \cdots, F_d(x_{d,(j)}, \alpha_d)) + \sum_{i=1}^{d} \sum_{j=1}^{n} log\ f_i(x_{i,(j)}, \alpha_i).$$

An alternative to the maximum likelihood (ML) estimation, a two stage estimation procedure is proposed by Joe(1997).

First stage: obtain estimators of α_i from the marginal log-likelihoods $\sum_{j=1}^{n} log\ \ f_i(x_{i,(j)})$ for $i = 1, \cdots, d$.

Second stage: estimate θ by maximizing $\sum_{j=1}^{n} log\ \ c_\theta(F_1(x_{1,(j)}, \alpha_1), \cdots, F_d(x_{d,(j)}, \alpha_d))$ with $\alpha_i' s$ replaced by their estimators obtained form the first stage.

- The two stage estimators are shown to be consistent and asymptotically normal under the usual regularity conditions.
- If the d-variate d.f. is not absolutely continuous with respect to the Lebesgue measure, different estimation procedures have to be considered.
- Yilmaz and Lawless (2011) discuss inference procedures of copula parameters and model assessment in parametric and semiparametric copula models when the lifetimes are censored.

Empirical Copula :

Let $(U_1, V_1), (U_2, V_2), \cdots, (U_n, V_n)$ be a random sample from a continuous distribution. Define the joint empirical d. f.

$$H_n(u, v) = \frac{1}{n} \sum_{k=1}^{n} \mathcal{I}[U_k \leq u, V_k \leq v],$$

and let $F_n(u) = H_n(u, \infty)$ and $G_n(v) = H_n(\infty, v)$ be its associated marginal distributions. Define the empirical copula function C_n by

$$C_n(u, v) = H_n(F_n^{-1}(u), G_n^{-1}(u)),$$

where $F^{-1}(u) = inf\{t \in R | F(t) \geq u\}, \;\; 0 \leq u \leq 1.$

Deheuvels (1979, 1981): the empirical copula converges uniformly to the true copula a.s. and established the law of iterated logarithm for the empirical copula process $\sqrt{n}(C_n - C)$.

Assume: the bivariate d.f. H has continuous marginals and the associated copula function C(x, y) has continuous partial derivatives.

Gaenssler and Stute (1987): weak convergence of $\sqrt{n}(C_n - C)$ to a Gaussian process in the Skorokhod space $D([0,1]^2)$.

Fermanian et al. (2004): weak convergence in $L^{\infty}([0,1]^2)$ and argue that the assumptions on the partial derivatives of C cannot be dispensed with.

References

1. Joe, H. (1997). Multivariate Models and Dependence Concepts. London: Chapman and Hall.

2. Nelsen, R. B. (2006). An Introduction to Copulas. New York: Springer.

3. U. Cherubini, E. Luciano and W. Vecchiato (2004). Copula Methods in Finance. John Wiley I& Sons, Ltd.

4. A. Patton, (2012). A review of copula models for economic time series. Journal of Multivariate Analysis.

Thank you