

Copula function theory and classification

Yueguang Hu ^{1,a} and Zhigang Zhang ^{1,b}

Department of Information and Computing Science, University of Science and Technology Beijing
30 Xueyuan Road, Beijing, 10083, China

[a17888803371@163.com](mailto:17888803371@163.com), [bzzgcyf@263.net](mailto:zzgcyf@263.net)

Keywords: copula, the joint distribution function, independence

Abstract: "copula" comes from Latin "copulare", meaning "together", and connecting to the one-dimensional marginal distribution. Copula function is the multivariate distribution function on $[0,1]$, is also a usual method of measuring multivariate extreme value theory dependent function. According to the statistics, the copula theory can be traced back to the multivariate non-Gauss distribution in twentieth Century. Later In 1959, Professor Sklar has proposed the copula function, he believes that a joint distribution is too complex, and it can be decomposed into k marginal distribution, also, The copula function is not affected by the edge distribution, The copula function describes the correlation structure between variables and studies the correlation between variables, if the edge function is continuous, then the copula function is the only sure.

Introduction

With the continuous integration of the global economy, the risk and destruction of multiple factors make people have to start to study the risk and to avoid it. The appearance of Copula function has aroused people's attention. The application of Copula is also very extensive, in which the financial risk measurement is the most. In fact, it is not surprising. With the deepening of economic globalization, the capital flows in the world are more frequent, the relationship among financial markets becomes increasingly complex, but also to strengthen the linkage of financial market. The market interdependence greatly enhance the risk and destruction of financial market. Therefore, financial risk management is particularly important. The copula function has played a very important role in financial risk management.

The definition and theorem of Copula function

Definition of the N-copula function C:

1. $C = I^N = [0, 1]^N$;
2. C for each variable is monotonically increasing, that is, a monotonically increasing function of a single variable.
3. to meet the edge consistency, that is to say, $C_n(u_n) = C(0, \dots, 0, u_n, 0 \dots 0) = 0$; $C_n(u_n) = C(1, \dots, 1, u_n, 1 \dots 1) = u_n, u_n \in [0,1], n \in [1, N]$

The function C satisfying the above properties is the copula function.

Given the definition, the corresponding should also be given Sklar's theorem.

(Sklar's theorem) set X_1, \dots, X_n are n random variables, $K(X_1, \dots, X_n)$ are their joint distribution functions, and if $F_1(x_1), \dots, F_n(x_n)$ is its edge distribution function, there is a copula function $C(u_1, \dots, u_n)$, so that for any x, y , $K(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$

If $F(x)$, $k(x)$ is continuous, the copula function C is unique.

The properties of copula function

Copula function has very perfect properties, a variety of multivariate distribution can be constructed by copula function. We use copula theory to build financial models can study random variables and

their related structures independently, because the corollaries and relating measurements of the copula function will be not changed for strictly monotone increasing transformation. In order to facilitate the understanding. We use two variables Copula function $C(U, V)$ as an example to illustrate the nature of the copula function.

(1) set X_1, X_2 and X_3, X_4 are two independent random variables, $F_1(x), F_2(x), F_3(x), F_4(x)$ respectively, corresponding to their distribution functions. $C(F_1(x), F_2(x))C(F_3(x), F_4(x))$ are still their joint distribution functions.

It is easy to prove that our $C(F_1(x), F_2(x))$ and $C(F_3(x), F_4(x))$ are joint functions of X_1, X_2 and X_3, X_4 , Are all independent of each other, so the joint distribution multiplication is the joint distribution function of X_1, X_2, X_3, X_4

(2) For the binary copula function $C(u, v)$, if u and v are independent, then $C(u, v) = u \cdot v$; that is, if the two variables are independent of each other, the copula function can be defined by the edge function.

(3) C has two multiplicities, that is, the joint distribution increases with the increase of the marginal distribution. Take the binary copula function as an example $\forall u_1, u_2, v_1, v_2 \in I$ and $u_1 \leq u_2, v_1 \leq v_2$, then $c(u_2, v_2) - c(u_1, v_2) - c(u_2, v_1) + c(u_1, v_1) \geq 0$

(4) copula function has nonlinear transformation invariance. If the variables X and Y have Copula function C , and h_1, h_2 are incremental nonlinear continuous functions, then $h_1(X)$ and $h_2(Y)$ and also have the same Copula function C .

We all know that the commonly used correlation is only a linear correlation under the constant change in a correlation index. If the correlation of the nonlinear function is involved, an erroneous conclusion arises. But the copula function is not the same, it has a more extensive basis, than the linear range is larger.

(5) Convex combination of Copula functions is also Copula function. $C_i(u, v)$ is a Copula function, and its convex combination $C(u, v) = \sum_{i=1}^n \lambda_i C_i(u, v)$ for any n , when $\sum_{i=1}^n \lambda_i = 1, \lambda_i \geq 0$ 时 $C(u, v)$ is also a copula function.

(6) Monotonical increase: $C(u, v)$ is incremented for variables u and v . If a marginal distribution is constant and another marginal distribution increases, then the joint distribution will be also increased.

(7) $C(u, 0) = C(0, v) = 0, C(u, 1) = C(1, v) = v$. Namely if the probability of any marginal distribution is zero, the probability distribution of the corresponding joint occurrence is also zero; if the probability of any marginal distribution is 1, then the joint distribution is determined by another marginal distribution.

(8) (Frechet-Hosffding constraint) Assuming that C is a two – varieties copula function, for each $u \in [0, 1], v \in [0, 1], W(u, v) \leq C(u, v) \leq M(u, v)$.

And, $W(u, v) = \max(u + v - 1, 0)$ is the Frechet-Hosffding lower bound。

$M(u, v) = \min(u, v)$, is the Frechet-Hosffding upper bound。

Classification of Copula functions

The copula function classification have many ways. For example, according to the type of distribution function, it can be classified into Elliptic copula and Archimedean copula are ; According to the number of parameters, it can be classified into a single parameter copula, two parameter copula and multi parameter copula. We still take two variables copula function as an example to introduce some common forms of Copula function.

1、 The product copula function:

Definition: satisfying the form of the copula function $\prod uv = u \cdot v$ is product copula function.

The product copula function is the copula function of independent random variables, and it is the simplest form of copula function.

Theorem: if U, V is continuous random variables, only when the random variable copula function $C = \prod u.v$ was satisfied, U, V is independent for each other.

2、 Elliptic copula function:

Elliptic copula function includes normal copula function and t-copula function, they all belong to the elliptic copula function. However, they have different properties.

*normal copula function:*The normal copula function has many advantages, such as random sampling, few parameters, and the correlation structure is easy to calculate.

the form of Two variables normal copula function is:

$$c_{Ga}(u, v) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi(1-R_{12}^2)^{\frac{1}{2}}} \exp\left\{-\frac{s^2 - 2R_{12}st + t^2}{2(1-R_{12}^2)}\right\} dsdt \quad (1)$$

And R_{12} is the correlation coefficient matrix Φ is the standard normal distribution function.

Normal distribution random variable X_1, \dots, X_n , its mean value is μ_1, \dots, μ_n And variance is $\sigma_1, \dots, \sigma_n$. The covariance matrix is R, then the random variable $U_i = \Phi\left(\frac{X_i - \mu_i}{\sigma_i}\right), i \in I$ whose distribution function $C_R(\mu_1, \dots, \mu_n)$ is normal copula function whose covariance matrix is R.

t-copula function: the form of Two variables t-copula function is:

$$C_t(u, v, \gamma, \rho) = \int_{-\infty}^{v^{-1}(u)} \int_{-\infty}^{v^{-1}(v)} \frac{1}{2\pi(1-\rho^2)} \left\{1 + \frac{s^2 - 2\rho st + t^2}{v(1-\rho^2)}\right\}^{\frac{v+1}{2}} dsdt \quad (2)$$

Normal distribution random variable X_1, \dots, X_n , its mean value is 0, the variance is 1, and the covariance matrix is R. Y is χ^2 random variable, with V degrees of freedom, is also independent of

(X_1, \dots, X_n) . The random variable $U_i = t_v\left[\frac{\sqrt{v}}{\sqrt{Y}}X_i, i \in I\right]$ distribution function $C_{v,R}(\mu_1, \dots, \mu_n)$ is the copula function, which is called the t-copula function with the V degree of freedom, the covariance matrix R.

3、 Archimedean Copula function:

The Archimedean Copula function has the characteristics of combination, easy to construct, symmetry and easy calculation. Therefore, it is widely used in many fields.

Before defining the Archimedean Copula function, we first introduce the generating element of the copula function.

Definition: the quasi inverse function: for continuous strictly monotone function: $\varphi : [0,1] \rightarrow [0, \infty]$ satisfied $\varphi(1) = 0$. Then φ quasi inverse function φ^{-1} is:

$$\varphi^{-1}(t) = \begin{cases} \varphi^{-1}(t), & 0 \leq t \leq \varphi(0) \\ \varphi(0) \leq t \leq \infty \end{cases} \quad (3)$$

φ^{-1} domain is $[0, \infty]$, range is $[0,1]$. So in the case of multiple function φ is the generating element of Archimedean Copula function, $C(\mu_1, \dots, \mu_n) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_n))$. Since the generator is a strictly decreasing convex function on $[0,1]$, each Archimedean Copula function corresponds to a unique generator.

Set U, V is a uniform random variable on $[0,1]$, and their joint distribution function is generated by the generated Archimedean function $C(u,v)$, in the case of two variable, the distribution function of random variables is:

$$K_C(t) = P(C(U,V) \leq t) = t - \frac{\varphi(t)}{\varphi'(t^+)}, \quad t \in [0,1] \quad (4)$$

$\varphi'(t^+)$ Represents the right inverse at t.

When the generator meets the necessary conditions, in accordance with the above approach, you can get the form of multivariate Archimedean Copula function, such as Gumbel copula, Clayton copula and Frank copula, etc.. The Gumbel copula function usually has only the upper tail correlation, Clayton copula has only the lower tail correlation, and the Frank copula has neither the tail dependence nor the tail dependence.

The following describes the kinds of Copula function of the form of the two element function.

(1) Gumbel copula :

$$\varphi(t) = (-\ln t)^\alpha, \quad \alpha \in [1, \infty)$$

$$C_{Gu}(u, v, \alpha) = \exp\{-[(-\ln u)^\alpha + (-\ln v)^\alpha]^{\frac{1}{\alpha}}\} \quad (5)$$

(2) Clayton copula

$$\varphi(t) = \frac{1}{\alpha}(t^{-\alpha} - 1), \quad \alpha \in [-1, \infty) \setminus \{0\}$$

$$C_{cl}(u, v, \alpha) = \max\{(u^{-\alpha} + v^{-\alpha} - 1)^{\frac{1}{\alpha}}, 0\} \quad (6)$$

(3) Frank copula

$$\varphi(t) = \ln \frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}, \quad \alpha \in (-\infty, \infty) \setminus \{0\}$$

$$C_F(u, v, \alpha) = -\frac{1}{\alpha} \ln\left(1 + \frac{(e^{-\alpha u} - 1)(e^{-\alpha v} - 1)}{e^{-\alpha} - 1}\right) \quad (7)$$

4、Two-parameter copula function:

The study of two-parameter copula function is relatively small, the relevant literature and information is not much. The copula function has two parameters to represent the relationship, i.e. a copula function can capture two related patterns, for example, a parameter to describe the tail dependence, another parameter to describe the consistency; or, a description on the tail correlation, a description of lower tail dependence and so on.

Its function form is:

$$C(u_1, u_2) = \psi(-\log K(e^{-\psi^{-1}(v_1)}, e^{-\psi^{-1}(v_2)})) \quad (8)$$

And, ψ is a Laplace transform, K is one of the biggest infinitely divisible copula. ψ and K can take different forms, then the corresponding copula function and different in different forms, so the two variables copula is flexible.

The product copula function is the most basic copula function. And The elliptic copula function can be regarded as an indirect copula function, according to Sklar theorem and copula definition, the degree of correlation among the basic copula function. The Archimedean Copula function can be regarded as a direct copula function, because it is directly obtained by the concrete generating function. It should be noted that these are a single parameter model, the general formula for $C(u, v; \delta)$, where the parameters of the delta correlation. The same form of Copula function, the greater the δ , the stronger the correlation. The single parameter copula parameter is used to represent the consistency of the relevant metrics, such as Tau Kendall' and Spearman 's Rho, or tail correlation metric, such as the Archimedean Copula function. The parameters of the two parameters are used to represent the correlation parameters are two, so a copula function can capture the correlation of the two models.

As the copula function is a new branch of science developed recently, the related theories are still being studied and perfected. Therefore, in the near future, the copula function theory will be improved and enriched.

References:

- [1] Ming Han: Copula-a new econometric tools, statistics and information.Forum Vol.19 No.5(2004), p.93
- [2] Yaoting Zhang: connection function (copula) technology and financial risk analysis , statistical research No.4(2002)
- [3] Yanhua Wei, Shiyong Zhang, Lifeng Meng: the application of Copula theory in finance, Journal of NSTUAF .Vol.3 No.5(2003), p.97
- [4] Xiao Wang: Research on the properties of Copula function, Education space, p.139
- [5] Shushan Li:Copula regression and prediction, statistics and decision 18(20150), p.70
- [6] Ni Zhang, Yiwen Zhang: Research on the correlation between macroeconomic and stock market based on Copula theory, Value Engineering, p.3

- [7] Yan Mu, Zhongzhi Wang: A note on symmetric Copula, J. of Math. Vol.33(2013)No.6, p.1085
- [8] Jialong Liu: application of Copula function in financial risk measure, graduate thesis(2012)
- [9] Liqin Zhao: The Study of Financial Risk Measurement Based on Copula Function doctoral thesis(2009)