1. Copula

(1) Definition

The copula represents the true interdependence structure of a random variable.

Intuitively, the copula is a standardized version of the purely joint features of a multivariate distribution, which is obtained by filtering out all the purely one-dimensional features (the marginal distribution) of each entry X_n.

(2) How to calculate the copula

We can define a new random variable, called the grade of X:

$$U \equiv F_X(X)$$

The grade of X is a deterministic transformation of the random variable X that assumes values in the interval $[0,1]$. We can prove that the grade is uniformly distributed on this interval[1]:

$$U \sim \mathrm{U}([0,1])$$

We can standardize each marginal component X_n of $\boldsymbol{X}$ by means of the uniform distribution. Consider a new random variable $\boldsymbol{U}$, which is the vector of the grades:

$$\mathbf{U} \equiv \begin{pmatrix} U_1 \\ \vdots \\ U_N \end{pmatrix} \equiv \begin{pmatrix} F_{X_1}(X_1) \\ \vdots \\ F_{X_N}(X_N) \end{pmatrix}$$

Each entry of $\boldsymbol{U}$ has the same distribution (uniform distribution on $[0,1]$). Hence, through transforming $\boldsymbol{X}$ into $\boldsymbol{U}$, we factor out the marginal components of $\boldsymbol{X}$.

The copula of the multivariate random variable $\boldsymbol{X}$ is the joint distribution of its grades $\boldsymbol{U}$. The pdf of the copula reads[2]:

$$f_{\mathbf{U}}(u_1, \ldots, u_N) = \frac{f_{\mathbf{X}}(Q_{X_1}(u_1), \ldots, Q_{X_N}(u_N))}{f_{X_1}(Q_{X_1}(u_1)) \cdots f_{X_N}(Q_{X_N}(u_N))}$$

where Q_{X_N} is the quantile of the generic n-th marginal entry of $\boldsymbol{X}$.

The cdf of the copula of $\boldsymbol{X}$ reads:

$$F_{\mathbf{U}}(u_1, \ldots, u_N) = F_{\mathbf{X}}(Q_{X_1}(u_1), \ldots, Q_{X_N}(u_N))$$

Since the marginal distribution of the generic n-th entry is uniform, we obtain:

---

[1] $\because\ U = F_X(X)$

$\therefore\ F_U(u) = P(U \le u) = P(F_X(X) \le u)$

$\because\ F_X(x)$ is non − decreasing

$\therefore\ P(F_X(X) \le u) = P[F_X^{-1}(F_X(X)) \le F_X^{-1}(u)] = P[X \le F_X^{-1}(u)] = F_X[F_X^{-1}(u)] = u$

$\therefore\ F_U(u) = u$, which is the CDF of uniform distribution

$\because\ F_X(X) \in [0,1]$

$\therefore\ U$ is uniformly distributed on [0,1]

Therefore, $U := F_X(X) \sim U[0,1]$

[2] $\because\ Q_X(U) = F_X^{-1}(U) = F_X^{-1}(F_X(X)) = X$

$\therefore X \sim Q_X(U)$

$$F_{\mathbf{U}}(1, \ldots, u_n, \ldots, 1) = u_n$$

(3) Properties

Decomposition formula: the joint pdf of a generic variable $\boldsymbol{X}$ is the product of the pdf of its copula and the pdf of the marginal densities of its entries.

$$f_{\mathbf{X}}(x_1, \ldots, x_N) = f_{\mathbf{U}}(F_{X_1}(x_1), \ldots, F_{X_N}(x_N)) \prod_{n=1}^{N} f_{X_n}(x_n)$$

Given the copula of $\boldsymbol{X}$, i.e., the distribution of the grades $\boldsymbol{U}$, we can reconstruct the distribution of $\boldsymbol{X}$ with a deterministic transformation of each grade separately:

$$\mathbf{X} \stackrel{d}{=} \left(\begin{array}{c} Q_{X_1}(U_1) \\ \vdots \\ Q_{X_N}(U_N). \end{array} \right)$$

The copulas of co-monotonic variables are equal:

$$(\mathbf{X}, \mathbf{Y}) \text{ co-monotonic } \Leftrightarrow \text{ copula of } \mathbf{X} = \text{ copula of } \mathbf{Y}$$