Copula: An Introduction

March 29, 2018

An n-dimensional copula $C(u_1, \dots, u_n)$ is a joint cumulative distribution function in the unit hypercube $[0,1]^n$ with uniform margins.

Example

Let $n = 2$, a copula function $C(u_1, u_2)$ can be written as

$$C(u_1, u_2) \equiv \mathrm{P}\left(U_1 \leqslant u_1, U_2 \leqslant u_2\right),$$

where the random variables U_1, U_2 are uniformly distributed on $[0, 1]$.

Given a random variable X, the cumulative distribution function (abbreviated as cdf) F_X of X is defined as

$$F_X(x) \equiv \mathrm{P}(X \leqslant x).$$

The quantile function (generalized inverse) function F_X^{-1} of F_X is defined as

$$F_X^{-1}(u) = \inf\{x | F_X(x) \geqslant u\}, \quad u \in (0, 1).$$

Theorem

1. *Given a* $[0,1]$*-uniformly distributed random variable* U *and a cdf* F*, the random variable* $X \equiv F^{-1}(U)$ *has the cdf* F*.*
2. *The random variable* $U \equiv F_X(X)$ *is uniformly distributed on* $[0,1]$*.*

Proof.

1.

$$F_X(x) = \mathrm{P}\left(X \leqslant x\right) = \mathrm{P}\left(F^{-1}(U) \leqslant x\right) = \mathrm{P}\left(U \leqslant F(x)\right) = F(x).$$

The last equality holds for U is $[0,1]$-uniformly distributed.

2.

$$\begin{aligned} F_U(u) &= \mathrm{P}\left(U \leqslant u\right) = \mathrm{P}\left(F_X(X) \leqslant u\right) \\ &= \mathrm{P}\left(X \leqslant F_X^{-1}(u)\right) = F_X\left(F_X^{-1}(u)\right) = u, \quad u \in [0,1] \end{aligned}$$

which is the cdf of a $[0,1]$-uniformly distributed random variable.

□

Let $X_1, X_2, \ldots, X_n$ be random variables with cdfs $F_{X_1}, F_{X_2}, \ldots, F_{X_n}$; the random variables $U_i \equiv F_{X_i}(X_i), i = 1, 2, \ldots, n$ are uniformly distributed on $[0, 1]$. Furthermore, the function

$$\begin{aligned} C_{(X_1, X_2, \ldots, X_n)}&(u_1, u_2, \ldots, u_n) \\ &\equiv \mathrm{P}\left(U_1 \leqslant u_1, U_2 \leqslant u_2, \ldots, U_n \leqslant u_n\right) \\ &= \mathrm{P}\left(F_{X_1}(X_1) \leqslant u_1, F_{X_2}(X_2) \leqslant u_2, \ldots, F_{X_n}(X_n) \leqslant u_n\right) \\ &= \mathrm{P}\left(X_1 \leqslant F_{X_1}^{-1}(u_1), X_2 \leqslant F_{X_2}^{-1}(u_2), \ldots, X_n \leqslant F_{X_n}^{-1}(u_n)\right) \end{aligned}$$

is a copula. The joint cdf of $(X_1, X_2, \ldots, X_n)$ can be recover as

$$\begin{aligned} \mathrm{P}&\left(X_1 \leqslant x_1, X_2 \leqslant x_2, \ldots, X_n \leqslant x_n\right) \\ &= \mathrm{P}\left(F_{X_1}(X_1) \leqslant F_{X_1}(x_1), F_{X_2}(X_2) \leqslant F_{X_2}(x_2), \ldots, F_{X_n}(X_n) \leqslant F_{X_n}(x_n)\right) \\ &= \mathrm{P}\left(U_1 \leqslant F_{X_1}(x_1), U_2 \leqslant F_{X_2}(x_2), \ldots, U_n \leqslant F_{X_n}(x_n)\right) \\ &= C_{(X_1, X_2, \ldots, X_n)}(F_{X_1}(x_1), F_{X_2}(x_2), \ldots, F_{X_n}(x_n)) \end{aligned}$$

The above discussion culminates in the

Theorem (Sklar [31]; c.f. Durante and Sampi [9], Joe [17])

For a n-variate distribution F with j-th univariate margin F_j, the copula associated with F is a distribution function $C : [0,1]^n \to [0,1]$ with $U(0,1)$ margins that satisfies

$$F(y) = C(F_1(y_1), F_2(y_2), \ldots, F_n(y_n)), \quad y \in \mathbb{R}^n.$$

If F is a continuous n-variate distribution function with univariate margins $F_1, F_2, \ldots, F_n$ and quantile functions $F_1^{-1}, F_2^{-1}, \ldots, F_n^{-1}$, then

$$C(u) = F(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_n^{-1}(u_n))$$

is the unique choice.

The copula uniquely determined in $[0,1]^n$ for distributions F under absolutely continuous margins F_i has the density function

$$f(x_1,\ldots,x_n) = c\left(F_1\left(x_1\right),\ldots,F_n\left(x_n\right)\right)\prod_{i=1}^{n} f_i\left(x_i\right)$$

where f_i are the marginal densities and c is the density function of the copula given by

$$c\left(u_1,\ldots,u_n\right) = \frac{f\left(F_1^{-1}\left(u_1\right),\ldots,F_n^{-1}\left(u_n\right)\right)}{\prod\limits_{i=1}^{n} f_i\left(F_i^{-1}\left(u_i\right)\right)}.$$

The copula function $C : [0,1]^n \to [0,1]$ has the following properties:

1. $C(u_1, u_2, \ldots, u_n) = 0$ if at least one $u_i = 0$.
2. $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$.
3. For each n-dimensional rectangle
$\mathbb{T} \equiv [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$, $a_i, b_i \in [0,1]$, $a_i < b_i$,

$$0 \leqslant \sum_{(d_1, d_2, \ldots, d_n) \in \mathbb{T}} (-1)^{|\{j : d_j = a_j\}|} C(d_1, d_2, \ldots, d_n) \leqslant 1.$$

For example, if $n = 2$, the condition reads

$$0 \leqslant (-1)^0 C(b_1, b_2) + (-1)^2 C(a_1, a_2) + (-1)^1 C(b_1, a_2) + (-1)^1 C(a_1, b_2)$$

[1] C. Bernard and C. Czado. Multivariate option pricing using copulae. *Applied Stochastic Models in Business and Industry*, 29(5):509–526, 2013.

[2] X. Chen and Y. Fan. Estimation of copula-based semiparametric time series models. *J. Econometrics*, 130:307–335, 2006.

[3] X. Chen, W. B. Wu, and Y. Yi. Efficient estimation of copula-based semiparametric markov models. *Ann. Stat.*, 37:4214–4253, 2009.

[4] U. Cherubini and E. Luciano. Bivariate option pricing with copulas. *Applied Mathematical Finance*, 9(2):69–85, 2002.

[5] S.-C. Chiou and R.-S. Tsay. A copula-based approach to option pricing and risk assessment. *Journal of Data Science*, 6:273–301, 2008.

[6] J. Da Fonseca and J. Ziveyi. Valuing variable annuity guarantees on multiple assets. *Scandinavian Actuarial Journal*, 2017(3):209–230, 2017.

[7] W. F. Darsaw and B. Nguyen. Copulas and Markov processes. *Illinois J. Math.*, 36:600–642, 1992.

[8] J.-C. Duan. The GARCH option pricing model. *Mathematical Finance*, 5:13–32, 1995.

[9] F. Durante and C. Sampi. *Principles of Copula Theory*. Chapman & Hall/CRC, Boca Raton, F.L., 2016.

[10] P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and dependency in risk management: Properties and pitfalls. In M. Dempster, editor, *Risk Management: Value at Risk and Beyond*, pages 176–223. Cambridge University Press, 2002.

[11] J.-D. Fermanian and M. H. Wegkamp. Time-dependent copulas. *J. Multivariate Anal.*, 110:19–29, 2012.

[12] C. Genest and B. Rémillard. Tests of independence and or randomness based on the empirical copula process. *Test*, 13:335–369, 2004.

[13] C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas: A review and a power study. *Insurance Math. Econom.*, 44: 199–213, 2009.

[14] Alexios Ghalanos. *rugarch: Univariate GARCH models*, 2018. URL `https://cran.r-project.org/web/packages/rugarch/index.html`. R package version 1.4-0.

[15] Marius Hofert, Ivan Kojadinovic, Martin Mächler, and Jun Yan. *copula: Multivariate Dependence with Copulas*, 2017. URL

http://cran.r-project.org/package=copula. R package version 0.999-18.

[16] H. Joe. *Multivariate Models and Multivariate Dependence Concepts*. Chapman & Hall/CRC, Boca Raton, F.L., 1997.

[17] H. Joe. *Dependence Modeling with Copulas*. Chapman & Hall/CRC, Boca Raton, F.L., 2014.

[18] E. Jondeau and M. Rockinger. The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance*, 25(5):827–853, 2006.

[19] J. Klemelä. *Nonparametric Finance*. Wiley, New York, 2018.

[20] J.-F. Mai and M. Scherer. *Financial Engineering with Copulas Explained*. Palgrave Macmillan, New York, 2014.

[21] J.-F. Mai and M. Scherer. *Simulating Copulas: Stochastic Models, Sampling Algorithms, and Applications*. Imperial College Press, London, Second edition, 2017.

[22] R. Nelsen. *An Introduction to Copulas*. Springer-Verlag, New York, Second edition, 2006.

[23] C.-Y. Ng and S.-H. Li. Pricing and hedging variable annuity guarantees with multiasset stochastic investment models. *North American Actuarial Journal*, 17(1):41–62, 2013.

[24] A. Patton. Modelling asymmetric exchange rate dependence. *Internat. Econ. Rev.*, 47:527–556, 2006.

[25] A. Patton. Copula methods for forecasting multivariate time series. In *Handbook of Economic Forecasting. Volume 2B*, pages 899–960. Elsevier, Oxford, 2011.

[26] B. Rémillard. Non-parametric change point problems using multipliers. Social Science Research Network (SSRN) Working Paper Series, 2012. URL `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2043632`.

[27] B. Rémillard. *Statistical Methods for Financial Engineering*. Chapman & Hall/CRC, Boca Raton, F.L., 2013.

[28] B. Rémillard. Goodness-of-fit tests for copulas of multivariate time series. *Econometrics*, 5(1):1–23, 2017.

[29] J.V. Rosenberg. Pricing multivariate contigent claims using estimated risk-neutral density functions. *J. Int. Money Finan.*, 17:229–247, 1998.

[30] J.V. Rosenberg. Nonparametric pricing of multivariate contigent claims. *J. Derivatives*, 10:9–26, 2003.

[31] A. Sklar. Fonctions de répartition à *n* dimensions et leurs marges. *Publications de L'Institut de Statistique de L'Université de Paris*, 8: 229–231, 1959.

[32] R.M. Stulz. Options on the minimum or the maximum of two risky assets: Analysis and applications. *J. Finan. Econ.*, 10:161–185, 1982.

[33] R.W.J. van den Goorbergh, C. Genest, and B.J.M. Werker. Bivariate option pricing using dynamic copula models. *Insurance Math. Econom.*, 37:101–114, 2005.