

A Sort of Multiple Life Model under Dependent Causes of Decrement

Cui Jiang¹, Dabuxilatu Wang^{2,3*}

¹School of Mathematics and Information Sciences, Guangzhou University, Guangzhou Guangdong

²School of Economics and Statistics, Guangzhou University, Guangzhou Guangdong

³Lingnan Research Center for Statistical Science, Guangzhou University, Guangzhou Guangdong

Email: *wangdabu@gzhu.edu.cn

Received: Aug. 17th, 2015; accepted: Sep. 5th, 2015; published: Sep. 8th, 2015

Copyright © 2015 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In this paper, both dependencies in decrement causes and in lives in life insurance are depicted with the copula functions, and the joint survival functions of the multiple life are compared upon the cases of independence as well as dependence. A general symmetric status of multiple life under dependent causes of decrement is considered, and some influence caused by the dependence between the individuals on the survival functions of the general symmetric status is explained.

Keywords

Decrement Causes, Multiple Life, Survival Function, Copula Function

一类相依衰减因多生命模型

蒋 翠¹, 王达布希拉图^{2,3*}

¹广州大学数学与信息科学学院, 广东 广州

²广州大学经济与统计学院, 广东 广州

³广州大学岭南统计研究中心, 广东 广州

Email: *wangdabu@gzhu.edu.cn

收稿日期: 2015年8月17日; 录用日期: 2015年9月5日; 发布日期: 2015年9月8日

*通讯作者。

摘要

本文利用Copula函数刻画寿险中衰减因间以及个体之间的相依性, 比较个体间独立情形与相依情形的联合生存分布。考虑相依多生命的一般对称状态在相依衰减因下的生存分布, 并解释个体间相依性对一般对称状态的生存分布的影响。

关键词

衰减因, 多生命, 生存分布, Copula函数

1. 引言

在一个由多个个体组成的多生命生存模型中, 每个个体的生存时间往往会受到相同的经济环境或者自然环境的影响(例如地震、疾病等), 而这样就使得个体间的生存时间不再是相互独立的, 例如夫妻双方的寿命是正相依的[1]。另一方面, 在多元衰减模型中, 各衰减因可能是完全相依的, 从而导致各衰减时间完全相依, 譬如: 伤残与死亡之间的相依关系。因此, 在生存分析中, 研究各个衰减因之间的关系以及每个个体之间的关系时, 经常会碰到类似上述的双重相依关系。对此, 学者们多用相依竞争风险理论或多元衰减模型分别研究。

1994年, Carriere 基于相依竞争风险理论提出了相依衰减理论[2], 在该文中, 用 Copula 函数刻画了单生命个体受相依的多元衰减因影响的生存模型, 利用粗生存分布函数与净生存分布函数的非线性微分方程组, 通过粗生存分布得到了净生存分布的一个非参数估计。2007年, Kaishev 等用 Copula 函数模拟了竞争风险生存时间的联合分布函数, 分别用不同的 Copula 函数刻画了各风险之间的相依关系, 并且在计算技术上作了很大的改进[3]。除了用 Copula 方法模拟竞争风险之间的相依性, 还有许多文献试图解释个体在竞争风险下的生存时间的潜在机制[4] [5]。

在经典的多元衰减模型中, 为了计算方便, 往往假定各衰减因的衰减时间是独立的[6]。然而, 在实际情况中, 个体与个体之间以及各衰减因之间往往存在相依的情形。基于此, 本文将运用 Copula 方法探讨相依个体与相依衰减的多生命多减因衰减模型。

2. 基本概念

对于单生命的相依多元衰减模型, Carriere [2]和 Kaishev 等[3]在文中都有深刻的讨论。假设个体的寿命 $x \geq 0$ 的终止受制于 m 种原因中的任一个, 即个体的潜在寿命是 T_1, T_2, \dots, T_m , $0 \leq T_j < \infty$, $j = 1, \dots, m$ 。显然, 实际衰减时间是所有潜在时间 T_1, T_2, \dots, T_m 当中的最小者, 即 $\min \bar{T} := \min \{T_1, \dots, T_m\}$, 其可观测, 但每一 T_1, T_2, \dots, T_m 却不可观测。正如前所述, 实际情形中, 作用于个体的衰减因往往相依, 从而考虑随机变量 T_1, T_2, \dots, T_m 相依。而又如 Carriere [2]文中所提, 虽然生物统计学家和精算学家常会考虑随机变量 T_j 删失, 但这个假设会导致一些棘手的理论问题, 基于这样的事实, 假设 T_j 非删失, 即 $P(T_j < \infty) = 1$ 。再设 (T_1, T_2, \dots, T_m) 的联合分布 $F(t_1, \dots, t_m) = P(T_1 \leq t_1, \dots, T_m \leq t_m)$ 绝对连续, 则对应的生存联合分布 $S(t_1, \dots, t_m) = P(T_1 > t_1, \dots, T_m > t_m)$ 也是绝对连续。一个年龄为 $x \geq 0$ 的个体的总体生存分布函数定义为

$$\mathfrak{S}(t) = S(t, \dots, t) = P(T_1 > t, \dots, T_m > t) = P(\min(T_1, \dots, T_m) > t)$$

粗生存分布函数定义为

$$S^{(j)}(t) = P(\min(T_1, \dots, T_m) > t, \min(T_1, \dots, T_m) = T_j)$$

净生存分布函数定义为

$$S^{(j)}(t) = P(T_j > t)$$

其为单独考虑衰减因 j 作用于 (x) 个体在 t 时刻之前仍未衰减的概率, 用精算符号表示为 ${}_t p_x^{(j)}$, 为了说明方便, 暂用精算符号表示净生存概率, 对 $0 < h < t$ 有关系式

$${}_t p_x^{(j)} = P(T_j(x) > t) = P(T_j(x) > h)P(T_j(x) > t | T_j(x) > h) = {}_h p_x^{(j)} {}_{t-h} p_{x+h}^{(j)} \quad (1)$$

注意, 净生存分布不可直接观测, 而所能观测的是 $\min(T_1, \dots, T_m)$, 因此, 粗生存分布是可观测的。

由以上定义可知, $S^{(j)}(t)$ 为总体生存分布函数 $S(t_1, \dots, t_m)$ 的边缘分布函数, 亦称为边缘生存分布函数。

2.1. Copula 函数

众所周知, 联合分布函数包含了随机变量间的相依信息。而 Copula 函数将这种复杂的相依关系用边缘变量的信息完整表现出来。

定义 2.1.1 Copula 是一个的函数, 且满足

- 1) $C(t_1, \dots, t_{k-1}, 0, t_{k+1}, \dots, t_n) = 0, \forall k \in \{1, 2, \dots, n\}$;
- 2) $\Delta_{a_n}^{b_n} \Delta_{a_{n-1}}^{b_{n-1}} \dots \Delta_{a_1}^{b_1} C(t_1, t_2, \dots, t_n) \geq 0, \forall 0 \leq a_i \leq b_i \leq 1$, 其中 $\Delta_{a_i}^{b_i}$ 表示的是一阶差分, 即

$$\Delta_{a_i}^{b_i} C(t_1, t_2, \dots, t_n) = C(t_1, \dots, t_{i-1}, b_i, t_{i+1}, \dots, t_n) - C(t_1, \dots, t_{i-1}, a_i, t_{i+1}, \dots, t_n)$$

- 3) $C(1, \dots, 1, t_k, \dots, 1) = t_k, \forall k \in \{1, 2, \dots, n\}$.

引理 2.1.1 (Sklar 定理(Nelsen) [6]) 设 $F(x_1, \dots, x_n)$ 是 n 维向量的联合分布函数, 其边缘分布函数分别为 $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$, $(x_1, \dots, x_n) \in R^n$, 则必存在一个 Copula 函数 C , 使得对所有的 $(x_1, \dots, x_n) \in R^n$ 都有

$$F(x_1, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n))$$

若边缘分布函数 $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ 都是连续的, 那么 C 是唯一确定的, 否则 C 在 $RanF_1 \times RanF_2 \times \dots \times RanF_n$ 上是不唯一确定的。反之, 若 C 是一个 Copula 函数, $F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ 是分布函数, 则由上式所定义的 $F(x_1, \dots, x_n)$ 是一个联合分布函数, 其边缘分布分别是

$F_1(x_1), F_2(x_2), \dots, F_n(x_n)$ 。

由此可知, Copula 函数包含了随机变量的所有相依信息, 使得处理随机变量信息更加容易, 至少在确定 Copula 函数后可知各随机变量之间的相依结构。另一方面, 对于不同类型的边缘分布函数, 可以有相同的 Copula 函数, 使得处理随机变量相依结构更加灵活。

Copula 函数包括很多分布族, 其中椭球分布族(例如多元的正态 Copula, t-Copula 等)和 Archimedean 分布族(例如 Clayton Copula, Frank Copula 等) [6] 是常见的两个分布族, 本文将针对 Archimedean 分布族中的两类 Copula 函数进行讨论。

2.2. 生存 Copula

对于生存 Copula C , 它是连接联合生存分布函数和边缘生存分布函数的 Copula 函数。与之对应的, 有类似于 Sklar 定理的表述。因为对 $(u_1, \dots, u_n) \in [0, 1]^n$ 有

$$\begin{aligned}
C(u_1, \dots, u_n) &= P\left(\bigcap_{i=1}^n \{S_{X_i} \leq u_i\}\right) \\
&= P\left(X_1 \geq S_1^{-1}(u_1), \dots, X_n \geq S_n^{-1}(u_n)\right) \\
&= S\left(S_1^{-1}(u_1), \dots, S_n^{-1}(u_n)\right)
\end{aligned}$$

其中, $S_i := S_{X_i}$, $S_i(x_i) = S^{(i)}(x_i) = u_i$, $u_i \sim U(0,1)$ 。因此, 有如下的结论:

引理 2.2.1 (Carriere [2], Kaishev 等[3]) 设 $S(x_1, \dots, x_n)$ 为联合生存分布函数, $S^{(1)}(x_1), \dots, S^{(n)}(x_n)$ 为其边缘分布函数且连续, 则存在唯一的生存 Copula C 使得

$$S(x_1, \dots, x_n) = C\left(S^{(1)}(x_1), \dots, S^{(n)}(x_n)\right) \quad (2)$$

引理 2.2.2 (Carriere [2]) 粗生存分布函数与联合生存分布函数之间有关系式:

$$S^{(j)}(t) = \int_t^{\infty} S_j(r, \dots, r) dr \quad (3)$$

其中 $S_j(r, \dots, r) = \frac{\partial}{\partial t_j} S(t_1, \dots, t_n) \Big|_{t_k=r, \forall k}$ 。

由引理 2.2.1 与引理 2.2.2 可得下列由非线性微分方程组表达的粗生存分布函数与净生存分布函数之间的关系式[3]:

$$\begin{cases} \frac{d}{dt} S^{(1)}(t) = C_1\left(S^{(1)}(t), S^{(2)}(t), \dots, S^{(n)}(t)\right) \times \frac{d}{dt} S^{(1)}(t) \\ \frac{d}{dt} S^{(2)}(t) = C_2\left(S^{(1)}(t), S^{(2)}(t), \dots, S^{(n)}(t)\right) \times \frac{d}{dt} S^{(2)}(t) \\ \vdots \\ \frac{d}{dt} S^{(n)}(t) = C_n\left(S^{(1)}(t), S^{(2)}(t), \dots, S^{(n)}(t)\right) \times \frac{d}{dt} S^{(n)}(t) \end{cases} \quad (4)$$

其中 $C_j(\mu_1, \dots, \mu_n) = \frac{\partial}{\partial \mu_j} C(\mu_1, \dots, \mu_n)$ 。

3. 一类相依衰减因下的多生命一般对称状态

本节考察相依多生命的一般对称状态在相依衰减因下的生存分布, 并用生存 Copula 函数模拟个体间及衰减因间的相依性。

对于由 n 个独立个体组成的群体, 可通过剩余未衰减时间 T_u 来定义状态 u , 记 ${}_t p_u$ 为在 0 时刻状态 u 已存在, 且在时刻 t 状态 u 仍存在的条件概率。设 $u(k)$ 表示一般对称状态[7], 即至少还有 k 个个体没有退出, 也就是第 $(n-k+1)$ 个个体退出时状态 $u(k)$ 方不存在。因此, 如果衰减因是死亡, 那么 $u(n)$ 代表联合生存状态, $u(1)$ 代表最后生存者状态。一般对称状态的未衰减概率 ${}_t p_{u(k)}$ 可以利用由一般概率加法公式推得的 Schuette-Nesbitt 公式[7]

$${}_t p_{u(k)} = \sum_{w=k}^n (-1)^{w-k} C_{w-1}^{k-1} {}_t S_w \quad (5)$$

计算, 其中 ${}_t S_w$ 表示 n 个个体中任意 w 个个体活过 t 的联合概率。文献[7]仅考虑个体之间独立的情况, 而现实中个体间可能是相依的, 同时衰减因间也可能是相依的, 故本文考虑 n 个相依个体暴露在 m 个相依衰减因的情形。为此, 需计算 n 个衰减时间变量 $\min(\bar{T}_{x_1}), \min(\bar{T}_{x_2}), \dots, \min(\bar{T}_{x_n})$ 的一般对称状态的未衰减概率 ${}_t p_{u(k)}$, 这里 $\min(\bar{T}_{x_i}) = \min\{T_1(x_i), \dots, T_m(x_i)\}$ 。记事件 $\{\min \bar{T}_{x_i} > t\}$ 为 ${}_t B_i$, $i = 1, \dots, n$, 设

$${}_t S_k := \sum_{I_k \in CL(n,k)} P({}_t B_{i_1}, {}_t B_{i_2}, \dots, {}_t B_{i_k}), \quad k=1, \dots, n$$

其中, $I_k := \{i_1, i_2, \dots, i_k\}$, $CL(n, k)$ 表示由 $\{1, 2, \dots, n\}$ 的子集 $\{i_1, i_2, \dots, i_k | i_l \in \{1, 2, \dots, n\}, l=1, 2, \dots, k\}$ 所构成的类, $CL(n, k)$ 中含有 C_n^k 个元素, $k=1, 2, \dots, n$ 。 $P({}_t B_{i_1}, {}_t B_{i_2}, \dots, {}_t B_{i_k})$ 表示 n 个个体中任意 k 个相依个体活过 t 的联合概率, 由 Sklar 定理知, 存在某一生存 Copula 函数 C_I , 使得

$$P({}_t B_{i_1}, {}_t B_{i_2}, \dots, {}_t B_{i_k}) = C_I(S_{\min \bar{T}_{i_1}}(t), \dots, S_{\min \bar{T}_{i_k}}(t)), \quad \text{于是}$$

$${}_t S_k = \sum_{I_k \in CL(n,k)} C_I(S_{\min \bar{T}_{x_{i_1}}}(t), \dots, S_{\min \bar{T}_{x_{i_k}}}(t)), \quad k=1, \dots, n$$

综上所述可得

$${}_t P_{u(k)} = \sum_{w=k}^n (-1)^{w-k} C_{w-1}^{k-1} \sum_{J_w \in CL(n,w)} C_I(S_{\min \bar{T}_{x_{i_1}}}(t), \dots, S_{\min \bar{T}_{x_{i_w}}}(t)) \quad (6)$$

其中 $S_{\min \bar{T}_{x_{i_l}}}(t)$, ($l=1, 2, \dots, w$) 为 x_{i_l} 岁个体在 m 个相依衰减因下未衰减概率分布, $\min(\bar{T}_{x_{i_l}})$ 是 x_{i_l} 岁个体实际观测的衰减时间, 从而 $\min(\bar{T}_{x_{i_l}})$ 在 t 时刻的未衰减概率亦是 x_{i_l} 岁个体的粗生存分布在 t 时刻的值。由于假定 n 个个体的终止皆受制于 m 种相依衰减因, 此 m 种衰减因的相依性用生存 Copula 函数 C_D 刻画, 则对于 x_{i_l} 岁个体有

$$\begin{aligned} {}_t P_{\min \bar{T}_{x_{i_l}}} &:= S_{\min \bar{T}_{x_{i_l}}}(t) = P(\min\{T_1(x_{i_l}), \dots, T_m(x_{i_l})\} > t) \\ &= P(T_1(x_{i_l}) > t, \dots, T_m(x_{i_l}) > t) = S(t, \dots, t) \\ &= C_D(S_{T_1(x_{i_l})}(t), \dots, S_{T_m(x_{i_l})}(t)) \end{aligned}$$

其中 $S_{T_j(x_{i_l})}(t)$ 为 x_{i_l} 岁个体在第 j 个衰减因下的净生存分布, 其可通过式(4)可得。

4. 实证分析

考虑 5 个相依个体在三个相依衰减因下的一般对称状态的生存分布。Frank copula (Fc), Clayton copula (GCc) 以及 Independence copula (Ic) 函数如下:

$$1) \text{ Fc: } C(u_1, \dots, u_n) = -\frac{1}{\theta} \ln \left(1 + \frac{\prod_{i=1}^n (e^{-\theta u_i} - 1)}{(e^{-\theta} - 1)^{n-1}} \right);$$

$$2) \text{ GCc: } C(u_1, \dots, u_n) = [u_1^{-\theta} + \dots + u_n^{-\theta} - n + 1]^{-\frac{1}{\theta}};$$

$$3) \text{ Ic: } C(u_1, \dots, u_n) = \prod_{i=1}^n u_i,$$

其中, 参数 $\theta \in (0, \infty)$ 。

以文献[7]中的多元衰减表为基础, 建立暴露在工作期间死亡(d), 工作期间退职(w), 工作期间伤残(i)的 30 岁群体在整数年龄的粗生存分布函数值, 该模型中的个体在 60 岁时退休, 则工作期间退职以及工作期间伤残在 60 岁以后(包括 60 岁)将退出该模型, 因此 60 岁及之后的情形就不作讨论。假设各衰减因之间的相依结构用 Clayton Copula ($\theta=2$) 刻画, 分别用过 (0,1) 的直线拟合工作期间死亡和伤残的粗生存分布, 样条函数拟合工作期间退职的粗生存分布, 由式(4), 利用 MATLAB, 得图 1 所示的净生存分

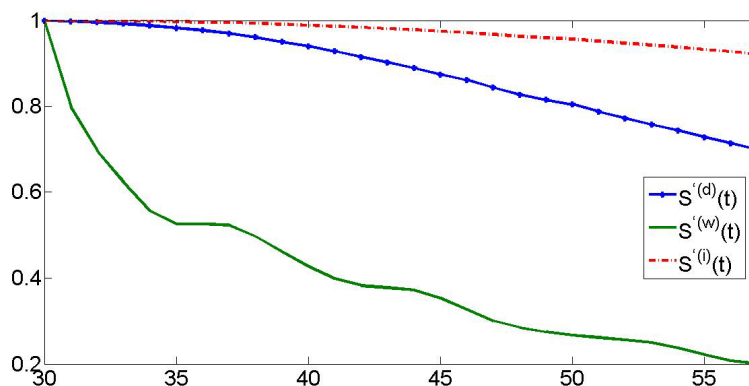


Figure 1. The net survival functions for decrement causes d, w, i under Clayton Copula

图 1. Clayton Copula 下的各衰减因 d, w, i 的净生存分布

Table 1. The values of the net survival function under the decrement causes d, w, i and $p_{\min \bar{T}_{x_j}}$ for five individuals

表 1. 五个个体的衰减因 d, w, i 净生存分布值及 $p_{\min \bar{T}_{x_j}}$

	$x_1 = 32$	$x_2 = 35$	$x_3 = 40$	$x_4 = 45$	$x_5 = 56$
$p_x^{(d)}$	0.9968	0.9941	0.9875	0.9849	0.9787
$p_x^{(w)}$	0.8947	0.9994	0.9349	0.9244	0.9672
$p_x^{(i)}$	0.9994	0.9990	0.9976	0.9968	0.9937
$p_{\min \bar{T}_{x_j}}$	0.8920	0.9925	0.9228	0.9100	0.9425

Table 2. Comparisons of probabilities $p_{u(k)} (k = 1, \dots, 5)$ under three copulas

表 2. 三种 Copula 函数下 $p_{u(k)} (k = 1, \dots, 5)$ 的比较

	$P_{u(1)}$	$P_{u(2)}$	$P_{u(3)}$	$P_{u(4)}$	$P_{u(5)}$
Fc	0.99994529	0.99853343	0.98868401	0.94109543	0.73154184
GCc	0.99990305	0.99757521	0.98355358	0.92915634	0.74961182
Ic	0.99999968	0.99994174	0.99768231	0.96148662	0.70068966

布函数。考虑该群体中分别活过 32 岁, 35 岁, 40 岁, 45 岁以及 56 岁的五生命群体, 他们各衰减因在未来一年未衰减的概率值由式(1)可得, 其值及 $p_{\min \bar{T}_{x_j}}$ 的值见表 1。

基于表 1 的结果, 分别用 Frank copula, Clayton copula 以及 Independence Copula 模拟个体间的相依性, 参数 $\theta = 2$ 。由式(5)得表 2 的结果。

由表 2 可见, 除了 $u(5)$ 的生存分布值, 其它在 Fc, GCc 相依假设下的生存分布值皆比独立假设的小, 说明通常的个体间独立假设容易低估组合风险。同时表明若五个个体同时暴露在相同的生存环境且正相依, 则只要其中一个个体衰减, 其它个体暴露在风险的可能性就会增加, 从而其相应状态下的生存概率就会减少。进一步, GCc 相依比 Fc 相依的生存分布值小。对 $u(5)$, 可见其情况与 $u(i) (i = 1, \dots, 4)$ 相反, 因为 $u(5)$ 相当于五个个体的联合生存状态, 只要有一个个体衰减, 则五个个体都暴露在衰减风险中。在此联合生存状态下 Fc, GCc 相依的生存概率大于独立情形的生存概率, 说明上述相依结构相对于独立结构更加稳固, 加强了联合状态对衰减的抵御能力。

5. 结论

由以上的讨论可知, 对于上述一类的多个衰减因共同作用下的多生命衰减模型, 联合状态下的独立性假设高估了风险, 而对于除了联合状态的一般的对称状态, 其独立性假设又会低估风险, 而且不同的相依结构其生命状态的分布也不同。因此, 设计上述一类团体保险保单时应充分考虑个体之间的相依关系。建议在不同险种的定价中, 以独立性假设为基础, 然后做相应的调整。

致 谢

作者衷心感谢国家自然科学基金项目(No. 11271096)的资助。

参考文献 (References)

- [1] Ribas, C., Marín-Solano, J. and Alegre, A. (2003) On the computation of the aggregate claims distribution in the individual life model with bivariate dependencies. *Insurance: Mathematics and Economics*, **32**, 201-215.
[http://dx.doi.org/10.1016/S0167-6687\(02\)00210-X](http://dx.doi.org/10.1016/S0167-6687(02)00210-X)
- [2] Carriere, J.F. (1994) Dependent decrement theory. *Trans. Soc. Act.*, **46**, 45-65.
- [3] Kaishev, V.K., Dimitrova, D.S. and Haberman, S. (2007) Modelling the joint distribution of competing risks survival times using copula functions. *Insurance: Mathematics and Economics*, **41**, 339-361.
<http://dx.doi.org/10.1016/j.insmatheco.2006.11.006>
- [4] Gasbarra, D., Kulathinal, S.B., Dewan, I. and Nissinen, A. (2006) Testing dependence between the failure time and failure modes: An application of enlarged filtration. *Journal of Statistical Planning and Inference*, **136**, 1669-1686.
<http://dx.doi.org/10.1016/j.jspi.2005.07.002>
- [5] Tiwari, R., Kulasekera, K.B. and Park, C. (2006) Nonparametric tests for cause specific hazard rates with censored data for competing risks among several groups. *Journal of Statistical Planning and Inference*, **136**, 1718-1745.
<http://dx.doi.org/10.1016/j.jspi.2004.12.016>
- [6] Nelsen, R.B. (2006) An introduction to copulas. 2nd Edition, Springer-Verlag, New York.
- [7] Gerber, H.U. (1997) Life insurance mathematics. Spring-Verlag.