Copula theory and applications: Part II

Hans Colonius

Department of Psychology
Oldenburg University
hans.colonius@uni-oldenburg.de
https://uol.de/en/hans-colonius

Outline

Fréchet-Hoeffding bounds for copulas

For any copula $C(u, v)$,

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

- Proof:

Fréchet-Hoeffding bounds for copulas

For any copula $C(u, v)$,

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

- Proof: For $u, v \in [0, 1]$,

* $C(u, v) \leq C(u, 1) = u$ and $C(u, v) \leq C(1, v) = v$, so $C(u, v) \leq \min(u, v)$ (upper bound)

Fréchet-Hoeffding bounds for copulas

For any copula $C(u, v)$,

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

- Proof: For $u, v \in [0, 1]$,
  * $C(u, v) \leq C(u, 1) = u$ and $C(u, v) \leq C(1, v) = v$, so $C(u, v) \leq \min(u, v)$ (upper bound)
  * $C(1, 1) - C(1, v) - C(u, 1) + C(u, v) \geq 0$, so
    $1 - v - u + C(u, v) \geq 0$
    $\Rightarrow C(u, v) \geq v + u - 1$ (lower bound)
- $W(u, v) = \max(u + v - 1, 0)$ (max. **negative** dependence)

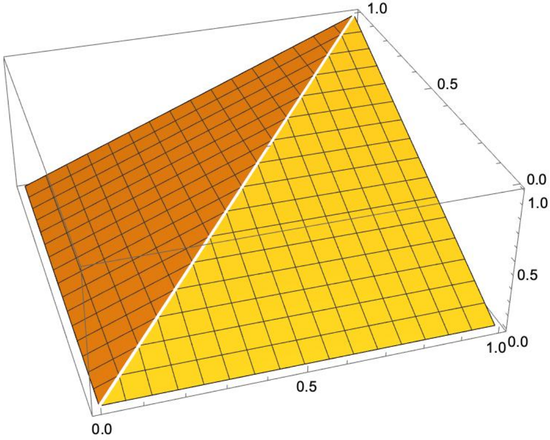
Fréchet-Hoeffding bounds for copulas

For any copula $C(u,v)$,

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

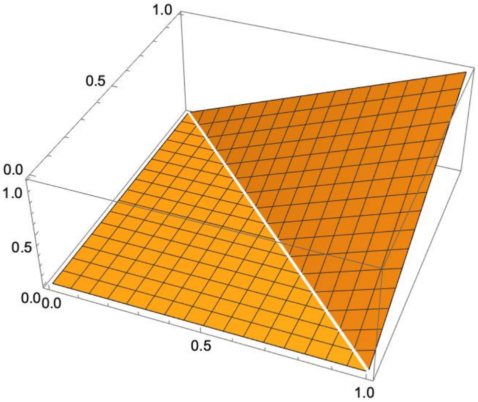
- Proof: For $u, v \in [0,1]$,
  - $C(u,v) \leq C(u,1) = u$ and $C(u,v) \leq C(1,v) = v$, so $C(u,v) \leq \min(u,v)$ (upper bound)
  - $C(1,1) - C(1,v) - C(u,1) + C(u,v) \geq 0$, so $1 - v - u + C(u,v) \geq 0$ $\Rightarrow C(u,v) \geq v + u - 1$ (lower bound)
- $W(u,v) = \max(u+v-1,0)$ (max. **negative** dependence)
- $M(u,v) = \min(u,v)$ (max. **positive** dependence)

Lower Fréchet-Hoeffding bound $W(u, v)$ (countermonotonicity)



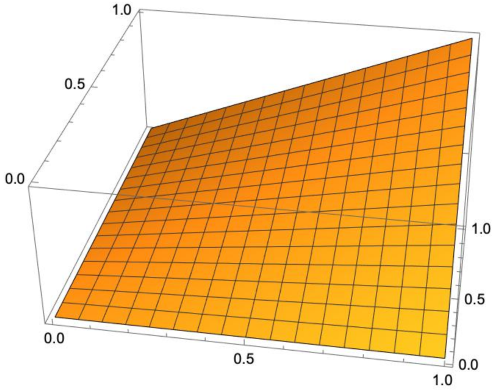
CopulaDistribution["Maximal",UniformDistribution[],UniformDistribution[]]
Plot3D[CDF,x,y]//Evaluate,x,0,1,y,0,1

Upper Fréchet-Hoeffding bound $M(u, v)$ (comonotonicity)



CopulaDistribution["Minimal",UniformDistribution[],UniformDistribution[]]
Plot3D[CDF,x,y]//Evaluate,x,0,1,y,0,1

Independence copula ($\Pi(u, v)$)



CopulaDistribution["Product", UniformDistribution[], UniformDistribution[]]
Plot3D[CDF,x,y]//Evaluate,x,0,1,y,0,1

Multivariate distribution and copula

Distinguish 2 cases:

(1) Given some copula, define a multivariate distribution by adding some margins.

(2) Given some multivariate distribution, find the margins and identify the copula;

Case (1) Gumbel' s bivariate logistic distribution

Consider copula

$$C(u, v) = \frac{uv}{u + v - uv} \tag{1}$$

Case (1) Gumbel' s bivariate logistic distribution

Consider copula

$$C(u, v) = \frac{uv}{u + v - uv} \tag{1}$$

Take *standard logistic distributions* as margins:
$u = F(x) = (1 + e^{-x})^{-1}$ and $v = G(y) = (1 + e^{-y})^{-1}$
$(-\infty < x, y < \infty$

Case (1) Gumbel' s bivariate logistic distribution

Consider copula

$$C(u, v) = \frac{uv}{u + v - uv} \tag{1}$$

Take *standard logistic distributions* as margins:
$u = F(x) = (1 + e^{-x})^{-1}$ and $v = G(y) = (1 + e^{-y})^{-1}$
$(-\infty < x, y < \infty$
Then, $x = F^{-1}(u) = \ln(1/u - 1)$ and $y = G^{-1}(v) = \ln(1/v - 1)$

Case (1) Gumbel' s bivariate logistic distribution

Consider copula

$$C(u,v) = \frac{uv}{u+v-uv} \tag{1}$$

Take *standard logistic distributions* as margins:
$u = F(x) = (1+e^{-x})^{-1}$ and $v = G(y) = (1+e^{-y})^{-1}$
$(-\infty < x, y < \infty$
Then, $x = F^{-1}(u) = \ln(1/u - 1)$ and $y = G^{-1}(v) = \ln(1/v - 1)$

$$\begin{aligned} C(u,v) &= H(F^{-1}(u), G^{-1}(v)) \\ &= H(x,y) \\ &= \frac{1}{1+e^{-x}+e^{-y}} \end{aligned} \tag{2}$$

Case (2) Bivariate extreme value distribution

$$F(x, y) = \exp[-(e^{-\delta x} + e^{-\delta y})^{1/\delta}]$$

for $-\infty < x, y < \infty, \delta \geq 1$.

Case (2) Bivariate extreme value distribution

$$F(x, y) = \exp[-(e^{-\delta x} + e^{-\delta y})^{1/\delta}]$$

for $-\infty < x, y < \infty, \delta \geq 1$.

Margins: $F_1(x) = \exp\{-e^{-x}\}$ and $F_2(y) = \exp\{-e^{-y}\}$

Case (2) Bivariate extreme value distribution

$$F(x,y) = \exp[-(e^{-\delta x} + e^{-\delta y})^{1/\delta}]$$

for $-\infty < x, y < \infty, \delta \geq 1$.

Margins: $F_1(x) = \exp\{-e^{-x}\}$ and $F_2(y) = \exp\{-e^{-y}\}$

$x = F_1^{-1}(u) = -\ln(-\ln u)$ and $y = F_2^{-1}(v) = -\ln(-\ln v)$

Case (2) Bivariate extreme value distribution

$$F(x, y) = \exp[-(e^{-\delta x} + e^{-\delta y})^{1/\delta}]$$

for $-\infty < x, y < \infty, \delta \geq 1$.

Margins: $F_1(x) = \exp\{-e^{-x}\}$ and $F_2(y) = \exp\{-e^{-y}\}$

$x = F_1^{-1}(u) = -\ln(-\ln u)$ and $y = F_2^{-1}(v) = -\ln(-\ln v)$

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

an example of the class of *bivariate extreme value* copulas characterized by $C(u^t, v^t) = C^t(u, v)$ for all $t > 0$.

Corresponding exercises from Hofert p. 25–27

```
### 2.3 Sklar's Theorem ##########################################################
### First part of Sklar's Theorem - decomposition

library(mvtnorm)
d <- 2 # dimension
rho <- 0.7 # off-diagonal entry of the correlation matrix P
P <- matrix(rho, nrow = d, ncol = d) # build the correlation matrix P
diag(P) <- 1
set.seed(64)
u <- runif(d) # generate a random evaluation point
x <- qnorm(u)
pmvnorm(upper = x, corr = P) # evaluate the copula C at u

nc <- normalCopula(rho) # normal copula (note the default dim = 2)
pCopula(u, copula = nc) # value of the copula at u

nu <- 3 # degrees of freedom
x. <- qt(u, df = nu)
pmvt(upper = x., corr = P, df = nu) # evaluate the t copula at u

try(pmvt(upper = x., corr = P, df = 3.5))

tc <- tCopula(rho, dim = d, df = nu)
pCopula(u, copula = tc) # value of the copula at u
```

Corresponding exercises from Hofert p. 25–27

```
### 2.3 Sklar's Theorem ########################################################

### Second part of Sklar's Theorem - composition

H.obj <- mvdc(claytonCopula(1), margins = c("norm", "exp"),
paramMargins = list(list(mean = 1, sd = 2), list(rate = 3)))

set.seed(1979)
z <- cbind(rnorm(5, mean = 1, sd = 2), rexp(5, rate = 3)) # evaluation points
pMvdc(z, mvdc = H.obj) # values of the df at z

dMvdc(z, mvdc = H.obj) # values of the corresponding density at z

set.seed(1975)
X <- rMvdc(1000, mvdc = H.obj)

plot(X, cex = 0.5, xlab = quote(X[1]), ylab = quote(X[2]))
contourplot2(H.obj, FUN = dMvdc, xlim = range(X[,1]), ylim = range(X[,2]),
n.grid = 257)
```

Copulas and Sklar's theorem for $d \geq 2$ variables

$$C(u_1, \ldots, u_d) = P(U_1 \leq u_1, \ldots, U_d \leq u_d), \quad u_1, \ldots, u_d \in [0, 1].$$

Copulas and Sklar's theorem for $d \geq 2$ variables

$$C(u_1, \ldots, u_d) = P(U_1 \leq u_1, \ldots, U_d \leq u_d), \quad u_1, \ldots, u_d \in [0, 1].$$

Let $F(x_1, \ldots, x_d)$ be an d-variate distribution function with margins $F_1(x_1), \ldots, F_d(x_d)$;
then there exists an d-copula $C : [0, 1]^d \longrightarrow [0, 1]$ that satisfies

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_n(x_d)), \quad (x_1, \ldots x_d) \in \mathbb{R}^d.$$

Copulas and Sklar's theorem for $d \geq 2$ variables

$$C(u_1, \ldots, u_d) = P(U_1 \leq u_1, \ldots, U_d \leq u_d), \;\; u_1, \ldots, u_d \in [0, 1].$$

Let $F(x_1, \ldots, x_d)$ be an d-variate distribution function with margins $F_1(x_1), \ldots, F_d(x_d)$;
then there exists an d-copula $C : [0, 1]^d \longrightarrow [0, 1]$ that satisfies

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_n(x_d)), \quad (x_1, \ldots x_d) \in \mathbb{R}^d.$$

If all univariate margins $F_1, \ldots, F_d$ are continuous, then the copula is unique. Otherwise, C is uniquely determined on $\mathrm{Ran} F_1 \times \mathrm{Ran} F_2 \times \ldots \mathrm{Ran} F_n$ ($\mathrm{Ran} F$ is the range/image of F).

Copulas and Sklar's theorem for $d \geq 2$ variables

$$C(u_1, \ldots, u_d) = P(U_1 \leq u_1, \ldots, U_d \leq u_d), \quad u_1, \ldots, u_d \in [0, 1].$$

Let $F(x_1, \ldots, x_d)$ be an d-variate distribution function with margins $F_1(x_1), \ldots, F_d(x_d)$;
then there exists an d-copula $C : [0, 1]^d \longrightarrow [0, 1]$ that satisfies

$$F(x_1, \ldots, x_d) = C(F_1(x_1), \ldots, F_n(x_d)), \quad (x_1, \ldots x_d) \in \mathbb{R}^d.$$

If all univariate margins $F_1, \ldots, F_d$ are continuous, then the copula is unique. Otherwise, C is uniquely determined on $\mathrm{Ran} F_1 \times \mathrm{Ran} F_2 \times \ldots \mathrm{Ran} F_n$ ($\mathrm{Ran} F$ is the range/image of F).
If $F_1^{-1}, \ldots, F_d^{-1}$ are the quantile functions of the margins, then for any $(u_1, \ldots, u_d) \in [0, 1]^d$

$$C(u_1, \ldots, u_d) = F(F_1^{-1}(u_1), \ldots, F_d^{-1}(u_d)).$$

Copula density

For an (absolutely continuous) copula C there exists a **copula density** $c : [0,1]^n \to [0,\infty]$ almost everywhere unique such that

$$C(u_1, \dots, u_n) = \int_0^{u_1} \cdots \int_0^{u_n} c(v_1, \dots, v_n)\, dv_n \dots dv_1, \quad u_1, \dots, u_n \in [0,1].$$

Such an absolutely continuous copula is n times differentiable and

$$c(u_1, \dots, u_n) = \frac{\partial}{\partial u_1} \cdots \frac{\partial}{\partial u_n} C(u_1, \dots, u_n), \quad u_1, \dots, u_n \in [0,1].$$

For example, the independence copula is absolutely continuous with density equal to 1:

$$\Pi(u_1, \dots, u_n) = \prod_{k=1}^{n} u_k = \int_0^{u_1} \cdots \int_0^{u_n} 1\, dv_n \dots dv_1$$

Copula density: bivariate case

For $d = 2$,

$$\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2} C(F_1(x_1), F_2(x_2)) = f_1(x_1)\, f_2(x_2)\, c_{12}(F_1(x_1), F_2(x_2))$$
$$= f_{12}(x_1, x_2)$$

This equation shows how independence case is "distorted" by copula density c whenever c is different from 1.

Copula density: bivariate case

For $d = 2$,

$$\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}C(F_1(x_1), F_2(x_2)) = f_1(x_1)\, f_2(x_2)\, c_{12}(F_1(x_1), F_2(x_2))$$
$$= f_{12}(x_1, x_2)$$

This equation shows how independence case is "distorted" by copula density c whenever c is different from 1.
The log-density splits into

$$\log f_{12}(x_1, x_2) = \log c_{12}(F_1(x_1), F_2(x_2)) + \log f_1(x_1) + \log f_2(x_2)$$

allowing for a two-stage estimation (marginal and copula parameters separately).

Copula density: bivariate case

For $d = 2$,

$$\frac{\partial}{\partial x_1}\frac{\partial}{\partial x_2}C(F_1(x_1), F_2(x_2)) = f_1(x_1)\ f_2(x_2)\ c_{12}(F_1(x_1), F_2(x_2))$$
$$= f_{12}(x_1, x_2)$$

This equation shows how independence case is "distorted" by copula density c whenever c is different from 1.
The log-density splits into

$$\log f_{12}(x_1, x_2) = \log c_{12}(F_1(x_1), F_2(x_2)) + \log f_1(x_1) + \log f_2(x_2)$$

allowing for a two-stage estimation (marginal and copula parameters separately).
Moreover,

$$c_{12}(u, v) = \frac{f_{12}(F_1^{-1}(u), F_2^{-1}(v))}{f_1(F_1^{-1}(u))\ f_2(F_2^{-1}(v))}$$

Simulation of copulas

Generate random variables

- Obtain an observation x of a random variable with df F:
  1. Generate a variate u that is uniform on $(0, 1)$
  2. Set $x = F^{-1}(u)$ (inverse distribution fct method)

Generate random variables

- Obtain an observation x of a random variable with df F:
  1. Generate a variate u that is uniform on $(0, 1)$
  2. Set $x = F^{-1}(u)$ (inverse distribution fct method)
- Obtain observations (x, y) of a random vector (X, Y) with distribution function H

Generate random variables

- Obtain an observation x of a random variable with df F:
  1. Generate a variate u that is uniform on $(0, 1)$
  2. Set $x = F^{-1}(u)$ (inverse distribution fct method)
- Obtain observations (x, y) of a random vector (X, Y) with distribution function H
- Special case: (X, Y) standard normal with correlation ρ
- **Exercise:** Let X and Z independent standard normal and

$$a := \frac{\sqrt{1-\rho^2}}{1-\rho} = \sqrt{\frac{1+\rho}{1-\rho}}.$$

Set $Y := \rho X + \sqrt{1-\rho^2}\, Z$.

Show: (X, Y) bivariate standard normal with correlation ρ.

Solution of exercise

$$\mathrm{E}(X\,Y) = \mathrm{E}(\rho\,X^2 + \sqrt{1-\rho^2}\,X\,Z) = \rho,$$

so

$$\operatorname{cov}(X, Y) = \mathrm{E}(X\,Y) - \mathrm{E}(X)\,\mathrm{E}(Y) = \rho,$$

so

$$\operatorname{corr}(X, Y) = \rho.$$

Simulation of bivariate copulas

- Task: Generate observations (x_1, x_2) of a pair of random variables (X_1, X_2) with joined distribution function H.

Simulation of bivariate copulas

- Task: Generate observations (x_1, x_2) of a pair of random variables (X_1, X_2) with joined distribution function H.

  Using Sklar's theorem, we need only generate a pair (u_1, u_2) of observations from uniform (U_1, U_2) with joint distribution C, the copula from X_1 and X_2 and then use transforms $x_1 = F^{-1}(u_1)$ and $x_2 = G^{-1}(u_2)$.

Simulation of bivariate copulas

- Task: Generate observations (x_1, x_2) of a pair of random variables (X_1, X_2) with joined distribution function H.

  Using Sklar's theorem, we need only generate a pair (u_1, u_2) of observations from uniform (U_1, U_2) with joint distribution C, the copula from X_1 and X_2 and then use transforms $x_1 = F^{-1}(u_1)$ and $x_2 = G^{-1}(u_2)$.
- Problem: Given C of H, how to generate (u_1, u_2)?

Simulation of bivariate copulas

- Task: Generate observations (x_1, x_2) of a pair of random variables (X_1, X_2) with joined distribution function H.

  Using Sklar's theorem, we need only generate a pair (u_1, u_2) of observations from uniform (U_1, U_2) with joint distribution C, the copula from X_1 and X_2 and then use transforms $x_1 = F^{-1}(u_1)$ and $x_2 = G^{-1}(u_2)$.
- Problem: Given C of H, how to generate (u_1, u_2)?
- There are many methods; one is the *conditional distribution method* (aka *Rosenblatt transform*).

Conditional distribution method

(see Nelsen 2006, p.36 or Mai & Scherer 2012, p.22)

$$\begin{aligned}
\frac{\partial}{\partial u_2} C(u_1, u_2) &= \frac{\partial}{\partial u_2} \int_0^{u_2} \int_0^{u_1} c(v_1, v_2) dv_1 \, dv_2 \\
&= \int_0^{u_1} c(v_1, u_2) dv_1 = \int_0^{u_1} f_{U_1|U_2=u_2}(v_1) dv_1 \\
&= P(U_1 \leq u_1 \mid U_2 = u_2) \\
&= F_{U_1|U_2=u_2}(u_1)
\end{aligned}$$

1. Simulate U_2 and fix the value u_2;
2. Compute $F_{U_1|U_2=u_2}(u_1) = \frac{\partial}{\partial u_2} C(u_1, u_2)$:
3. Compute the inverse of $F_{U_1|U_2}(u_1)$: $F^{-1}_{U_1|U_2}(v)$:
4. Simulate uniform V independent U_2
5. Set $U_1 = F^{-1}_{U_1|U_2}(V)$ and return $(U_1, U_2) \sim C$.

Simulation of upper Fréchet-Hoeffding bound

$M(u_1, u_2) = \min(u_1, u_2)$; with fixed $u_2 \in (0, 1)$

$$M(u_1, u_2) = \begin{cases} u_2, \text{if } u_2 < u_1 \\ u_1, \text{if } u_2 > u_1 \end{cases}$$

and

$$F_{U_1|U_2=u_2}(u_1) = \frac{\partial}{\partial u_2} M(u_1, u_2) = \begin{cases} 1, \text{if } u_2 < u_1 \\ 0, \text{if } u_2 > u_1 \end{cases}$$

for $u_1 \in [0, 1]$, $u_1 \neq u_2$.

Simulation of upper Fréchet-Hoeffding bound

$M(u_1, u_2) = \min(u_1, u_2)$; with fixed $u_2 \in (0, 1)$

$$M(u_1, u_2) = \begin{cases} u_2, \text{if } u_2 < u_1 \\ u_1, \text{if } u_2 > u_1 \end{cases}$$

and

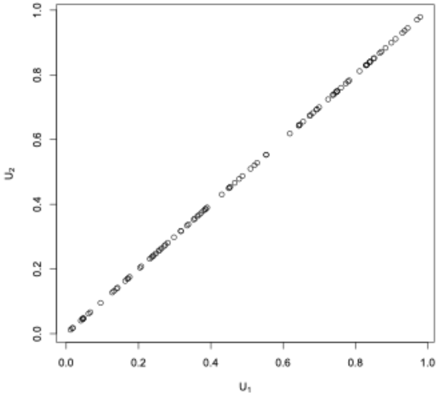
$$F_{U_1|U_2=u_2}(u_1) = \frac{\partial}{\partial u_2} M(u_1, u_2) = \begin{cases} 1, \text{if } u_2 < u_1 \\ 0, \text{if } u_2 > u_1 \end{cases}$$

for $u_1 \in [0, 1]$, $u_1 \neq u_2$. Note that $F_{U_1|U_2=u_2}(u_1)$ is not defined at point u_2, we set it equal to 1. The inverse is

$$F^{-1}_{U_1|U_2=u_2}(v) = u_2, \quad v \in (0, 1).$$

Thus, the algorithm implies simulating U_2 and then setting $F^{-1}_{U_1|U_2=u_2}(V) = U_2$.

Upper Fréchet-Hoeffding bound $M(u_1, u_2)$ (comonotonicity)




```
 set.seed(1980)\[  \]
U <- runif(100)
plot(cbind(U, U),   xlab = quote(U[1]), ylab = quote(U[2]))
#plot(cbind(U, 1-U), xlab = quote(U[1]), ylab = quote(U[2]))
```

Archimedean copulas

Let

$$C(u, v) = \psi^{-1}(\psi(u) + \psi(v))$$

- What are necessary and sufficient conditions for $C(u, v)$ to be a copula?

Archimedean copulas

Let

$$C(u,v) = \psi^{-1}(\psi(u) + \psi(v))$$

- What are necessary and sufficient conditions for $C(u,v)$ to be a copula?

1. $\psi : [0,1] \mapsto [0,\infty)$ strictly decreasing and continuous
2. $\psi(0) = \infty$ and $\psi(1) = 0$
3. ψ is convex.

ψ is called *(strict) generator* of $C(u,v)$.

Archimedean copulas

Let

$$C(u,v) = \psi^{-1}(\psi(u) + \psi(v))$$

- What are necessary and sufficient conditions for $C(u,v)$ to be a copula?

1. $\psi : [0,1] \mapsto [0,\infty)$ strictly decreasing and continuous
2. $\psi(0) = \infty$ and $\psi(1) = 0$
3. ψ is convex.

ψ is called *(strict) generator* of $C(u,v)$.
Is the generator of a copula uniquely determined?

Archimedean copulas

Let

$$C(u, v) = \psi^{-1}(\psi(u) + \psi(v))$$

- What are necessary and sufficient conditions for $C(u, v)$ to be a copula?

1. $\psi : [0, 1] \mapsto [0, \infty)$ strictly decreasing and continuous
2. $\psi(0) = \infty$ and $\psi(1) = 0$
3. ψ is convex.

ψ is called *(strict) generator* of $C(u, v)$.
Is the generator of a copula uniquely determined?
Show: Archimedean copulas are (1) *symmetric* (2) *associative* !

Example of Archimedean copula: Gumbel-Hougaard

(aka “bivariate extreme value”)

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

  has generator

  $$\psi(t) = (-\ln t)^\delta.$$

Measures of dependence

- Measures of association/dependence are scalar measures which summarize the dependence in terms of a single number.
- Linear measures of dependence, like (Neyman-Pearson) correlation and covariance depend on both the marginals and the copula.
- Task: find measures of dependence for copulas

Measures of dependence

- Measures of association/dependence are scalar measures which summarize the dependence in terms of a single number.
- Linear measures of dependence, like (Neyman-Pearson) correlation and covariance depend on both the marginals and the copula.
- Task: find measures of dependence for copulas
- Let $(X_1, X_2) \sim F$ with margins F_1, F_2 and copula C; if T_i $(i = 1, 2)$ are strictly increasing transformations of $X_i, (i = 1, 2)$, then $(T_1(X_1), T_2(X_2))$ has the same copula C. (*Invariance principle*)

Measures of dependence

- Measures of association/dependence are scalar measures which summarize the dependence in terms of a single number.
- Linear measures of dependence, like (Neyman-Pearson) correlation and covariance depend on both the marginals and the copula.
- Task: find measures of dependence for copulas
- Let $(X_1, X_2) \sim F$ with margins F_1, F_2 and copula C; if T_i $(i = 1, 2)$ are strictly increasing transformations of $X_i, (i = 1, 2)$, then $(T_1(X_1), T_2(X_2))$ has the same copula C. (*Invariance principle*)
- Thus, one can study dependence independently of the margins via $(U_1, U_2) = (F_1(X_1), F_2(X_2))$ instead of (X_1, X_2).

Linear dependence: Hoeffding's formula (aka, lemma)

$X_i \sim F_i,\ i = 1, 2$ random variables with $\mathrm{E}(X_i^2) < \infty$ and joint distribution function F. Then

$$\begin{aligned}\mathrm{Cov}(X_1, X_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(F(x_1, x_2) - F_1(x_1)\,F_2(x_2))dx_1\,dx_2 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(C(F_1(x_1), F_2(x_2)) - F_1(x_1)\,F_2(x_2))dx_1\,dx_2\end{aligned}$$

Correlation fallacies:

1. F_1, F_2 and correlation ρ uniquely determine F.
2. Uncorrelatedness implies independence.
3. Given F_1, F_2, any level of correlation $-1 < \rho < +1$ can be attained.

Definition: Kendall's tau (τ) (population version)

Let F be a continuous bivariate distribution function and let $(X_1, X_2), (X_1', X_2')$ be independent pairs with distribution F. *Kendall's tau* equals the probability of concordant pairs minus the probability of discordant pairs, i.e.,

$$\tau = P[(X_1 - X_1')(X_2 - X_2') > 0] - P[(X_1 - X_1')(X_2 - X_2') < 0]$$

Proposition

If F has copula C, then

$$\tau = 4 \int_{[0,1]^2} C(u_1, u_2)\, dC(u_1, u_2) - 1.$$

Definition: Kendall's tau (τ) (population version)

Let F be a continuous bivariate distribution function and let $(X_1, X_2), (X_1', X_2')$ be independent pairs with distribution F. *Kendall's tau* equals the probability of concordant pairs minus the probability of discordant pairs, i.e.,

$$\tau = P[(X_1 - X_1')(X_2 - X_2') > 0] - P[(X_1 - X_1')(X_2 - X_2') < 0]$$

Proposition

If F has copula C, then

$$\tau = 4 \int_{[0,1]^2} C(u_1, u_2)\, dC(u_1, u_2) - 1.$$

Interpretation of tau as expected value:

$$\tau = 4\,\mathrm{E}[C(u_1, u_2)] - 1$$

Equivalent version of Kendall's tau (τ)

Instead of

$$\tau = 4 \int_{[0,1]^2} C(u_1, u_2)\, dC(u_1, u_2) - 1,$$

compute

$$\tau = 1 - 4 \int_{[0,1]^2} \frac{\partial}{du_1} C(u_1, u_2) \frac{\partial}{du_2} C(u_1, u_2) du_1 du_2,$$

which is often more tractable.

Example: Kendall's tau for FGM copula

- Farlie-Gumbel-Morgenstern copula:

$$C_\theta(u_1, u_2) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad \theta \in [-1, 1].$$

Example: Kendall's tau for FGM copula

- Farlie-Gumbel-Morgenstern copula:

$$C_\theta(u_1, u_2) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad \theta \in [-1, 1].$$

-

$$\begin{aligned} dC_\theta(u_1, u_2) &= \frac{\partial^2 C_\theta(u_1, u_2)}{\partial u_1 \partial u_2} du_1 du_2 \\ &= [1 + \theta(1 - 2u_1)(1 - 2u_2)] du_1 du_2 \end{aligned}$$

Example: Kendall's tau for FGM copula

- Farlie-Gumbel-Morgenstern copula:

$$C_\theta(u_1, u_2) = u_1 u_2 + \theta u_1 u_2 (1 - u_1)(1 - u_2), \quad \theta \in [-1, 1].$$

- $$\begin{aligned} dC_\theta(u_1, u_2) &= \frac{\partial^2 C_\theta(u_1, u_2)}{\partial u_1 \partial u_2} \, du_1 du_2 \\ &= [1 + \theta(1 - 2u_1)(1 - 2u_2)] du_1 du_2 \end{aligned}$$

- $$\int_{[0,1]^2} C(u_1, u_2) \, dC(u_1, u_2) = \frac{1}{4} + \frac{\theta}{18}$$

$$\tau = \frac{2\theta}{9}, \quad \text{thus} \quad -2/9 < \tau < 2/9.$$

Kendall's tau for Archimedean copulas

Proposition

For an Archimedean copula C generated by $\psi(t)$,

$$\tau = 1 + 4 \int_0^1 \frac{\psi(t)}{\psi'(t)}\, dt.$$

Kendall's tau for Archimedean copulas

Proposition

For an Archimedean copula C generated by $\psi(t)$,

$$\tau = 1 + 4\int_0^1 \frac{\psi(t)}{\psi'(t)}\,dt.$$

- No need to compute a double integral !
- Example: Gumbel-Hougaard with $(-\ln t)^{\delta}$:

$$\frac{\psi(t)}{\psi'(t)} = \frac{t\ln t}{\delta},$$

then

$$\tau_{\delta} = \frac{\delta - 1}{\delta}$$

Survival copula

- Survival distribution of random variable X: $\overline{F}(x) = P(X > x)$

Survival copula

- Survival distribution of random variable X: $\overline{F}(x) = P(X > x)$
- Bivariate survival distribution of (X, Y) with copula C:

$$\begin{aligned}\overline{H}(x,y) &= P(X > x, Y > y) = 1 - F(x) - G(y) + H(x,y)\\ &= \overline{F}(x) + \overline{G}(y) - 1 + C(1 - \overline{F}(x), 1 - \overline{G}(y))\end{aligned}$$

Survival copula

- Survival distribution of random variable X: $\overline{F}(x) = P(X > x)$
- Bivariate survival distribution of (X, Y) with copula C:

$$\begin{aligned}\overline{H}(x, y) = P(X > x, Y > y) &= 1 - F(x) - G(y) + H(x, y) \\ &= \overline{F}(x) + \overline{G}(y) - 1 + C(1 - \overline{F}(x), 1 - \overline{G}(y))\end{aligned}$$

- Margins of $\overline{H}(x, y)$: $\overline{H}(x, -\infty) = \overline{F}(x)$, $\overline{H}(-\infty, y) = \overline{G}(y)$

Survival copula

- Survival distribution of random variable X: $\overline{F}(x) = P(X > x)$
- Bivariate survival distribution of (X, Y) with copula C:

$$\begin{aligned}\overline{H}(x,y) &= P(X > x, Y > y) = 1 - F(x) - G(y) + H(x,y) \\ &= \overline{F}(x) + \overline{G}(y) - 1 + C(1 - \overline{F}(x), 1 - \overline{G}(y))\end{aligned}$$

- Margins of $\overline{H}(x,y)$: $\overline{H}(x,-\infty) = \overline{F}(x)$, $\overline{H}(-\infty,y) = \overline{G}(y)$
- Define *survival copula*:

$$\widehat{C}(u,v) = u + v - 1 + C(1-u, 1-v),$$

so

$$\overline{H}(x,y) = \widehat{C}(\overline{F}(x, \overline{G}(y))$$

Survival copula

- Survival distribution of random variable X: $\overline{F}(x) = P(X > x)$
- Bivariate survival distribution of (X, Y) with copula C:

$$\begin{aligned}\overline{H}(x, y) = P(X > x, Y > y) &= 1 - F(x) - G(y) + H(x, y) \\ &= \overline{F}(x) + \overline{G}(y) - 1 + C(1 - \overline{F}(x), 1 - \overline{G}(y))\end{aligned}$$

- Margins of $\overline{H}(x, y)$: $\overline{H}(x, -\infty) = \overline{F}(x)$, $\overline{H}(-\infty, y) = \overline{G}(y)$
- Define *survival copula*:

$$\widehat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v),$$

so

$$\overline{H}(x, y) = \widehat{C}(\overline{F}(x, \overline{G}(y))$$

- Extension to $d \geq 2$ possible.

Copula and order statistics

Diagonal section of a copula

$U_1, \ldots, U_d$ uniform $[0,1]$ random variables with joint distribution function C. For any $t \in [0,1]$

$$P(\max\{U_1, \ldots, U_d\} \leq t) = C(t, t, \ldots, t).$$

Diagonal section of a copula

$U_1, \ldots, U_d$ uniform $[0,1]$ random variables with joint distribution function C. For any $t \in [0,1]$

$$P(\max\{U_1, \ldots, U_d\} \leq t) = C(t, t, \ldots, t).$$

$\delta_C(t) = C(t, t, \ldots, t)$ is called the *diagonal section* of copula C.

- $d = 2$ with (U, V):

$$\begin{aligned} P(\min(U, V) \leq t) &= P(U \leq t) + P(V \leq t) \\ &\qquad - P(\{U \leq t\} \cap \{V \leq t\}) \\ &= 2t - \delta_C(t). \end{aligned}$$

Diagonal section of a copula

$U_1, \dots, U_d$ uniform $[0,1]$ random variables with joint distribution function C. For any $t \in [0,1]$

$$P(\max\{U_1, \dots, U_d\} \le t) = C(t, t, \dots, t).$$

$\delta_C(t) = C(t, t, \dots, t)$ is called the *diagonal section* of copula C.

- $d = 2$ with (U, V):

$$\begin{aligned} P(\min(U, V) \le t) &= P(U \le t) + P(V \le t) \\ &\qquad - P(\{U \le t\} \cap \{V \le t\}) \\ &= 2t - \delta_C(t). \end{aligned}$$

- $$\max(2t-1), 0) \le \delta_C(t) \le t$$

for any copula C and all $t \in [0,1]$

Diagonal section of a copula

Definition

A function $\delta : [0,1] \mapsto [0,1]$ is called *diagonal* if

1. $\delta(1) = 1$
2. $0 \leq \delta(t_2) - \delta(t_1) \leq 2(t_2 - t_1)$ for all $t_1, t_2 \in [0,1]$ with $t_1 \leq t_2$
3. $\delta(t) \leq t$ for all $t \in [0,1]$.

Proposition

Let δ be any diagonal and set

$$C(u,v) = \min\{u, v, 1/2[\delta(u) + \delta(v)]\}.$$

Then C is a copula with diagonal section δ (Nelsen-Fredricks copula).

Diagonal section of a copula

Definition

A function $\delta : [0,1] \mapsto [0,1]$ is called *diagonal* if

1. $\delta(1) = 1$
2. $0 \leq \delta(t_2) - \delta(t_1) \leq 2(t_2 - t_1)$ for all $t_1, t_2 \in [0,1]$ with $t_1 \leq t_2$
3. $\delta(t) \leq t$ for all $t \in [0,1]$.

Proposition

Let δ be any diagonal and set

$$C(u,v) = \min\{u, v, 1/2[\delta(u) + \delta(v)]\}.$$

Then C is a copula with diagonal section δ (Nelsen-Fredricks copula).

C is not **unique**!

Copula of (extreme) order statistics

Random sample $X_1, \ldots, X_n$ independent identically distributed random variables with continuous distribution function F.

- What is the copula of $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$?

Copula of (extreme) order statistics

Random sample $X_1, \ldots, X_n$ independent identically distributed random variables with continuous distribution function F.

- What is the copula of $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$? Then $F_1(x) = P(X_{(1)} \leq x) = 1 - (1 - F(x)^n)$ and $F_n(y) = P(X_{(n)} \leq y) = F^n(y)$.

Copula of (extreme) order statistics

Random sample $X_1, \ldots, X_n$ independent identically distributed random variables with continuous distribution function F.

- What is the copula of $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$? Then $F_1(x) = P(X_{(1)} \leq x) = 1 - (1 - F(x)^n)$ and $F_n(y) = P(X_{(n)} \leq y) = F^n(y)$.
- Then,

$$F_{1,n}(x, y) = P(\{X_{(1)} \leq x\} \cap \{X_{(n)} \leq y\}) = \begin{cases} F^n(y) - (F(y) - F(x))^n & \text{if } x < y \\ F^n(y) & \text{if } x \geq y. \end{cases}$$

Copula of (extreme) order statistics

Random sample $X_1, \ldots, X_n$ independent identically distributed random variables with continuous distribution function F.

- What is the copula of $X_{(1)} = \min\{X_1, \ldots, X_n\}$ and $X_{(n)} = \max\{X_1, \ldots, X_n\}$?
  Then $F_1(x) = P(X_{(1)} \leq x) = 1 - (1 - F(x)^n)$ and $F_n(y) = P(X_{(n)} \leq y) = F^n(y)$.
- Then,

$$F_{1,n}(x, y) = P(\{X_{(1)} \leq x\} \cap \{X_{(n)} \leq y\}) \\ = \begin{cases} F^n(y) - (F(y) - F(x))^n & \text{if } x < y \\ F^n(y) & \text{if } x \geq y. \end{cases}$$

- Find copula $C(F_1(x), F_n(y)) = F_{1,n}(x, y)$ by setting

$$C(u, v) = F_{1,n}(F_1^{-1}(u), F_n^{-1}(v)) \;\; u, v \in [0, 1].$$

Copula of (extreme) order statistics

- With $F_1^{-1}(u) = F^{-1}(1-(1-u)^{1/n})$ and $F_n^{-1}(v) = F^{-1}(v^{1/n})$

$$C_n(u,v) = \begin{cases} v - (v^{1/n} + (1-u)^{1/n} - 1)^n, & \text{if } 1-(1-u)^{1/n} < v^{1/n}, \\ v & \text{if } 1-(1-u)^{1/n} \geq v^{1/n}, \end{cases}$$

a copula describing the dependence structure of the minimum and maximum of n independent random variables.

Copula of (extreme) order statistics

- With $F_1^{-1}(u) = F^{-1}(1-(1-u)^{1/n})$ and $F_n^{-1}(v) = F^{-1}(v^{1/n})$

$$C_n(u,v) = \begin{cases} v - (v^{1/n} + (1-u)^{1/n} - 1)^n, & \text{if } 1-(1-u)^{1/n} < v^{1/n}, \\ v & \text{if } 1-(1-u)^{1/n} \geq v^{1/n}, \end{cases}$$

a copula describing the dependence structure of the minimum and maximum of n independent random variables.
What happens for $n \Rightarrow \infty$?

Copula of (extreme) order statistics

- With $F_1^{-1}(u) = F^{-1}(1-(1-u)^{1/n})$ and $F_n^{-1}(v) = F^{-1}(v^{1/n})$

$$C_n(u,v) = \begin{cases} v - (v^{1/n} + (1-u)^{1/n} - 1)^n, & \text{if } 1-(1-u)^{1/n} < v^{1/n}, \\ v & \text{if } 1-(1-u)^{1/n} \geq v^{1/n}, \end{cases}$$

  a copula describing the dependence structure of the minimum and maximum of n independent random variables.
  What happens for $n \Rightarrow \infty$?

- Kendall's tau

$$\tau_n(X_{(1)}, X_{(n)}) = \frac{1}{2n-1}$$

Tail dependence

Vine copulas