学校编码：10384　　　　　　　　　　　　　　分类号 ______ 密级 ______
学号：200423041　　　　　　　　　　　　　　UDC ____________________

厦　门　大　学

硕　士　学　位　论　文

保险中的相依性研究

Research of dependence in insurance

杨　雪

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: 2007 年 5 月

论文答辩日期: 2007 年 　月

学位授予日期: 2007 年 　月

答辩委员会主席: ______________

评　　阅　　人: ______________

2007 **年** 5 **月**

厦门大学学位论文原创性声明

兹呈交的学位论文，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的其它个人或集体的研究成果，均在文中以明确方式标明．本人依法享有和承担由此论文而产生的权利和责任．

声明人（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定. 厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版．保密的学位论文在解密后适用本规定.

本学位论文属于

1、保密（　　），在　　　年解密后适用本授权书.

2、不保密（　　）.

（请在以上相应括号内打"✓"）

作者签名：　　　　　　　　日期：　　　年　月　日

导师签名：　　　　　　　　日期：　　　年　月　日

目 录

Contents

摘　　要

对于保险公司来说，人们更关心地是保单组合在未来 n 年内的聚合赔付情况. 在经典的保险精算学中为了应用大数定律往往假设风险之间是相互独立的. 而实际上由于受一些共同因素的影响，风险之间存在着或多或少的关系，也就是说风险之间常常具有相依性，而且通常都具有较强的正相依性. 比如，在地震或者洪水灾害以及传染病中，考虑赔付时风险之间就存在着相依性，在大雾的天气里汽车的赔付率也互相影响；在夫妻险中，丈夫和妻子的赔付也有很密切的关系等等. 在这种情况下传统的方法就显得无能为力. 因此，探索相依风险下的精算问题就成了近十余年精算界的热点问题. 本文在前人研究的基础上讨论了相依结构下停止损失纯保费的近似算法及 Copula 函数在两重生命风险模型下的应用. 具体内容如下：

第一章, 介绍了相依风险下保险精算在国内外的研究概况. 第二章, 给出了必要的基础知识. 第三章对相依结构下停止损失纯保费的近似算法进行了研究，并在本章末给出数值例子. 第四章讨论了 Copula 函数在两重生命风险模型下的应用.

关键词: Copula; 相依风险; 同单调.

Abstract

Insurance company usually pay more attention to aggregate claim in the future n years. In traditional actuarial theory, The Law of Large Numbers will be hold under the assumption of independent, But actually due to some common influnce, risks are usually related to each other more or less. For example, the individual risks of earthquake or flooding risks or epidemic portfolio are dependent; on a foggy day all cars of a region have higher probability to be involved in an accident. In life insurance, there is ample evidence that the living time of husbands and their wives are positively associated. So actuarial specialists expect to find effective methods to describe dependent risks.In this thesis, we find convex-order for stop-loss premims, and discuss applications of copula in two-life model.

In chapter 1, we present the research of dependent risks in and abroad. In chapter 2, we introduce basic definitions and properties used in the sequel; In chapter 3, we study the method to approximate the stop-loss premiums of dependent risks, it is an application of comonotonicity theory, then we give numerical example; In chapter 4, We study applications of comonotonicity and copula in two-life model.

Key words: copula; dependent risks; comonotonicity.

第一章　引　言

传统的风险理论在涉及多个风险时往往假定他们是相互独立的，因为有了这一假设，大数定律才有了应用的前提，从而保险公司通过增加保单来有效的管理风险.

实际上由于受一些共同因素的影响，风险之间往往存在着较强的相依性. 比如，在财产保险中，如果在某一地区或组织中所承保的风险密度过大，如冰雹、爆炸、地震、流行病等巨灾事故发生时，会导致风险的扩散，从而引起赔付率的增加. 在人寿保险中的夫妻险中，由于他们属于同一社会阶层，有相似的生活环境和生活方式，丈夫和妻子的寿命具有正相关性，并且一般来说，配偶死后，另一位的死亡率也会上升；另一个例子就是养老金险，它一般以一个公司中的员工为一个保单组合，这些人处在相同的工作环境里，很明显他们的寿命在一定程度上是相关的.

因此独立性假设在很多时候是不符合实际的，寻求描述风险之间相依性的有效办法成为精算界的热点问题. 如果可以假定某个确定的相依函数，通过这个相依函数用已知的边际分布构造联合分布，这是最彻底的一种方法，因为联合分布一旦确定，则所有的性质都可以通过分析联合分布函数来得到. 同单调就是这样的一种方法，这种方法的另一个典型代表是连接函数 Copula.

同单调是一种特殊的相依结构，最早由 Yarri(1987) 和 Schmeidler(1986) 引入，在数学处理上，它和独立性假设有相同的优点，即在已知边际分布时，它的联合分布可以确定，并且在形式上比较简单. 同时在一定意义下，它又是最强的一种正相依结构. 尽管如此，同单调理论在 20 世纪 90 年代才开始引起精算界的重视，例如 Wang(1998)[27], Denuit(1999)[5] 和 Goovaerts(2002)[17], 他们对同单调性的研究往往是和随机序与保费原理等结合在一起的. 2002 年，Dhaene, Denuit, Goovaerts, Kaas 和 Vyncke(2002a,b)[7,8] 分理论和应用总结了关于同单调随机变量的一些最新研究成果，考虑的模型为 $S=\sum\limits_{i=1}^{n}X_i$, 其中 X_i 表示某保单组合中第 i 张保单的理赔量. Tom Hoedemakers(2005)[26] 将模型扩展为两个随机向量的积，即 $S=\sum\limits_{i=1}^{n}X_iY_{t_i}$.

其中 X_i 表示在时刻 $i=t_i,(i=1,2,,\cdots,n)$ 的随机损失， Y_t 是在时间区间 [0,t] 上的贴现函数.

1959 年 Abe Sklar 在 Sklar 定理中第一次把 Copula 一词引入统计学中； Paul Embrechts(2001)[23] 系统的总结了这个领域的主要研究成果，讨论了 Copula 的基本性质; Frees Valdez(1998)[15] 着重讨论了 Copula 作为相关结构度量工具的优点，并且给出 Copula 在实际中的应用； Matteis(2001)[21] 详细介绍了 Archimedean Copulas 在数据建模中的应用，并运用 Copula 对丹麦火灾险损 失进行了度量，并且对再保险的信用风险和投资风险进行了分析；Bouy$\acute{e}$(2001)[3] 对 Copula 函数在风险管理中的应用问题进行了比较深入的讨论. 在精算领域中 Copula 主要被用来构建死亡率和损失率之间的相关关系， Frees, Carriere, Vaklez(1996)[14] 和 Frees Valdez(1998)[15] 以及 Free, Wang(2005)[16] 分别进行了这方面的研究.

本文在以上工作的基础上, 研究精算模型中的风险相依性问题, 主要完成了以下几个方面的工作:

第二章介绍了基础知识, 首先给出有关相依风险的一些基本概念: 逆分布函数、随机序和同单调, 然后对 Copula 的性质做了系统的总结. 第三章对相依结构下停止损失纯保费的近似算法进行了研究. 第四章主要讨论了同单调和 Copula 函数在两重生命模型中的应用：运用 Copula 函数的一些性质，得到了几种特殊的趸缴保费和年金的计算方法；在被保险人剩余寿命为同单调时，得到了联合生存状态和最后生存状态的剩余寿命的分布；最后给出数值例子，从数据的分析中看到了独立性假设的不合理性.

第二章 预备知识

§2.1 逆分布函数

若随机变量 X 的分布函数 $F_X(x)=P[X\le x]$ 是非降右连续的，且

$$F_X(-\infty)=\lim_{x\to-\infty}F_X(x)=0, \tag{2.1}$$

$$F_X(+\infty)=\lim_{x\to+\infty}F_X(x)=1. \tag{2.2}$$

下面来定义逆分布函数．约定 $\inf\Phi=+\infty$, $\sup\Phi=-\infty$.

定义 **2.1** 记

$$F_X^{-1}(p)=\inf\{x\in R|F_X(x)\ge p\},\ p\in[0,1], \tag{2.3}$$

$$F_X^{-1+}(p)=\sup\{x\in R|F_X(x)\le p\},\ p\in[0,1]. \tag{2.4}$$

并对任意 $\alpha\in[0,1]$, 记

$$F_X^{-1(\alpha)}(p)=\alpha F_X^{-1}(p)+(1-\alpha)F_X^{-1+}(p),\ p\in(0,1) \tag{2.5}$$

则称 $F_X^{-1(\alpha)}(p)$ 为 $F_X(x)$ 的 $\alpha-$ 混合逆分布函数．显然

$$F_X^{-1(0)}(p)=F_X^{-1+}(p),$$

$$F_X^{-1(1)}(p)=F_X^{-1}(p),$$

并对任意 $\alpha\in[0,1]$,

$$F_X^{-1}(p)\le F_X^{-1(\alpha)}(p)\le F_X^{-1+}(p). \tag{2.6}$$

事实上，我们通常所说的逆分布函数是指 $F_X^{-1}(p)$.

现在令 d 满足 $0 < F_X(d) < 1$, 则 $F_X^{-1}(F_X(d))$ 和 $F_X^{-1+}(F_X(d))$ 是有限的，并且 $F_X^{-1}(F_X(d)) \leq d \leq F_X^{-1+}(F_X(d))$. 那么总存在一个 $\alpha_d \in [0,1]$, 使得

$$d = \alpha_d F_X^{-1}(F_X(d)) + (1-\alpha_d)F_X^{-1+}(F_X(d)) = F_X^{-1(\alpha_d)}(F_X(d)). \tag{2.7}$$

性质 **2.1**

$$F_X^{-1}(p) \leq x \Leftrightarrow p \leq F_X(x). \tag{2.8}$$

性质 **2.2** $F_X^{-1}(p)$ 是 p 的非降左连续函数; $F_X^{-1+}(p)$ 是 p 的非降右连续函数.

性质 **2.3** 设 X 和 g(X) 均为实值随机变量, $0 < p < 1$,

(a). 如果 g 是非降左连续的，则 $F_{g(X)}^{-1}(p) = g(F_X^{-1}(p))$;

(b). 如果 g 是非降右连续的，则 $F_{g(X)}^{-1+}(p) = g(F_X^{-1+}(p))$;

(c). 如果 g 是非增左连续的，则 $F_{g(X)}^{-1+}(p) = g(F_X^{-1}(1-p))$;

(d). 如果 g 是非增右连续的，则 $F_{g(X)}^{-1}(p) = g(F_X^{-1+}(1-p))$.

§2.2 同单调

同单调 (Comonotonicity) 的概念是经济领域一个重要的概念, 下面给出它的定义和主要性质.

定义 **2.2** 如果对 n 维向量集 A 中的任意两个元素 $\underline{x}$ 和 $\underline{y}$, 要么 $\underline{x} \leq \underline{y}$, 要么 $\underline{y} \leq \underline{x}$, 则称集合 $A \subseteq R^n$ 是同单调的.

定义 **2.3** 如果 $P\{X \in B\} = 1$, 且对任意非零测集 $B_1 \subset B$, $P\{X \in B_1\} < 1$, 则称集合 B 为 X 的最小支撑.

性质 **2.4** 随机向量 $X = (X_1, X_2, \cdots, X_n)$ 同单调，当且仅当下述等价条件之一成立:

(a). X 的支撑同单调;

(b). 对所有的 $x = (x_1, x_2, \cdots, x_n)$, X 的联合分布函数为

$$F_X(x) = \min\{F_{X_1}(x_1), F_{X_2}(x_2), \cdots, F_{X_n}(x_n)\};$$

(c). 对 $U \sim U(0,1)$, 有

$$X \stackrel{\mathrm{d}}{=} (F_{X_1}^{-1}(U), F_{X_2}^{-1}(U), \cdots, F_{X_n}^{-1}(U));$$

(d). 存在随机变量 Z 和非降函数 f_i, $(i = 1, 2, \cdots, n)$ 使得

$$X \stackrel{\mathrm{d}}{=} (f_1(Z), f_2(Z), \cdots, f_n(Z)).$$

§2.3 凸序

描述风险的随机向量是不能直接比较大小的，因此只能对风险在特定意义下进行优劣比较. 统计学家和精算学家一般在随机序的意义下对风险进行比较. Shaked 等 (1994) 对各种随机序做了详尽的论述. 这里我们给出其中的一种 – 凸序的定义和性质. 本节所说的随机变量都是非负随机变量，并且假定期望和方差都存在.

定义 **2.4** 如果对于任意使得 $E[f(X)]$ 和 $E[f(Y)]$ 存在的凸函数 f 有

$$E[f(X)] \leq E[f(Y)]$$

则称 X 在凸序意义下小于 Y, 记作 $X \leq_{cx} Y$.

凸序也有几个等价定义，我们把这些等价定义罗列如下.

性质 **2.5** (Shaked(1994) 和 Dhaene(2002a)) 下面的结论是等价的:

(a). $X \leq_{cx} Y$;

(b). $-X \leq_{cx} -Y$;

(c). E[X] =E[Y], 并且对任意的 $d \in R$ 有 $E[(X-d)_+] \leq E[(Y-d)_+]$;

(d). E[X] =E[Y], 并且对任意的 $d \in R$ 有 $E[(d-X)_+] \leq E[(d-Y)_+]$;

(e). 对任意的 $d \in R$ 有 $E[(X-d)_+] \leq E[(Y-d)_+]$, $E[(d-X)_+] \leq E[(d-Y)_+]$;

(f). 对任意使得 $E[g(-X)]$ 和 $E[g(-Y)]$ 存在的凹函数 g, 有 $E[g(-X)] \geq E[g(-Y)]$.

在效用理论里, 效用函数是表示收入的函数, 对于保守型决策者来说它是一个非降的凹函数, 这体现了期望效用递增和边际效用递减的思想; 与效用函数相对应的损失函数是非降凸函数, 即对损失的厌恶程度是递增并且是加速递增的. 如果随机变量 X 和 Y 表示损失, 则 $-X$ 和 $-Y$ 表示收益, 由定义 2.4, 凸序 $X \leq_{cx} Y$ 可以理解为, 保守型决策者都选择承受风险 X 而不是风险 Y, 由性质 2.5(f) 保守型决策者都会选择接受收益 $-X$ 而不是收益 $-Y$, 因此保守型的效用函数和损失函数是一一对应的.

§2.4 **连接函数** (Copula)

§2.4.1 Copula **函数定义、性质**

Copula 被 Sklar(1959) 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. 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 $\overline{a} = (a_1, a_2, \cdots, a_n)$,$\overline{b} = (b_1, b_2, \cdots, b_n)$. 并且对任意的 k 有 $a_k \leq b_k$. 令 $B = [a_1, b_1] \times \cdots \times [a_n, b_n]$,B 的所有顶点都在 $DomH$ 内. $\overline{c} = (c_1, c_2, \cdots, c_n)$ 表示 B 的顶点, 定义

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

其中

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

(1) 如果 $V_H(B) \geq 0$ 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 $\overline{t} = (t_1, t_2, \cdots, t_n) \in DomH$, 若存在一个 $t_k = a_k$ 便有 $H(\overline{t}) = 0$ 成立, 那么则称 H 是基础的.

定义 2.5 Copula 是多元分布函数，其边缘分布是定义在 [0,1] 上的均匀分布 (即 U(0,1)), 以符号 C 表示，并满足以下三个条件：

(a). $C : [0,1]^n \to [0,1]$;

(b). C 是基础的、 n 元增函数;

(c). 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 $u \in [0,1]$, $i = 1, 2, \ldots, n$.

性质 2.6 $M^n(u)$, $W^n(u)$ 分别是定义在 $[0,1]^n$ 上的函数,

$$M^n(u) = \min(u_1, u_2, \cdots, u_n),$$

$$W^n(u) = \max(u_1 + u_2 + \cdots + u_n - n + 1, 0),$$

则 $M^n(u)$ 在 $n \geq 2$ 时为 Copula 函数，则 $W^n(u)$ 在 $n \geq 3$ 时不一定是 Copula 函数. 若 $C(u)$ 为 $[0,1]^n$ 上的 n-Copula, 则有

$$W^n(u) \leq C(u) \leq M^n(u).$$

性质 2.7 若 $F(\cdot)$ 是一个联合分布函数，其边缘分布函数为 $F_1, \ldots, F_n$, 那么一定存在一个Copula 函数 C, 使得:

$$F(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)).$$

如果边缘分布函数是连续的，那么 Copula 形式是唯一的.

假如 $F_1, \ldots, F_n$ 是单变量的累积分布函数，则 $C(F_1, \ldots, F_n)$ 是表示一多变量的累积分布函数，其边缘分布函数为 $F_1, \ldots, F_n$. 由此可见 Copula 是一个把边缘分布函数连起来构造联合分布的函数.

性质 2.8 随机向量 $(X_1, X_2, \cdots, X_n)$ 相互独立当且仅当

$$C(F_1(x_1), \cdots, F_n(x_n)) = F_1(x_1) \cdot F_1(x_2) \cdots F_n(x_n).$$

§2.4.2 基于 Copula 的相关性度量

定义 **2.6** $Kendall$ 相关系数

设 (X_1,Y_1) 与 (X_2,Y_2) 是随机向量 (X,Y) 的样本， (X_1,Y_1) 与 (X_2,Y_2) 的和谐概率减去不和谐概率的差定义为随机向量 (X,Y) 的 $Kendall$ 相关系数 τ

$$\tau(X,Y)=P[(X_1-X_2)(Y_1-Y_2)>0]-P[(X_1-X_2)(Y_1-Y_2)<0]$$

性质 **2.9** 用 Copula 函数表示的 $Kendall$ 相关系数 τ

$$\tau(X,Y)=4\int\int_{[0,1]^2}C(u_1,u_2)dC(u_1,u_2)-1,$$

性质 **2.10** 随机向量 (X,Y) 的边际分布连续，$Copula$ 函数为 C, 则 $Kendall$ 相关系数具有以下性质：

(a). $|\tau(X,Y)|\le 1$.

(b). $\tau(X,Y)=1\Leftrightarrow C=M^2$.

(c). $\tau(X,Y)=-1\Leftrightarrow C=W^2$.

(d). 如果 X,Y 互相独立，则 $\tau(X,Y)=0$.

(e). X,Y 同单调 $\Leftrightarrow\tau(X,Y)=1$.

其中 W^2,M^2 的定义同性质 2.6.

§2.4.3 **常用的二元** Copula

1. 定义 **2.7** 椭球族 Copula(Ellipitical Copulas)

假设 F 是一个椭球分布的多元累积分布函数， F_i 表示其边缘分布， F_i^{-1} 是边缘分布函数的逆函数. 那么由 F 确定的连接函数椭球 Copula 形式为

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

(1) Normal Copula

Normal Copula 又称 Gaussian Copula, 是多元正态分布的连接函数. 假设 $X=(X_1,X_2,\cdots,X_n)$ 服从多元正态分布，当且仅当其边缘分布函数 $F_1,F_2,\cdots,F_n$ 皆