学校编码：10384　　　　分类号 ______ 密级 ______

学号：17020061151763　　　　UDC ____________________

厦门大学

硕 士 学 位 论 文

关于风险中相依测度的研究

On measures of dependence for the risks

吴 倩

指导教师姓名: 张 志 强 副教授

专 业 名 称: 概率论与数理统计

论文提交日期: 2009 年 5 月

论文答辩日期: 2009 年 　月

学位授予日期: 2009 年 　月

答辩委员会主席: ____________

评　　阅　　人: ____________

2009 **年** 5 **月**

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　　　　　　　）课题（组）的研究成果，获得（　　　　　　　　）课题（组）经费或实验室的资助，在（　　　　　　　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文。并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于

（ ）1. 经厦门大学保密委员会审查核定的保密学位论文，于　年　月　日解密，解密后适用上述授权。

（ ）2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

目　录

Contents

摘　　要

关于相依的研究, 现在越来越多的出现在多元统计和概率论中. 它不仅只是个理论上的挑战, 更具有现实的意义. 这是因为在概率论与统计学中, 为了简化研究的目的，常常假设随机变量之间是相互独立，但实际上它们之间却存在着非常复杂的联系. 鉴于这些问题，对于随机变量相依测度的研究就显得尤为重要. 本文是基于前人对两维连续型相依测度研究的基础上，进行了多维的推广. 此外，还讨论了两维离散情形的相应结果. 具体内容如下：第一章, 介绍了本文的写作背景, 并简要介绍本文的主要研究工作. 第二章, 给出必要的基础知识. 第三章, 主要讨论了两维情形下的相依测度. 第四章，给出了相依性在多维连续情形下的推广.

关键词: 连接函数；和谐性；相依测度

Abstract

The study of dependence is one of the most frequently considered topic in multivariate statistics and probability theory, not only as a theoretical challenge, but also for its great importance in various practical applications. In light of complexity of correlations between random variables in probability theory and statistics, it is invariably postulated that random variables are independent from one another. Nevertheless, in practice absolute independence is hard to be satisfied. The paper improves the measurement which depict the dependence among the random variables, on the basis of the previous studies. Furthermore, this paper also focues heavily on dependence measures with non–continuous marginals. The outline for this paper is the following: In chapter 1, we present the background of the paper and simply introduce the main research; In chapter 2, we presents the necessary mathematical background; In chapter 3, we mainly descuss the measure of dependence in bivariate case; In chapter 4, we consider multivariate measures of dependence with continuous distributions.

Key words: Copulas ; Concordance; Measures of dependence.

第一章 引 言

关于相依的研究, 现在越来越多的出现在多元统计和概率论中. 它不仅仅是个理论上的挑战, 更具有现实的意义. 这是因为在概率论和统计学中, 为了使研究和计算方便的目的, 常常会假设随机变量间是相互独立的. 但实际上它们之间却存在着复杂的联系.

我们有许多方法去描述和度量随机变量之间的相依关系. 这种关系主要蕴涵在联合分布函数中. 由于边缘分布无法唯一确定联合分布, 因此若想借助于边缘分布来讨论相依性很难. 联合分布的每一个自变量的取值为整个实数轴, 是所要研究的问题更加显得复杂. 如何通过边缘分布函数来研究相依关系, 并且把每个自变量的取值变为某个特定区间上的取值? “ Copula ”就可以解决了这些问题. “ Copula ”是一个拉丁语名词, 它实质就是连接联合分布函数与边际分布函数的函数. 现在国内很多文献把它称为连接函数. 在本文中仍用原文“ Copula ”表示. 它可以把联合分布函数用其一维边缘分布函数表示, 把联合分布的定义域, 即整个平面转化为一个边长为 1 的正方形区域. 对随机变量的相依性研究就可以转化为对边际分布函数之间的研究, 利用 Copula 对相依性的研究就成为数学家关注的对象.

对这个问题的研究最早在 1940 和 1941 年, Hoeffding 发现“标准化函数”及其一些结论 [3].1951 年, Fréchet 也取得了很多和 Hoeffding 相同的结论, 还发现了“ Fréchet–Hoeffding 界”等重要的性质. 在 1959 年 Sklar 第一次把 Copula 这个定义引入到数学中来, 并且证明了“ Copula ”的存在性. 二十世纪七十年代 Kimeldorf 和 Smapson , Galmabor 和 Deheuvels 合作研究概率度量空间也得到了很多关于 Copula 的重要结论 [5][6]. 通过在二十世纪九十年代这十年四次著名的国际数学会议, “ Copula ”在统计与概率的应用中得到发展. Nelesn 在 1999 年把 Copula 的定义及其重要结论做了系统总结, 从而使 Copula 这个概念得以进一步的推广. 现在 Copula 还用在随机过程等领域中, 它也越来越焕发出勃勃生机.

最早利用 Copula 研究随机变量相依性的是 Schiweiezer 和 Wolff. 他们讨论了

两个随机变量之间的相依测度, 在对随机变量进行严格单调变换下的不变性. Scarsini 在 1984 年基于 Copula 给出了一系列关于两维连续型随机变量和谐测度的公理. 常见的 Kendall's τ, Spearman's ρ, Blomqvi-st's β, Gini's γ 都属于和谐测度这个范畴. 在 Scarsini 提出的公理中, $I^n = [0,1]^n$ 的对称性, 对多维理论起到了重要的作用. 这个结论由 Taylor 在 2007 年推广至多维. 但是和谐测度有一个不足之处, 就是和谐测度为零, 与随机变量间的独立并不充分必要. 这使得判断随机变量间是否独立充满了障碍. 于是, 改进得到一个满足条件充要的测度, 既是所谓的相依测度. 比如 Schweizer 和 Wolff 在 1981 年构造了一个测度, 这个测度与 Spearman's ρ 具有相似的形式. 并由 Nelsen 构造出类似于 Kendall's τ 和 Gini's γ 的相依测度. 但是这些相依测度都是针对于两维连续型随机变量. 基于前人的研究上, 本文将给出一个多维连续型相依测度的构建方法, 并给出实例.

随着理论的发展, 用 Copula 去描述随机变量间的关系是恰当而合理的. 但大部分的研究都是基于边缘连续的假设之上的. 有时对这种连续模型进行讨论是不符合事实的. 因此, 我们希望构建出相对于连续情形类似的, 离散条件下的测度. 然而, 正如 Albert Marshall 在 1996 年指出, 在边缘非连续时, 部分在连续时已知的结论就不再成立了. 2004 年, Johana Nešlehov$\acute{a}$ 构建了模型, 用于描述非独立随机变量间的相依结构. 在边缘离散并具有有限支撑的前提下, 讨论了在单调变换作用下 Copula 的变化. 虽然会存在一系列的 Copula 满足前面所提到的 Sklar's 方程, 但在它们之中, 有一类比较特殊, 性质优良的 Copula , 叫做标准延拓 Copula. 它可以反映出大量联合分布函数的相依结构. 许多好的性质也可以得以保留. 在连续情形中的基本 Copula , 也将被这里的标准延拓 Copula 所代替. 本文的另一个目的, 就是给出非连续情形下, 类似与连续相依测度的六条性质. 并且给出基于 L_p–距离的测度构建. 但是由于非连续情形下的不确定性, 使我们只能研究具有有限支撑的两维离散分布.

第二章 预备知识

§2.1 连接函数（ Copula ）

§2.1.1 Copula 函数定义、性质

Copula 被 Sklar 首次引入，用于统计学，是表示将一元分布函数“连接”起来形成多元分布函数 [1]. 为了给出它的定义和基本性质，首先给出几个预备概念.

$\overline{R}$ 表示扩展后的实轴 $[-\infty,+\infty]$, $S_1,S_2,\cdots,S_n$ 是 $\overline{R}$ 的非空子集， H 是定义在 $DomH=S_1\times S_2\times\cdots\times S_n$ 上的实函数，向量 $\mathbf{a}=(a_1,a_2,\cdots,a_n)$,$\mathbf{b}=(b_1,b_2,\cdots,b_n)$. 并且对任意的 k 有 $a_k\le b_k$. 令 $B=[a_1,b_1]\times\cdots\times[a_n,b_n]$,B 的所有顶点都在 $DomH$ 内. $\mathbf{c}=(c_1,c_2,\cdots,c_n)$ 表示 B 的顶点，定义

$$V_H(B)=\sum sgn(\mathbf{c})H(\mathbf{c}),$$

其中

$$sgn(\mathbf{c})=\begin{cases}1 & \text{若有偶数个 } c_k=a_k\\ -1 & \text{若有奇数个 } c_k=a_k.\end{cases}$$

(1) 如果 $V_H(B)\ge 0$ 成立，其中 B 的所有顶点都落在 DomH 内，则称 H 是 n 元增函数.

(2)H 是定义在 $DomH=S_1\times S_2\times\cdots\times S_n$ 上的实函数，假设每个 S_k 都有一个最小的元素 a_k, 对于任意的向量 $\mathbf{t}=(t_1,t_2,\cdots,t_n)\in DomH$, 若存在一个 $t_k=a_k$ 便有 $H(\mathbf{t})=0$ 成立，那么则称 H 是基础的.

(3) 若每个 S_k 是非空的和都有一个最大元素 b_k, 则函数 H 有边缘函数而且其一维的边缘函数 H_k 满足 $DomH_k=S_k$ 以及 $H_k(x)=H(b_1,\ldots,b_{k-1},x,b_{k+1},\ldots,b_n)$ 对任意的 $x\in S_k$ 都成立. 类似地，我们也可以定义更高维度的边缘函数。

引理 2.1 若 $S_1,S_2,\cdots,S_n$ 是 $\overline{R}$ 的非空子集， H 是定义在 $DomH=S_1\times S_2\times\cdots\times S_n$ 上的实函数且是基础的 n 元增函数. 则 H 是关于每个变量是递增的,

即对任意的 $(t_1,\ldots,t_{k-1},x,t_{k+1},\ldots,t_n),(t_1,\ldots,t_{k-1},y,t_{k+1},\ \ldots,t_n)\in DomH$ 及 $x\le y$, 都有 $H(t_1,\ldots,t_{k-1},x,t_{k+1},\ldots,t_n)\le H(t_1,\ldots,\ t_{k-1},y,t_{k+1},\ldots,t_n)$.

引理 **2.2** 若 $S_1,S_2,\cdots,S_n$ 是 $\overline{R}$ 的非空子集, H 是定义在 $DomH=S_1\times S_2\times\cdots\times S_n$ 上的实函数且是基础的 n 元增函数. 则对于任意的 $\mathbf{x}=(x_1,\ldots,x_n),\mathbf{y}=(y_1,\ldots,y_n)\in S_1\times S_2\times\cdots\times S_n$, 都有

$$\mid H(\mathbf{x})-H(\mathbf{y})\mid\le\sum_{k=1}^{n}\mid H_k(x_k)-H_k(y_k)\mid$$

定义 **2.1** 若 H 是定义在 $DomH=S_1\times S_2\times\cdots\times S_n$ 上基础的 n 元增函数. 且 $H(\infty,\ldots,\infty)=1$. 则称 H 是 n 维的分布函数.

由引理 2.1 可得：任一 n 维的分布函数其一维的边缘函数也是分布函数，我们在这把它们记为 $F_1,\ldots,F_n$.

定义 **2.2** 一个 n 维的 Copula 是定义在 $[0,1]^n$ 上的实函数，以符号 C 表示，并满足以下二个条件：

(a). C 是基础的、 n 元增函数;

(b). C 的所有边缘分布 C_i 满足： $C_i(u)=C(1,\cdots,1,u,1,\cdots,1)=u$, 其中 $u\in[0,1]$, $i=1,2,\ldots,n$.

对于任一个 n-copula $C,n\ge3$, 其每一 k 维的边缘分布是 k-copula. 或者说一个 n-copula C 是 $[0,1]^n\to[0,1]$ 的函数且满足以下两个性质：

1. 对任 $\mathbf{u}\in[0,1]^n$ ， $C(\mathbf{u})=0$ 当 $\mathbf{u}$ 至少有一个分量为 0 时， $C(\mathbf{u})=u_k$ 当 $\mathbf{u}$ 的所有的分量除 u_k 以外都为 1 时.

2. 对于任意的 $\mathbf{a},\mathbf{b}\in[0,1]^n$, 若满足 $a_i\le b_i$, $\forall$ i, 有 $V_C([\mathbf{a},\mathbf{b}])\ge0$.

因为 n-copula 是 $[0,1]^n$ 上的分布函数，所以其 $[0,1]^n$ 在概率测度可由以下式子确定：

$$V_C([0,u_1]\times\cdots\times[0,u_n])=C(u_1,\ldots,u_n).$$

这个概率测度就记为 V_C.

§2.1.2 Sklar **定理**

以下定理就是著名的 Sklar 定理，它是关于 Copula 理论中最重要的结论，也是 Copula 在各个方面应用中最有意义的。

定理 **2.1** 若 H 是一个 n 维联合分布函数，其边缘分布函数为 $F_1, \ldots, F_n$. 那么一定存在一个 n-copula 函数 C，使得：对所有的 $\mathbf{x} \in \bar{R}^n$, 有

$$H(x_1, \ldots x_n) = C(F_1(x_1), \ldots, F_n(x_n)). \tag{2.1}$$

如果边缘分布 $F_1, \ldots, F_n$ 是连续的, 那么 Copula 形式是唯一的. 否则 C 在 $RanF_1$, $\ldots, RanF_n$ 上就不能唯一确定. 反过来，如果 C 是 Copula 函数， $F_1, \ldots, F_n$ 边缘分布函数，那么 H 就是边缘分布为 $F_1, \ldots, F_n$ 的联合分布函数 [7].

若 F 单变量的分布函数，我们定义 F 的广义逆函数为 $F^{-1}(t) = inf\{x \in R \mid F(x) \geq t\}$ $\forall t \in [0,1]$, 并规定 $inf\emptyset = -\infty$.

引理 **2.3** 若 H 是一个 n 维连续的联合分布函数,其边缘分布函数为 $F_1, \ldots, F_n$ 及 Copula C (这时 C 满足式 (2.1)). 那么对 $\forall \mathbf{u} \in [0,1]^n$, 有

$$C(u_1, \ldots, u_n) = H(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)).$$

若没有连续的条件，则结论就不一定成立 [10].

§2.1.3 Copula **函数的** “ Fréchet-Hoeffding **界**”

首先定义 $[0,1]^n$ 上的三个函数 M^n, Π^n 和 W^n 如下：

$$M^n(\mathbf{u}) = min(u_1, \ldots, u_n), \Pi^n(\mathbf{u}) = u_1 \cdots u_n, W^n(\mathbf{u}) = max(u_1 + \cdots + u_n - n + 1, 0).$$

其中函数 M^n 和 Π^n 对 $\forall n \geq 2$ 都是 n-copulas，而函数 W^n 对 $\forall n \geq 3$ 则不是一个 Copula.

下面的定理是 n 维 Copula 函数的 Fréchet-Hoeffding 界的不等式.

定理 **2.2** 如果 C 是任一个 n-copula, 那么对于 $[0,1]^n$ 中的任一向量 $\mathbf{u}$, 都有

$$W^n(\mathbf{u}) \leq C(\mathbf{u}) \leq M^n(\mathbf{u}).$$ [11]

尽管 FréchetHoeffding 的下界 W^n 不是一个 copula 当 $n \geq 3$ 时，但它却是在以下意义下最好的下界.

定理 **2.3** 对于 $\forall \mathbf{u} \in [0,1]^n$ 且 $n \geq 3$, 存在一个 n-copula C(其依赖于 $\mathbf{u}$), 使得

$$C(\mathbf{u}) = W^n(\mathbf{u}).$$[7]

我们记 $\bar{C}$ 为 n 个随机变量的联合生存函数，且这 n 个随机变量的联合分布函数为 C, 即若 $(U_1, \ldots, U_n)^T$ 的分布函数是 C, 则有 $\bar{C}((u_1, \ldots, u_n) = P\{U_1 > u_1, \ldots, U_n > u_n\}$.

定义 **2.3** 设 C_1 和 C_2 都是 n 维 Copula 函数，如果对于 $[0,1]^n$ 中任意的 $\mathbf{u}$ 都有

$$C_1(\mathbf{u}) \leq C_2(\mathbf{u}), \quad \bar{C}_1(\mathbf{u}) \leq \bar{C}_2(\mathbf{u}).$$

我们称 C_1 小于 C_2(或者说 C_2 大于 C_1)，我们记做 $C_1 \prec C_2$(或者 $C_1 \succ C_2$). 注意对于双变量有如下结论：

$$\begin{aligned}\bar{C}_1(u_1,u_2) \leq \bar{C}_2(u_1,u_2) \quad &\Leftrightarrow \quad 1 - u_1 - u_2 + C_1(u_1,u_2) \leq 1 - u_1 - u_2 + C_2(u_1,u_2) \\ &\Leftrightarrow \quad C_1(u_1,u_2) \leq C_2(u_1,u_2).\end{aligned}$$

因此，Fréchet-Hoeffding 下界 W^n 小于任何一个 n 维 Copula 函数, 而 Fréchet-Hoeffding 上界又大于任何一个 n 维 Copula 函数. 对于含参数的 Copula 函数，比如单参数连接函数集 C_θ, 如果对于任何的 $\theta_1 \leq \theta_2$, 有：

$$C_{\theta_1} \prec C_{\theta_2}.$$

那么，我们称 Copula 函数集 C_θ 是正序的；相反地，如果 $\theta_1 \leq \theta_2$, 有：

$$C_{\theta_1} \succ C_{\theta_2}.$$

那么我们称 Copula 函数集 C_θ 是逆序的.

§2.1.4 Copula **函数和随机变量**

因为 $X_1,\ldots,X_n$ 相互独立当且仅当对于 $\forall x_1,\ldots,x_n\in\bar{R}$, 有 $H(x_1,\ldots,x_n)=F_1(x_1)\cdots F_n(x_n)$, 则以下结论可从定理 2.1 的得到.

定理 **2.4** 若 $(X_1,\ldots,X_n)^T$ 是由连续的随机变量构成的 n 维向量, 其 copula 函数为 C, 则 $X_1,\ldots,X_n$ 相互独立当且仅当 $C=\Pi^n$.

Copula 函数的一个重要性质是对于随机变量的严格单调变换具有不变性或者以简单形式改变. 注意到, 如果随机变量 X 的分布函数是连续的, 并且函数 α 在 $RanX$ 上是严格单调函数，那么新的随机变量的 $\alpha(X)$ 的分布函数也是连续的.

定理 **2.5** $(X_1,\ldots,X_n)^T$ 是由连续的随机变量构成的 n 维向量，其 Copula 函数为 C. 若 $\alpha_1,\ldots,\alpha_n$ 分别在 $RanX_1,\ldots,RanX_n$ 是严格的单调递增函数，则 $(\alpha_1(X_1),\ldots,\alpha_n(X_n))^T$ 的 Copula 函数也是 C[8].

从 Sklar 定理，我们得出 Copula 函数 C 把 n 维分布函数分成了边缘分布和一个表示相关结构的函数. 以下的这个定理表明，也同样存在一函数 $\hat{C}$ ，将 n 维生存函数分解成单一的边缘分布和一个表示相关结构的函数. 进一步可以证明出这个函数也是 Copula 及这个生存 Copula 函数能表示成 C 和它的 k 维边缘分布的形式.

定理 **2.6** 设 $X_1,\ldots,X_n$ 是连续的随机变量, 有 Copula 函数 $C_{X_1,\ldots,X_n}$. 函数 $\alpha_1,\ldots,\alpha_n$ 分别在 $RanX_1,\ldots,RanX_n$ 上是严格单调函数, 且设 $\alpha_1(X_1),\ldots,\alpha_n(X_n)$ 有 Copula 函数 $C_{\alpha_1(X_1),\ldots,\alpha_n(X_n)}$. 更进一步，我们假定对于某一个 k ， α_k 是一个严格递减函数，此处 $1\le k\le n$ ，不失一般性，假设 $k=1$ ，那么

$$\begin{aligned}C_{\alpha_1(X_1),\ldots,\alpha_n(X_n)}(u_1,u_2,\ldots,u_n) &= C_{\alpha_2(X_2),\ldots,\alpha_n(X_n)}(u_2,\ldots,u_n)\\ &\quad -C_{X_1,\alpha_2(X_2),\ldots,\alpha_n(X_n)}(1-u_1,u_2,\ldots,u_n).\end{aligned}$$

使用上面两个定理, 得知连接函数 $C_{\alpha_1(X_1),\ldots,\alpha_n(X_n)}$ 可以通过 $C_{X_1,\ldots,X_n}$ 来表示. 下面通过一个实例来加深理解.

例 **2.1** 考虑双变量得情况，设 α_1 是一个严格单调递减函数，而 α_2 是严格

单调递增函数，那么：

$$
\begin{aligned}
C_{\alpha_1(X_1),\alpha_2(X_2)}(u_1,u_2) &= u_2 - C_{X_1,\alpha_2(X_2)}(1-u_1,u_2) \\
&= u_2 - C_{X_1,X_2}(1-u_1,u_2).
\end{aligned}
$$

如果 α_1 和 α_2 都是严格单调递减的，那么：

$$
\begin{aligned}
C_{\alpha_1(X_1),\alpha_2(X_2)}(u_1,u_2) &= u_2 - C_{X_1,\alpha_2(X_2)}(1-u_1,u_2) \\
&= u_2 - (1-u_1-C_{X_1,X_2}(1-u_1,1-u_2)) \\
&= u_1+u_2-1+C_{X_1,X_2}(1-u_1,1-u_2).
\end{aligned}
$$

因此， $C_{\alpha_1(X_1),\alpha_2(X_2)}$ 是 $(X_1,X_2)^T$ 的一个生存 Copula 函数 $\hat{C}$, 也就是说：

$$H(x_1,x_2)=P\{X_1>x_1,X_2>x_2\}=\hat{C}(\bar{F}_1(x_1),\bar{F}_2(x_2)).$$

注意到：对于 n 个在 [0,1] 上服从均匀分布的随机变量，具有 Copula 函数 C, 那么其联合生存函数如下： $\bar{C}(u_1,\ldots,u_n)=\hat{C}(1-u_1,\ldots,1-u_n)$.

对任何的 Copula 函数 C ，设

$$C(u_1,\ldots,u_n)=A_C(u_1,\ldots,u_n)+S_C(u_1,\ldots,u_n). \tag{2.2}$$

其中,

$$A_C(u_1,\ldots,u_n)=\int_0^{u_1}\cdots\int_0^{u_n}C(s_1,\ldots,s_n)ds_1\ldots ds_n,$$

$$S_C(u_1,\ldots,u_n)=C(u_1,\ldots,u_n)-A_C(u_1,\ldots,u_n).$$

如果在 $[0,1]^n$ 上 $C=A_C$, 那么 C 被称为绝对连续的，在这个情况下， C 有密度函数 $\frac{\partial^n}{\partial u_1\cdots\partial u_n}C(u_1,\ldots,u_n)$.

如果在 $[0,1]^n$ 上 $C=S_C$, 那么 C 被称为是奇异的，且在 $[0,1]^n$ 中，几乎处处都有 $\frac{\partial^n}{\partial u_1\cdots\partial u_n}C(u_1,\ldots,u_n)=0$.

§2.2 M **的重组**