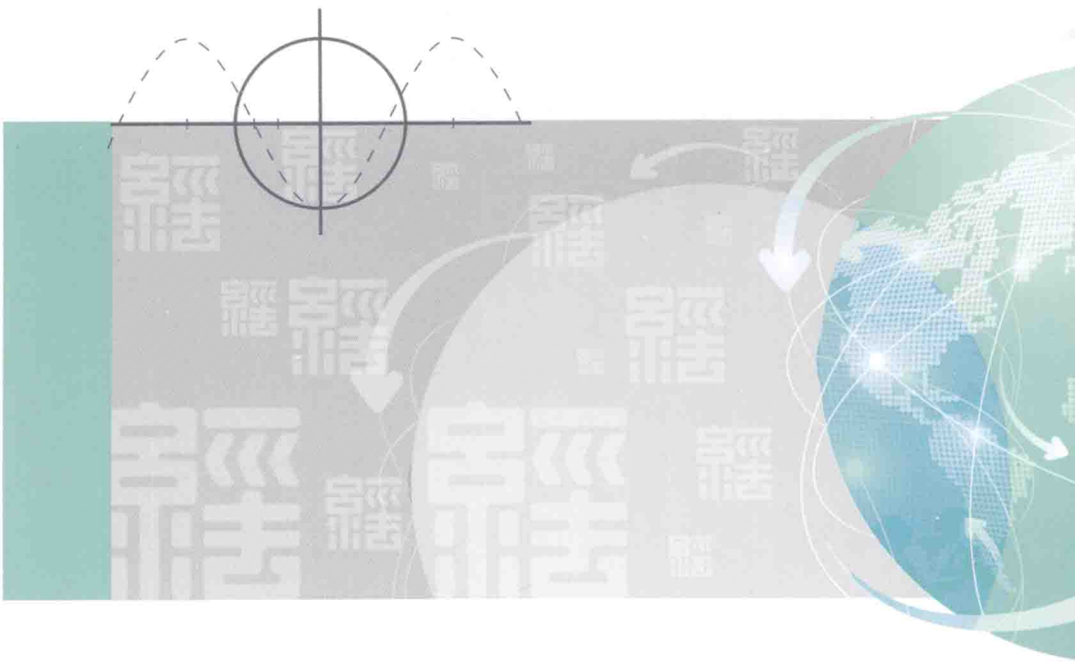
河南财经政法大学经济管理丛书

COPULA FANGFA JIQI YINGYONG

COPULA 方法及其应用

李霞（著）



经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

COPULA FANGFA JIQI YINGYONG

COPULA 方法及其应用

李霞（著）

经济管理出版社
ECONOMY & MANAGEMENT PUBLISHING HOUSE

图书在版编目（CIP）数据

COPULA方法及其应用/李霞著. —北京：经济管理出版社，2014.7
ISBN 978-7-5096-3158-4

Ⅰ.①C… Ⅱ.①李… Ⅲ.①时间序列分析—应用—金融风险—研究 Ⅳ.①F830.9 ②O211.61

中国版本图书馆CIP数据核字（2014）第125647号

组稿编辑：张　艳
责任编辑：张　艳　丁慧敏
责任印制：黄章平
责任校对：张　青

出版发行：经济管理出版社
（北京市海淀区北蜂窝8号中雅大厦A座11层　100038）
网　　址：www. E-mp. com. cn
电　　话：（010）51915602
印　　刷：大恒数码印刷（北京）有限公司
经　　销：新华书店
开　　本：710mm×1000mm/16
印　　张：13.5
字　　数：211千字
版　　次：2014年7月第1版　2014年7月第1次印刷
书　　号：ISBN 978-7-5096-3158-4
定　　价：45.00元

前 言

传统的概率统计常采用皮尔逊的相关系数来反映变量之间的相关关系，即便是多维随机变量的相关系数，也往往是二元相关系数的推广。我们知道，在变量之间的相关系数为非线性的情况下，线性系数并不能准确地反映变量之间的相关关系，那么，在变量之间的关系不能确定的情况下，我们应该选择什么指标来刻画变量之间的相关关系呢？针对这个问题，很多学者进行了大量分析，其中，张尧庭教授在文献中对我们应该选择什么样的相关性指标进行了分析，张教授对相关系数在刻画变量之间相依关系的局限性进行了阐述，无可厚非，长期以来，相关系数作为刻画变量之间线性相关关系的工具在描述变量之间的相依关系上作出了很大贡献，但是我们也应该看到这种方法的局限性。

首先，由两个随机变量的相关系数 ρ_{XY} 的计算公式 $\rho_{XY}=\dfrac{E\left(X-EX\right)\left(Y-EY\right)}{\sqrt{D\left(X\right)D\left(Y\right)}}$ 可以看出，若计算 ρ_{XY}，前提条件是 X、Y 必须有两阶矩，即必须有各自的方差及协方差，而金融市场中很多数据往往是厚尾分布的，即它们的二阶矩根本不存在，有的甚至连期望都不存在，显然不能用 ρ_{XY} 来刻画变量之间的相关性。

其次，若用 ρ_{XY} 刻画变量之间的相关性，会出现相关性很强的随机变量之间

的相关系数却为 0 的情况，如 $X \sim N(0,1)$，$Y = X^2$，Y 的取值由 X 确定，即 X、Y 之间具有很强的相关性，但是经计算得到的 $\rho_{XY} = \dfrac{EX^3 - EXEY^2}{\sqrt{D(X)D(Y)}}$，由于奇函数在对称区间内的积分为 0，故 $EX^3 = EX = 0$，即 $\rho_{XY} = 0$.

那么应该用什么来度量变量 X、Y 之间的相关性呢？哪种方法不要求 X、Y 有一阶矩，并且适合所有的分布呢？许多学者经过大量研究，提出了很多其他相依指标，如尾部相关性、象限相依性、Kendall τ 秩相关系数、Spearman ρ 相关系数和 Gini γ 相关系数等。

其中，尾部相关性，指的是当一个随机变量取较大值时，另一个随机变量也取较大值的概率增大，用式子表示即为，若 $P[Y \le y \mid X \le x]$ 是关于 x 的一个非降函数，则称两变量之间是尾部相关的。如金融市场中很多金融资产收益率之间的关系以及夫妻寿命之间的关系往往呈现出尾部相依性。

象限相依性，指的是对 R^2 中的所有 (x,y) 来讲，如果 $P[X \le x, Y \le y] \ge P[X \le x]P[Y \le y]$ 或者 $P[X > x, Y > y] \ge P[X > x]P[Y > y]$，我们称 X、Y 是象限相依的，若令 $\lambda(x,y) = P[X > x, Y > y] - P[X > x]P[Y > y]$，则 $\lambda(x,y)$ 作为 (x,y) 的函数，变化剧烈的点自然应该是相关性变化大的点，因此此函数可作为变量之间相关性的测度值。

Kendall τ 秩相关系数以及 Spearman ρ 相关系数主要是通过分析变量之间的变化是否协调来判断两变量之间的相依关系，Gini γ 相关系数和 Kendall τ 秩相关系数以及 Spearman ρ 相关系数相比，不仅可以衡量随机变量变化方向是否一致，而且可以衡量变化程度是否一致。

可以验证 Kendall τ、Spearman ρ 以及 Gini γ 等相关性指标与随机变量的边缘

分布无关，仅和联合分布的连接结构有关，而 Copula 函数作为联合分布函数和边缘分布函数之间的一个连接函数，自然也成了刻画随机变量之间相依结构的一个有效工具。

自从 Copula 函数被提出后，Copula 方法在变量之间相关性分析、金融风险及金融风险管理等方面得到了广泛的应用。笔者以变量之间的相依性为研究方向，一直关注着 Copula 理论的发展，并且进行了一些相关研究。本书主要是对 Copula 函数的一些基本理论、基本方法、方法的应用及目前常用的一些方法之间的比较、变量之间的随机模拟等问题进行了讨论，本书可以作为有志于 Copula 理论研究的学生的入门级图书。

本书的写作首先要感谢西南交通大学理学院数学系何平教授以及所有概率教研室的老师，何老师的 Copula 讨论班为本书的写作奠定了基础，在此对何老师及概率教研室的所有老师表示感谢！

其次，要感谢河南财经政法大学统计学院的领导及同事的支持和帮助，有了他们的鼓励和支持本书才顺利完成，在此对他们表示衷心的感谢！

最后，要感谢我的老公和儿子，他们的理解和支持是我写作此书的最大动力，家人的默默付出支持我顺利完成本书的写作，在此，我衷心地感谢他们！

由于作者水平有限，书中不当之处在所难免，恳请广大读者批评指正！

李　霞

2014 年 5 月

目 录

第1章　绪 论

1.1　Copula 理论的提出及研究现状

Copula 是拉丁语，原意是“连接”，Copula 的概念是 Sklar 在 1959 年回答 M. Frechet 关于多维分布函数和低维边缘之间关系的问题时首次引入的。起初，Copula 主要用于概率度量空间理论的发展，后来，随着理论的逐渐完善，它们又被用于确定随机变量之间相依性的非参数度量上。

Copula 之所以能受到统计学者的青睐主要有以下两个原因：第一个是 Copula 是一种研究相依性测度的方法；第二个是 Copula 作为构造二维分布族的起点，可用于多元模型分布和随机模拟。Copula 函数作为一种刻画变量之间相依机制的工具，几乎包含了随机变量所有的相依信息，在不能决定传统的线性相关系数能否正确度量变量之间的相关关系的情况下，Copula 函数对变量之间相关关系的分析很有用，Copula 函数的出现使变量之间的相依性刻画更加趋于完善。自从 Copula 方法被提出后，Copula 函数在金融资产收益率之间的相依性分析以及金融风险、金融风险管理等方面得到了广泛利用。

国外学者 Nelsen 在 1998 年对 Copula 函数的含义和性质做了全面详细的介绍。Patton、Cossette 和 Bouy 也分别在 2001 年、2002 年和 2003 年利用 Copula 理论做

了大量的研究工作。国外学者研究表明，Copula 理论在实际应用中有很多优点，如 Nelsen 在 2001 年指出，Copula 函数在变量单调增变换下保持不变，因此由 Copula 函数给出的一些相关性测度值，如 Kendall 的 τ 、Spearman 的 ρ 等在变量单调增变换下都不会发生变化，并且，一些特殊的 Copula 类，如 Archimedean Copula 函数，其上述测度值和 Copula 函数中的参数值之间有一一对应关系，这些结论为 Copula 函数模型形式的选择奠定了基础。另外，由于相关性测度值可以较准确地反映变量之间的相关程度，因此基于 Copula 函数的相关性测度值对时间序列的相关性分析及风险分析也都非常实用。Copula 理论也可用于分析时变的条件相关，国外学者 Bollerslev（1987）、Rockinger（2001）、Jondeau（2001）、Patton（2002）都提出，可以运用条件 Copula 建立具有时变、偏斜、尖峰、厚尾等特征的时间序列模型来处理非对称的相关性等问题，而利用 Copula 建立时间序列模型的关键问题是确定随机变量的边缘分布及一个能较好刻画变量之间相关关系的 Copula 函数。GARCH 类模型由于可以较好地描述时间序列的分布及波动特征，因此可用来表示边缘分布，而 Copula 函数可以假定其具有某种特定形式，如正态 Copula 函数，t-Copula 函数等，然后利用参数估计方法估计出参数。当然，对于边缘分布，除了 GARCH 类模型外，还可选择其他一些模型进行讨论，Nelson（1997）、Frees（1998）、Patton（2001）先后对此类问题进行了讨论。Copula 函数中某些 Copula 函数类的厚尾分布特征为分析金融资产收益率之间的相关性分析提供了方便，例如，利用 Copula 函数可以分析上海股市的较大波动是否会引起深圳股市的大波动，也可以分析一只股票的下跌是否会引起另一只股票的下跌，尾部相依关系的变化分析对于股市的波动溢出分析很有帮助，国外学者 Joe（1997）对 Copula 理论中尾部相依系数的应用进行了详细说明。其他学者 Boyer B. H.、M. S. Gibson 和 M. Loretan（2002）认为，在建立风险管理模型时仅仅考虑变量间的相关度是不够的，还必须考虑变量的相关结构。Hu（2002、2006、2010）

提出，混合 Copula 函数可以更好地捕捉到金融市场之间的尾部相依性。Okimoto（2008）利用 Copula 函数研究了证券市场的非对称相依结构，Bartram（2007），Chen、Tu 和 Wang（2008），Huang 等（2009），He 和 Gang（2009）等也对证券市场的相依机制进行了研究。

国内对 Copula 理论的研究起步较晚，国内学者张尧庭（2002）先探讨了 Copula 在金融领域应用的可行性。韦艳华、张世英（2004）结合 t-GARCH 模型和 Copula 函数，建立了 Copula-GARCH 模型对上海股市各板块指数收益率之间的条件相关性进行了分析，结果显示，不同板块的指数收益率序列具有不同的边缘分布，各序列间具有很强的正相关关系。司继文、蒙坚玲、龚木人（2005）揭示了 Copula 函数和 Kendall τ 统计量的内在关系，并基于 Copula 函数给出了变量尾部相关系数的表达式，实证分析结果显示，Copula 函数可以较好地描述国内外各股票市场之间的相依结构。李秀敏、史道济（2003）用混合 Copula 对上海，深圳股票市场进行相关性分析研究，用极值分布刻画了每只股票的边缘分布，用两步估计法对 Copula 参数进行了估计，结果显示，混合 Copula 比单个 Copula 更能捕捉金融市场间的尾部变化规律。秦伟良、王颖、达庆利（2007）用 Copula 函数研究了上证和深证市场的相关性，对上证指数和深证指数成指收益率的边缘分布分别用正则逆 Gamma 分布、偏 τ 分布进行拟合，然后在此基础上采用 Copula 函数方法建立上证和深证之间的联合分布，其中 Copula 函数采用的是金融领域常用的 Gumbel-Hougaard Copula、Frank Copula、Clayton Copula。Copula 函数中的参数采用推断函数法进行估计，分析结果显示沪深市场具有相关性。梁冯珍、钟君、史道济（2007）通过模拟方法研究了尾部相关性对投资组合风险度量 VaR 的影响，结果显示，当用 VaR 度量风险时，各资产收益率之间的线性相关系数小，不一定意味着投资风险小，风险还和尾部相关系数有关，若尾部相关系数大，则风险也可能较大。

1.2 Copula 方法研究相依结构的优越性

在传统的线性相关系数上，Granger 因果关系分析法是相关分析的常用方法。在金融时间序列的资产定价、投资组合等相关性分析问题中常被用到。国内学者也基于这两种方法对金融领域的相关问题进行了大量讨论，但是这些研究具有一定的局限性。众所周知，线性相关系数ρ刻画的是变量之间的线性相关程度，若$\rho = 0$，只能说明两者没有线性相关关系，如$X \sim N(0,1)$，$Y = X^2$，可计算$\rho = 0$，那么能否说明 X、Y 不相关呢，很显然，不能，X、Y 之间存在明显的非线性相关关系，因此用传统的线性相关系数ρ去度量变量之间非线性相关关系时，可能会得出错误的结论，再者，计算线性相关系数要求变量方差存在且有限，但金融时间序列多服从厚尾分布，方差有时并不存在，这一点也极大地限制了线性相关系数在金融领域中的应用。Granger 因果关系分析法只能对变量间因果关系给出定性描述，不能在数量上给予刻画。而 Copula 函数作为各变量边缘分布之间的连接函数，几乎包含了变量之间所有的相依信息，在传统的相关分析法不能正确衡量变量之间相关关系的情况下，为变量之间的相关分析提供了方便。

Copula 函数除了可以刻画变量之间的非线性关系外，还有许多其他的优越性质：

（1）Copula 函数的边缘分布较灵活，同一个 Copula 函数，其边缘分布可以为不同类型的分布，因此可利用 Copula 函数构造各种类型的多元分布，如可假设边缘分布为正态分布、对数正态分布、学生 t 分布、指数分布等构造 Copula 函数。Copula 函数作为各边缘分布之间的连接函数，其连接形式不受边缘分布的限制，相对于联合分布函数而言，这一点性质更加优越，联合分布函数一般受边缘分布函数形式的限制，多数情况下，多元联合分布函数是一元分布函数的拓广，如多元正态分布的边缘分布为一元正态分布，多元 t 分布的边缘分布为一元 t 分布，多

元均匀分布的边缘分布为一元均匀分布。

（2）Copula 函数在变量单调增变换下形式不发生变化，因此由 Copula 函数给出的一些相关性测度值如 Kendall 的 τ 、Spearman 的 ρ 等在变量单调增变换下都不会发生变化，这一点为 Copula 函数处理变量之间的非线性相关问题提供了方便，而由于传统的相关系数其测度的是变量之间线性相关程度，因此，只有在线性变换下，测度值才不会发生变化，这一点也限制了线性相关系数的应用。

（3）Copula 作为边缘分布函数的连接函数，其形式不受边缘分布的限制，因此，Copula 函数可以和边缘分布分开考虑，例如，在研究金融时间序列之间的相依结构时，单个序列的分布特征可由边缘分布描述，而金融时间序列之间的相依结构可以由 Copula 函数来刻画。

（4）Copula 函数形式的多样化，从结构上说，既可以是对称的，也可以是非对称的，还可以是对称和非对称组成的混合 Copula 函数形式；从相依性上讲，可以是上尾相依，也可以是下尾相依，或者是上尾相依和下尾相依组合在一起的混合 Copula。多变的 Copula 函数形式为变量之间非线性、非对称、上尾相依、下尾相依及混合相依关系的相依分析提供了方便，尤其是针对金融资产收益率之间较复杂多变的关系分析提供了一种新途径。

随着 Copula 理论的逐渐完善，其优越性愈加明显，近几年，Copula 理论越来越多地被应用到金融保险、金融风险管理、金融资产组合及其他研究领域，众多国内学者对金融资产收益率之间相依关系的研究为金融资产风险管理、金融资产组合等提供了理论依据。

1.3　本书的框架

本书共分五章：第 1 章是绪论，主要介绍 Copula 函数的国内外研究现状及

Copula 函数作为刻画变量之间相依关系的工具，其优越性；第 2 章是 Copula 函数的概述，主要介绍了 Copula 函数的概念、性质及其分类；第 3 章是阿基米德 Copula，作为 Copula 函数类的一种特殊 Copula，阿基米德 Copula 由于构造方便、计算简单，且具有许多良好的性质而在金融领域得到广泛应用，本章主要介绍了阿基米德 Copula 的定义及其性质、阿基米德 Copula 生成元的几种构造方法、几种常用的阿基米德 Copula 的尾部相依特征及阿基米德 Copula 函数在相依结构分析中的应用；第 4 章是具有阿基米德 Copula 函数的变量模拟生成的方法，分二维、多维分别对具有阿基米德 Copula 分布函数的变量模拟生成方法进行了讨论；第 5 章是 Copula 函数模型形式的选择，Copula 函数在实际应用中最为关键的问题就是 Copula 函数模型形式的选择，本章依据股票市场的样本数据对 Copula 函数模型的选择方法进行了讨论。

第 2 章　Copula 函数概述

2.1　Copula 函数的定义及 Sklar 定理

在给出 Copula 函数的定义之前，先给出 2-增函数的定义。设R表示实数域$(-\infty,+\infty)$，$\overline{R}$表示扩展的实数域$[-\infty,+\infty]$。

定义 2.1.1　设S_1和S_2是$\overline{R}$的两个非空子集，设H是定义在$D=S_1\times S_2$上的函数，$B=[x_1,x_2]\times[y_1,y_2]$是顶点位于区域$D$内的一个矩形区域，则$B$的$H$体积用$V_H(B)$表示，且

$$V_H(B)=H(x_2,y_2)-H(x_2,y_1)-H(x_1,y_2)+H(x_1,y_1)$$

若H在矩形区域的一阶差分定义如下：

$$\Delta_{x_1}^{x_2}H(x,y)=H(x_2,y)-H(x_1,y)$$

$$\Delta_{y_1}^{y_2}H(x,y)=H(x,y_2)-H(x,y_1)$$

则矩形区域B的H体积是H在区域B上的二阶差分，即

$$V_H(B)=\Delta_{y_1}^{y_2}\Delta_{x_1}^{x_2}H(x,y)$$

定义 2.1.2　一个二元实函数H是 2-增的，若对所有的顶点位于区域D内的

矩形区域B来说，都有$V_H(B)\geq 0$。

须注意，H是2-增的，并不意味着H关于每个变量都非降。

例题 2.1.1 设H是定义在I^2上的函数，$H(x,y)=\max(x,y)$，则H关于x,y分别是非降函数，但是$V_H(I^2)=1$，因此H不是2-增的。

例题 2.1.2 设H是定义在I^2上的函数，$H(x,y)=(2x-1)(2y-1)$，可以验证H是2-增的，但是在$\left(0,\frac{1}{2}\right)$区间内，$H$关于$x$是降的，同样在$\left(0,\frac{1}{2}\right)$区间内，$H$关于$y$也是降的。

根据例题2.1.1和例题2.1.2可看出，一个二元函数是2-增的，不一定关于每个变量都是单调非降的，但是如果我们对二元函数除了 2-增之外再附加一个条件，就可以得出二元函数关于每个变量非降的结论。

命题 2.1.1 设S_1和S_2是$\overline{R}$的两个非空子集，H是$S_1\times S_2\to R$上的2-增函数，设$x_1,x_2\in S_1$，$x_1\leq x_2$，$y_1,y_2\in S_2$，且$y_1\leq y_2$，则函数$t\to H(t,y_2)-H(t,y_1)$在S_1上是非降的，函数$t\to H(x_2,t)-H(x_1,t)$在S_2上是非降的。

定义 2.1.3 设S_1和S_2是$\overline{R}$的两个非空子集，a_1是S_1的最小值，a_2是S_2的最小值，H是$S_1\times S_2\to R$上的函数，则H称为有基底的，对$S_1\times S_2$所有的(x,y)都有$H(x,a_2)=0=H(a_1,y)$。

命题 2.1.2 设S_1和S_2是$\overline{R}$的两个非空子集，设H是定义在$S_1\times S_2$上的有基底的2-增函数，则H关于每个变量都是非降的。

证明：设a_1,a_2分别表示S_1和S_2的最小值，设$x_1=a_1$，$y_1=a_2$，则根据命题2.1.1知$t\to H(t,y_2)-H(t,y_1)$非降，设$x_2\geq x_1$，有

$$H(x_2,y_2)-H(x_2,y_1)\geq H(x_1,y_2)-H(x_1,y_1)$$

设$y_1=a_2$，则上式即为$H(x_2,y_2)\geq H(x_1,y_2)$，即$H(x,y)$是关于变量$y$的

非降函数；同样，由于 $t \to H(x_2,t)-H(x_1,t)$，设 $y_1 \le y_2$，有

$$H(x_2,y_2)-H(x_1,y_2) \ge H(x_2,y_1)-H(x_1,y_1)$$

设 $x_1 = a_1$，则上式即为 $H(x_2,y_2) \ge H(x_2,y_1)$，故 $H(x,y)$ 是关于变量 x 的非降函数。

例题 2.1.3 设 H 是定义在 $[-1,+1]\times[0,\infty]$ 上的函数，$H(x,y)=\dfrac{(x+1)(e^y-1)}{x+2e^y-1}$，因为 $H(x,0)=0$，$H(-1,y)=0$，因此 H 是有基底的，且 $V_H \ge 0$，根据命题 2.1.1 可知，H 分别关于 x, y 是非降的。

命题 2.1.3 设 S_1 和 S_2 是 $\overline{R}$ 的两个非空子集，设 H 是定义在 $S_1 \times S_2$ 上的有基底的 2-增函数，设 $(x_1,y_1),(x_2,y_2)$ 是 $S_1 \times S_2$ 上的任意一点，则 $\left|H(x_2,y_2)-H(x_1,y_1)\right| \le \left|F(x_2)-F(x_1)\right| + \left|G(y_2)-G(y_1)\right|$。

证明：由于 $\left|H(x_2,y_2)-H(x_1,y_1)\right| \le \left|H(x_2,y_2)-H(x_1,y_2)\right| + \left|H(x_1,y_2)-H(x_1,y_1)\right|$。

设 $x_1 \le x_2$，由于 H 是有基底的 2-增函数，根据命题 2.1.1 和命题 2.1.2 可得 $0 \le H(x_2,y_2)-H(x_1,y_2) \le F(x_2)-F(x_1)$，若 $x_1 \ge x_2$，我们可以得到类似的结论，故，对 S_1 中的任意 x_1, x_2，有 $0 \le \left|H(x_2,y_2)-H(x_1,y_2)\right| \le \left|F(x_2)-F(x_1)\right|$

同样的，我们可以得到对 S_2 中的任意 y_1, y_2，有

$0 \le \left|H(x_1,y_2)-H(x_1,y_1)\right| \le \left|G(y_2)-G(y_1)\right|$，证毕。

在上述定义基础之上，可引出 Copula 函数的定义。

定义 2.1.4 （Nelsen，1996，2006）二维 Copula 是满足以下条件的函数 C'：