A short introduction to copulas

Ivan Belohlavek

December 14, 2021

Consider a distribution function $H(x, y)$ with margins $F(x) = H(x, +\infty)$ and $G(y) = H(+\infty, y)$.

- We know that if H is the joint distribution function of a pair of independent random variables, then

$$H(x, y) = F(x) \cdot G(y).$$

- If F and G are injective, then there is clearly a well-defined *coupling function* $\wedge$ such that

$$H(x, y) = F(x) \wedge G(y).$$

- In fact, this is true without the injectivity assumption.

Theorem

Let $H(x,y)$ be a distribution function with margins $F(x)$ and $G(y)$ as before. Then

$$|H(x_1,y_1)-H(x_2,y_2)| \leq |F(x_1)-F(x_2)| + |G(y_1)-G(y_2)|.$$

- The value $H(x,y)$ thus depends only on $F(x)$ and $G(y)$.

- The value $u \wedge v$ is defined for u in $\mathrm{Img}(F)$ and v in $\mathrm{Img}(G)$.
- $u \wedge 0 = 0 \wedge v = 0$
- $u \wedge 1 = u$
- $1 \wedge v = v$
- The $\wedge$-area of any rectangle $[u_1, v_1] \times [u_2, v_2]$ with $u_1 \leq u_2$ and $v_1 \leq v_2$ is non-negative:

$$(u_2 \wedge v_2) - (u_1 \wedge v_2) - (u_2 \wedge v_1) + (u_1 \wedge v_1) \geq 0.$$

Definition

Let U and V be subsets of $[0, 1]$ containing 0 and 1. A function $C : U \times V \to [0, 1]$ is called a *subcopula* if

- $C(u, 0) = C(0, v) = 0$
- $C(u, 1) = u$
- $C(1, v) = v$
- For all $u_1 \le u_2$ and $v_1 \le v_2$,

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \ge 0.$$

If $U = V = [0, 1]$ then C is called a *copula*.

Sklar's Theorem

Theorem

Let $H(x, y)$ be a distribution with margins $F(x)$ and $G(y)$. Then there exists a unique subcopula C defined on Img$(F) \times$ Img(G) such that

$$H(x, y) = C(F(x), G(y)).$$

Theorem

Let $F(x)$ and $G(y)$ be distribution functions, and C be a subcopula defined on Img$(F) \times$ Img(G). Then the function

$$H(x, y) = C(F(x), G(y))$$

is a distribution function whose margins are F and G.

- The function $\Pi(u, v) = u \cdot v$ is a copula.

Theorem

Let $H(x, y)$ be a distribution function with margins $F(x)$ and $G(y)$. Then the function

$$C(u, v) = H(F^{-1}(u), G^{-1}(v))$$

defined for u in Img(F) and v in Img(G) is a subcopula.

We can use this theorem to construct e.g.:

- Normal copulas
- T copulas

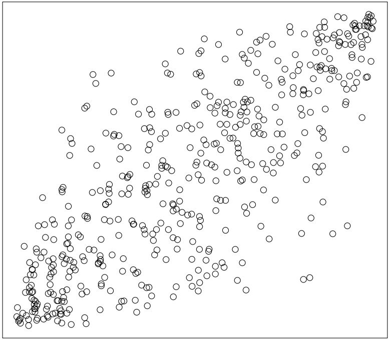

Figure: Normal copula with $\rho = 0.8$

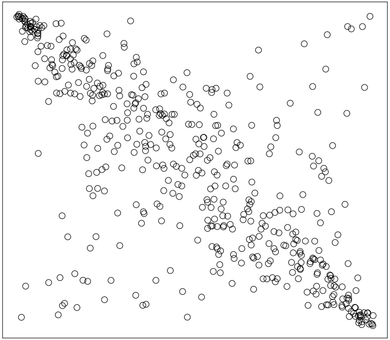

Figure: T copula with $\rho = -0.8$ and 1 degree of freedom

- Every subcopula can be extended to a copula.
- Every copula C is Lipshitz:

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

- Every copula is uniformly continuous on $[0, 1]^2$.
- Every copula is differentiable almost everywhere on $[0, 1]^2$.
- The partial derivatives of any copula C satisfy

$$0 \leq \frac{\partial C}{\partial u}(u, v) \leq 1 \quad \text{and} \quad 0 \leq \frac{\partial C}{\partial v}(u, v) \leq 1$$

  for all u, v in $[0, 1]$ such that the partial derivative exists.
- For any u in $[0, 1]$, the function $v \mapsto \frac{\partial C}{\partial u}(u, v)$ is defined almost everywhere in $[0, 1]$ and non-decreasing on its domain.

- The function $M(u, v) = \min\{u, v\}$ is a copula.
- The function $W(u, v) = \max\{0, u + v - 1\}$ is a copula.

Theorem (Fréchet-Hoeffding bounds)

Every copula C satisfies $M \geq C \geq W$.

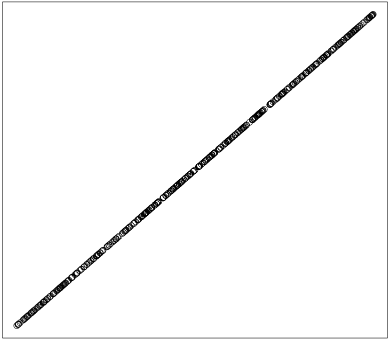

Figure: The copula M

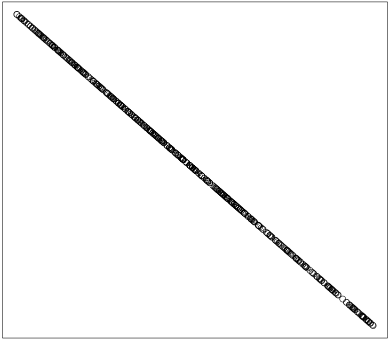

Figure: The copula W

Theorem

Let X and Y be non-atomic random variables.

The copula of X and Y is M if and only if there existst a non-decreasing function $T(x)$ such that $T(X) = Y$ almost surely.

The copula of X and Y is W if and only if there existst a non-increasing function $T(x)$ such that $T(X) = Y$ almost surely.

- Any copula induces a probability measure on $[0,1]^2$.
- This measure has uniform margins and is thus non-atomic.
- Thus it decomposes uniquely into an absolutely continuous and singular part.

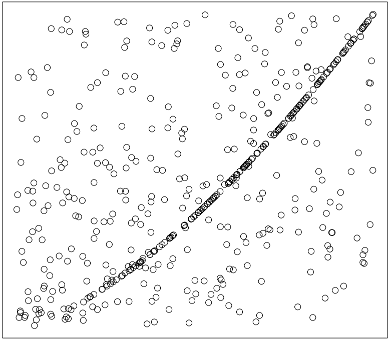

Figure: A copula with both an absolutely continuous and singular part

Theorem (Invariance principle)

Let X and Y be non-atomic random variables and $S(x)$ and $T(y)$ be strictly increasing functions. Then $S(X)$ and $T(Y)$ are also non-atomic and have the same copula as X and Y.

Let $C(u, v)$ be the copula of X and Y, and $C'(u, v)$ the copula of $S(X)$ and $T(Y)$.

- If S is strictly increasing and T strictly decreasing, then

$$C'(u, v) = u - C(u, 1 - v).$$

- If S and T are both strictly decreasing, then

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

Theorem

Let $(X_1, Y_1), \ldots (X_n, Y_n)$ *be a random sample, and* $C(u, v)$ *be the copula of* X_i *and* Y_i. *Then the copula of* $X_{(n)}$ *and* $Y_{(n)}$ *is*

$$C'(u, v) = \left(C(u^{1/n}, v^{1/n}) \right)^n .$$

- It is also possible to express copulas of other order statistics in terms of C.

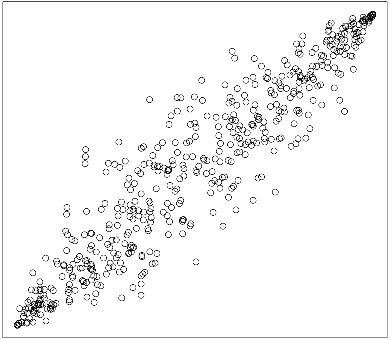

Figure: Normal copula with $\rho = 0.95$

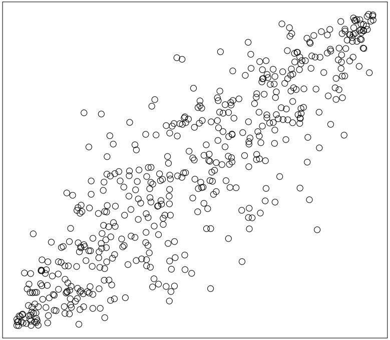

Figure: Transformed copula with $n = 10$

Dependence

Definition

Let X and Y be random variables with finite non-zero second moments. We define the *correlation* of X and Y by

$$\operatorname{Cor}(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}} = \frac{\mathrm{E}(X-\mathrm{E}X)(Y-\mathrm{E}Y)}{\sqrt{\operatorname{Var}(X)}\sqrt{\operatorname{Var}(Y)}}.$$

- $\operatorname{Cor}(X,Y) = 1$ if and only if $X = a + bY$ almost surely for some $a \in \mathbf{R}$ and $b > 0$.
- For $a \in \mathbf{R}$ and $b > 0$, $\operatorname{Cor}(X,Y) = \operatorname{Cor}(a + bX, Y)$.
- Correlation is not in general invariant under strictly increasing transformations.

Theorem

Let X and Y be non-atomic random variables with finite second moments and distribution functions $F(x)$ and $G(y)$. Let $C(u,v)$ be their copula. Then

$$\mathrm{Cov}(X,Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} C(F(x),G(y)) - F(x)G(y)\,dx\,dy.$$

- The correlation of X and Y is maximized when their copula is M, and minimized when it is W.

Let $\log(X) \sim \mathcal{N}(0,1)$ and $\log(Y) \sim \mathcal{N}(0,\sigma^2)$. Then

$$\frac{e^{-\sigma}-1}{\sqrt{e-1}\sqrt{e^{\sigma^2}-1}} \leq \mathrm{Cor}(X,Y) \leq \frac{e^{\sigma}-1}{\sqrt{e-1}\sqrt{e^{\sigma^2}-1}}.$$

Definition

We say that the pairs (x_1, y_1) and (x_2, y_2) are *concordant* if either

- $x_1 > x_2$ and $y_1 > y_2$, or
- $x_1 < x_2$ and $y_1 < y_2$.

Equivalently, (x_1, y_1) and (x_2, y_2) are concordant if

$$(x_1 - x_2)(y_1 - y_2) > 0.$$

We say (x_1, y_1) and (x_2, y_2) are *discordant* if either

- $x_1 < x_2$ and $y_1 > y_2$, or
- $x_1 > x_2$ and $y_1 < y_2$.

Equivalently, (x_1, y_1) and (x_2, y_2) are discordant if

$$(x_1 - x_2)(y_1 - y_2) < 0.$$

Definition

Let (X_1, Y_1) and (X_2, Y_2) be independent pairs of non-atomic random variables, such that their distribution functions $H_1(x, y)$ and $H_2(x, y)$ have common margins $F(x)$ and $G(y)$. We define the *concordance* of (X_1, Y_1) and (X_2, Y_2) by

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

Theorem

The concordance of (X_1, Y_1) and (X_2, Y_2) depends only on their copulas $C_1(u, v)$ and $C_2(u, v)$, and is equal to

$$Q(C_1, C_2) = 4 \int_{[0,1]^2} C_1(u, v) dC_2(u, v) - 1.$$

Theorem

The function $Q(C_1, C_2)$ is symmetric and non-decreasing in each of its arguments.

Recall the copulas following copulas:

- $M(u, v) = \min\{u, v\}$
- $W(u, v) = \max\{0, u + v - 1\}$
- $\Pi(u, v) = uv$

Their concordances are summarized in the following table.

Q	M	Π	W
M	1	$1/3$	0
Π	$1/3$	0	$-1/3$
W	0	$-1/3$	-1

Table: Concordances

Kendall's τ

Definition

Let X and Y be non-atomic random variables with copula $C(u, v)$. We define

$$\tau(X, Y) = \tau(C) = Q(C, C) = 4 \int_{[0,1]^2} C(u, v) dC(u, v) - 1.$$

Let (X_1, Y_1) and (X_2, Y_2) be independent copies of (X, Y). Then

$$\tau(X, Y) = P\Big[(X_1, Y_1) \text{ and } (X_2, Y_2) \text{ are concordant}\Big] - P\Big[(X_1, Y_1) \text{ and } (X_2, Y_2) \text{ are discordant}\Big].$$

Spearman's ρ

Definition

Let X and Y be non-atomic random variables with copula $C(u,v)$. We define

$$\rho(X,Y) = \rho(C) = 3Q(C,\Pi) = 12\int_{[0,1]^2} C(u,v)\,du\,dv - 3.$$

Let $F(x)$ and $G(y)$ be the distribution functions of X and Y. Then

$$\rho(X,Y) = \mathrm{Cor}(F(X), G(Y)).$$

References

- Nelsen, R. B.: An Introduction to Copulas, Second Edition, Springer, 2007
- Embrechts, P., Frey, R., & McNeil, A. Quantitative risk management. Princeton Series in Finance
- Embrechts, P., Hofert, M. A note on generalized inverses. Math Meth Oper Res 77, 423–432 (2013).