学校编码：10384　　　　　　　　　　分类号_______密级______
学号：19020111152538　　　　　　　　　　UDC_______

厦门大学

硕　士　学　位　论　文

二维相依样本随机序的比较

The comparison of two dimension dependent stochastic orders

尹　鑫

指导教师姓名：　李效虎教授
专 业 名 称：　概 率 论 与 数 理 统 计
论文提交日期：　2014 年　　月
论文答辩时间：　2014 年　　月
学位授予日期：　2014 年　　月

答辩委员会主席：__________
评　　阅　　人：__________

2014 年　　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

摘要

本文考虑二维相依样本的极值比较问题. 我们选取了几个特殊的阿基米德copula刻画相依性, 分别在齐次样本, 比例反失效率模型(PRH), 非齐次伽马随机变量的情形下, 得到了两组样本的极值间的随机比较结果, 将非齐次PRH模型的结果推广到一般阿基米德copula, 并研究了齐次样本下任意两个不同的copula的序的比较. 对于部分未能从理论上证明的随机比较结果, 我们通过例子进行探讨. 最后给出了理论结果在可靠性理论和拍卖理论中的应用.

关键词： 顺序统计量; 随机序; copula; Schur-Convex; PRH.

Abstract

This article considers bivariate dependent sample extremum problems. We choose some special Archimedean copulas to characterize the dependence, and obtain some results on stochastic comparison of two dependent samples' extremum in homogeneous samples, nonhomogeneous PRH samples and nonhomogeneous gamma samples, respectively. We also extend the stochastic comparison to two general Archimedean copulas. For stochastic orders that we cannot prove, we illustrate them through some numerical examples. Finally, we present two applications of our main results.

Key Words: order statistic; stochastic order; copula; Schur-Convex; PRH.

目 录

CONTENTS

第一章　背景

在实际应用中我们不可避免会遇到几个随机变量的大小比较, 比如在精算中风险实际上可以用一个随机变量来刻画, 比较哪个风险更危险的问题即可量化为对应的随机变量的大小比较. 但是因为随机变量不是常数, 用传统常数的大小比较方法显然不能得到令人满意的结果. 如何比较两个随机变量的大小从概率理论产生开始就一直困惑着人们. Littlewood和Polya 提出刻画非负向量的分散程度的majorization序后, Karamata (1932), Choquet (1963)以及Strassen (1965)把majorization序扩展到一般的概率理论中. 随后随机序进入快速发展中. 产生了如似然比序, 失效率序等一些有重要用途的随机序.

顺序统计量具有某些不依赖于总体分布的性质, 并且在诸多领域中扮演着重要的角色, 如统计推断, 优度拟合检验, 可靠性理论, 质量监管等, 因此近年来顺序统计量尤其是最大最小顺序统计量受到了广泛的关注. 考虑一组随机变量$X_1, \cdots, X_n$, 我们用$X_{i:n}$表示n个随机变量中第i小顺序统计量, 其中$X_{1:n}$, $X_{n:n}$分别表示最小, 最大顺序统计量. 对$X_1, \cdots, X_n$独立同分布的情形下的顺序统计量的随机序方面的研究, 很多文献已经进行了详尽的探讨. 可参阅Khaledi和Kochar (2000), Khaledi和Kochar (2006). 对于独立不同分布的情形, 针对顺序统计量的计算较为复杂, 因此研究较少, 可参阅David和Nagaraja (2003)和Balakrishnan (2007).

近年来, 对相依性的研究受到越来越多的关注, 人们提出了不同的方法对相依性进行刻画, 常用的有相关系数, 协方差, copula函数等. 有些度量指标过于粗糙, 如相关系数只是用一个数值来刻画总体的线性相关指标, 无法刻画非线性相关性. copula从函数的角度对相依性进行刻画, 用copula研究相依性我们可以不去关心随机向量的边际分布的具体形式而是把随机向量的依赖关系完全由copula来描述, 因此也有人称其为"连接函数" 所以相比其它度量指标, 能够更好, 更一般的描述变量之间的相依性. 其中, Genest和MacKay在1986年提出的阿基米德copula族能够刻画多种相依性结构, 且该copula族类被提出时主要应用于统计领域, 所以关于该copula类的统计估计方法有较多结果, 使得其在实际应用中备受关注. 近年来随着精算, 保险的发展, 越来越多的专家关注阿基米德copula在该领域的应用: Frees 和Valdez (1988), Klugman和Parsa (1999)以及Denuit, Purcaru和Van Keilegom (2004).为研究相依性对顺序统计量的影响, 本文选取阿基米德copula刻画二维随机变量的相依性, 首先研究了在边际分布相同, 同一类阿基米德copula但是参数不同时的序的关系, 随后研究了基于非齐次PRH 模型最大顺序统计量的随机比较结果. 而对于最小顺序统计量我们没有得到理论结果, 但用具体例子对它进行了进一步说明. 此外, 在伽马模型下我们证明了当形状参数大于零小于一时, 最小顺序统计量存在某些随机序关系.

第二章　预备知识

在陈述主要结果之前, 我们首先回顾一些重要的定义及已有的理论结果.

定义 2.1: 令X和Y为两个随机变量, 用f, F, $\bar{F}$和g, G, $\bar{G}$分别表示它们的概率密度函数, 分布函数和生存函数. 称

(1) X在通常的随机序下小于Y (记为$X \le_{\text{st}} Y$), 如果$\bar{F}(x) \le \bar{G}(x)$;

(2) X以失效率序小于Y(记为$X \le_{\text{hr}} Y$), 如果$\dfrac{\bar{G}(x)}{\bar{F}(x)}$是单调递增的;

(3) X以反失效率序小于Y(符号表示为$X \le_{\text{rh}} Y$), 如果$\dfrac{G(t)}{F(t)}$是单调递增的;

(4) X 以似然比序小于Y(符号表示为$X \le_{\text{lr}} Y$), 如果$\dfrac{g(x)}{f(x)}$是单调递增的.

并且关于上述随机序我们知道:

$X \le_{\text{lr}} Y \Rightarrow X \le_{\text{rh}} Y \Rightarrow X \le_{\text{st}} Y$;

$X \le_{\text{lr}} Y \Rightarrow X \le_{\text{hr}} Y \Rightarrow X \le_{\text{st}} Y$.

关于更多的随机序及其应用, 可参阅Shaked和Shanthikumar (2007).

记$x_{(1)} \le \cdots \le x_{(n)}$为向量$\boldsymbol{x} = (x_1, \cdots, x_n)$分量的递增排列, 我们有如下定义:

定义 2.2: 对于实向量$\boldsymbol{x} = (x_1, \cdots, x_n)$和$\boldsymbol{y} = (y_1, \cdots, y_n)$, 称

(i) $\boldsymbol{x}$以majorization序大于$\boldsymbol{y}$(记为$\boldsymbol{x} \overset{\text{m}}{\succeq} \boldsymbol{y}$), 如果

$$\sum_{i=1}^{j} x_{(i)} \le \sum_{i=1}^{j} y_{(i)}, \quad j = 1, \cdots, n-1,$$

并且$\displaystyle\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$;

(ii) $\boldsymbol{x}$以p-larger序大于$\boldsymbol{y}$(记为$\boldsymbol{x} \overset{\text{p}}{\succeq} \boldsymbol{y}$), 如果

$$\prod_{i=1}^{j} x_{(i)} \le \prod_{i=1}^{j} y_{(i)}, \quad j = 1, \cdots, n.$$

容易验证$\boldsymbol{x} \overset{\text{m}}{\succeq} \boldsymbol{y} \Rightarrow \boldsymbol{x} \overset{\text{p}}{\succeq} \boldsymbol{y}$.

定义 2.3: 设$\phi(\boldsymbol{x})$为定义在集合$\mathcal{A} \subseteq \mathcal{R}^n$上的实值函数,

(i) $\phi(\boldsymbol{x})$称为Schur-convex函数, 若

$$\boldsymbol{x} \overset{\text{m}}{\preceq} \boldsymbol{y} \Longrightarrow \phi(\boldsymbol{x}) \le \phi(\boldsymbol{y});$$

(ii) $\phi(\boldsymbol{x})$称为Schur-concave函数, 若

$$\boldsymbol{x} \overset{\text{m}}{\preceq} \boldsymbol{y} \Longrightarrow \phi(\boldsymbol{x}) \geq \phi(\boldsymbol{y}).$$

引理 2.1 (Marshall et al. (2011)): 定义在集合$\mathcal{A} \subseteq R$上的对称可微函数$\phi(\boldsymbol{x})$是Schur-convex的, 当且仅当对任意$\boldsymbol{x} \in \mathcal{A}$与$i \neq j$, 有

$$(x_i - x_j)\Big(\frac{\partial \phi(\boldsymbol{x})}{\partial x_i} - \frac{\partial \phi(\boldsymbol{x})}{\partial x_j}\Big) \geq 0.$$

关于majorization序和Schur-convex函数的更多理论结果, 可参阅Bon和Păltănea (1999)和Marshall et al. (2011).

Copula由Sklar (1959)首次引入刻画变量之间的相依性, 定义如下:

定义 2.4 (Sklar (1959)): 二维的copula C是一个定义域为$[0,1]^2$, 值域为$[0,1]$的递增, 右连续函数, 并且满足以下条件:

(i) $\lim\limits_{u_i \to 0} C(u_1, u_2) = 0, i = 1, 2$;

(ii) $\lim\limits_{u_1 \to 1} C(u_1, u_2) = u_2$, $\lim\limits_{u_2 \to 1} C(u_1, u_2) = u_1$;

(iii) C是supermodular, 即对任意的$u_1 \leq v_1, u_2 \leq v_2$下列不等式成立

$$C(v_1, v_2) - C(u_1, v_2) - C(v_1, u_2) + C(u_1, u_2) \geq 0.$$

由上面定理我们知道copula实际上是一个边际为均匀分布的二维向量的联合分布. 下面的引理说明了二维随机向量下边际分布, 联合分布, copula三者之间的关系.

引理 2.2: 设$\boldsymbol{X} \in \mathcal{R}_2(F_1, F_2)$有连续的边际分布函数$F_1$与$F_2$. 则存在唯一的copula函数$C$, 对所有的$\boldsymbol{x} \in \mathcal{R}^2$有

$$F_{\boldsymbol{X}}(x_1, x_2) = C(F_1(x_1), F_2(x_2)).$$

相反的, 如果C是一个copula, F_1和F_2分别表示两个随机变量的分布函数, 那么我们有上式定义的$F_{\boldsymbol{X}}$是一个以F_1和F_2为边际的二维分布函数.

以下列出在实际应用与理论分析中常见的几个copula:

- 独立copula $C_I(u, v) = uv$;
- Frechet下界copula $C_L(u, v) = \max\{0, u + v - 1\}$;
- Frechet上界copula $C_U(u, v) = \min\{u, v\}$;
- Clayton copula

$$C_1(u, v) = (u^{-\alpha} + v^{-\alpha} - 1)^{-\frac{1}{\alpha}}, \quad \alpha > 0;$$