

## 基于Copula函数的应力-强度模型可靠度的非参数估计\*

祁辉 管强

(三明学院信息工程学院, 三明, 365004)

**摘要:** 本文研究了相关的应力变量和强度变量在右删失的情形下, 应力-强度模型可靠度的非参数估计. 其中变量之间的相关关系采用常见的Farlie-Gumbel-Morgenstern copula函数和Clayton copula函数来度量. 采用经验过程的理论, 本文建立了所提出估计量的相合性及渐近正态性. 数值模拟的结果表明所提出的方法在有限样本下表现良好. 本文所提出的方法在实际中有广泛的应用前景.

**关键词:** 应力-强度模型; Kaplan-Meier估计量; 右删失; 连接函数

**中图分类号:** O213.2

**英文引用格式:** Qi H, Guan Q. Nonparametric estimation for reliability of the stress-strength model based on copula function [J]. Chinese J. Appl. Probab. Statist., 2017, 33(1): 91-101. (in Chinese)

### §1. 引言

在可靠性理论中, 应力-强度模型主要描述具备强度 $Y$ 的部件或系统在施加应力 $X$ 时的寿命. 当部件或系统强度 $Y$ 大于施加应力 $X$ 时, 部件或系统便正常工作, 否则便失效. 该模型在工程学、经济学、医学、心理学等领域有着广泛的应用<sup>[1-3]</sup>. 其在医学领域一个常见但很重要的应用就是我们一般用变量 $X$ 表示对照组的寿命, 变量 $Y$ 表示治疗组的寿命, 而模型的可靠度 $R = P(X < Y)$ 用来度量治疗效果.

对于变量 $X$ 和 $Y$ 相互独立的情形下应力-强度模型可靠度的估计问题引起了众多研究者的关注. Downton<sup>[4]</sup>得出了变量 $X$ 和 $Y$ 独立同正态分布时 $R$ 的最小无偏估计. Tong<sup>[5]</sup>基于同样的方法研究了变量 $X$ 和 $Y$ 独立同指数分布时 $R$ 的估计. Kundu和Raqab<sup>[6]</sup>考虑了独立变量 $X$ 和 $Y$ 均服从Weibull分布时 $P(X < Y)$ 的估计. Rezaei等<sup>[7]</sup>进一步研究了当独立的应力和强度变量服从不同参数Pareto分布时关于 $R$ 的估计问题. 基于MLE和Bootstrap方法, Ghitany等<sup>[8]</sup>研究了应力和强度相互独立并服从Power Lindley分布时 $R$ 的点估计和区间估计. Qi等<sup>[9]</sup>对这方面的研究情况和进展有比较全面的介绍.

在实际问题中, 应力变量 $X$ 和强度变量 $Y$ 之间存在一定的相关性. 例如, 在工程学领域, 如果应力变量 $X$ 和强度变量 $Y$ 分别表示具有相同动力装置的两个电子元件 $a$ 和 $b$ 的寿命, 则 $R = P(X < Y)$ 表示元件 $a$ 先于元件 $b$ 失效的概率. 在经济学领域存在许多类似的例子<sup>[10]</sup>,

\*国家自然科学基金青年项目(批准号: 11401341)、福建省自然科学基金项目(批准号: 2015J05014、2016J01681)和福建省教育厅A类基金项目(批准号: JA15476)资助.

本文2015年10月14日收到, 2016年4月11日收到修改稿.

例如, 如果应力变量和强度变量分别表示家庭可支配收入和消费, 则 $R$ 可以做为家庭财政状况脆弱性的衡量.

针对上述问题, 一些研究者假定相关的应力变量 $X$ 和强度变量 $Y$ 服从二元分布时, 对应力-强度模型可靠度进行了估计. Gupta和Subramanian<sup>[11]</sup>讨论了 $X$ 和 $Y$ 服从二元正态分布时关于 $R$ 的估计问题. Dieh等<sup>[12]</sup>基于排序集伴随变量抽样方法研究了上述问题. Hanagal<sup>[13]</sup>研究了在变量 $X$ 和 $Y$ 服从二元Pareto分布时关于模型可靠度的估计. Nadarajah<sup>[14, 15]</sup>分别考虑了变量 $X$ 和 $Y$ 服从二元伽码分布或贝塔分布时应力-强度模型可靠度的估计. 如果了解二元指数分布下关于 $P(X < Y)$ 的估计, 可参考文献[16]. Gupta等<sup>[17]</sup>提出了一种当变量 $X$ 和 $Y$ 服从二元正态分布时模型可靠度的大样本置信区间估计. 然而, 上述研究假定应力变量 $X$ 和强度变量 $Y$ 服从二元分布, 相当于限定变量之间具有特定的相关形式, 并且 $X$ 和 $Y$ 须同分布. 实际上, 上述假定是不现实的. 为此, Domma和Giordano<sup>[18]</sup>提出了一种较灵活的利用copula函数来度量变量 $X$ 和 $Y$ 之间相关关系的方法, 从而实现当变量服从Burr等分布时模型可靠度 $P(X < Y)$ 的估计.

在可靠性领域或临床试验中, 由于试验中途停止或财力的限制, 从而导致变量 $X$ 和 $Y$ 的样本观察值出现右删失. 在这种情况下, 对应力-强度模型可靠度 $R$ 的估计问题, 提出新的统计推断方法尤为紧迫和重要. Qi等<sup>[9]</sup>虽然提出了当变量 $X$ 和 $Y$ 右删失情况下模型可靠度 $R$ 的非参数估计方法, 但仅考虑了 $X$ 和 $Y$ 相互独立时的情形. 为此, 本文在该方法的基础上提出了一种对于相关的变量 $X$ 和 $Y$ 在右删失情况下模型 $R$ 的非参数估计方法.  $X$ 和 $Y$ 之间的相关关系采用常见的Farlie-Gumbel-Morgenstern copula函数和Clayton copula函数来度量. 上述方法尽可能多的考虑了变量 $X$ 和 $Y$ 不同的相关结构形式, 并对 $X$ 和 $Y$ 分布形式不做考虑, 因此灵活实用.

论文的组织结构安排如下: 首先我们提出了对于相关的应力变量 $X$ 和强度变量 $Y$ 在样本右删失的情形下模型可靠度 $R$ 的非参数估计方法. 接着我们对估计量的性质进行了证明. 论文第三部分对提出的方法进行数值模拟. 最后, 我们对论文的结论进行简要的讨论.

## §2. 研究方法

Copula函数提供了一种灵活的方式来度量应力变量 $X$ 和强度变量 $Y$ 之间的相关关系. 关于copula函数的详细介绍, 可参考文献[19]. 根据copula函数的理论, 对于边缘分布函数分别为 $F_X(x)$ 和 $F_Y(y)$ 的变量 $X$ 和 $Y$ , 其联合分布函数可表示为

$$H(x, y) = C(F_X(x), F_Y(y)),$$

其中 $C(\mu, \nu)$ 是度量 $X$ 和 $Y$ 相关关系的copula函数.  $(X, Y)$ 的联合概率密度函数可表示为

$$h(x, y) = c(F_X(x), F_Y(y))f_X(x)f_Y(y),$$

其中 $c(F_X(x), F_Y(y)) = \partial^2 C(F_X(x), F_Y(y)) / [\partial F_X(x) \partial F_Y(y)]$ 表示copula函数的密度函数.

对于非负变量 $X$ 和 $Y$ , 应力-强度模型可靠度可表示为

$$\begin{aligned} R &= \int_0^{+\infty} \int_x^{+\infty} h(x, y) dy dx \\ &= \int_0^{+\infty} \int_x^{+\infty} c(F_X(x), F_Y(y)) f_X(x) f_Y(y) dy dx. \end{aligned}$$

由Morgenstern<sup>[20]</sup>提出的Farlie-Gumbel-Morgenstern copula函数在统计学中有着广泛的应用, 其函数表达式为

$$C(\mu, \nu) = uv[1 + \theta(1 - \mu)(1 - \nu)], \quad -1 \leq \theta \leq 1.$$

它对应的密度函数可表示成

$$c(\mu, \nu) = 1 + \theta(1 - 2\mu)(1 - 2\nu).$$

通过FGM copula函数, 模型可靠度可表示成如下的形式:

$$\begin{aligned} R_1 &= \int_0^{+\infty} \int_x^{+\infty} [1 + \theta(1 - 2F_X(x))(1 - 2F_Y(y))] f_X(x) f_Y(y) dy dx \\ &= \int_0^{+\infty} (1 - F_Y(t)) dF_X(t) + \theta \int_0^{+\infty} (1 - 2F_X(t)) F_Y(t) (F_Y(t) - 1) dF_X(t) \\ &= R_{11} + \theta R_{12}, \end{aligned} \quad (1)$$

其中 $R_{11}$ 表示 $X$ 和 $Y$ 独立情形下的模型可靠度,  $R_{12}$ 度量了 $X$ 和 $Y$ 之间的相关关系对模型可靠度的贡献部分.

为了表达方便, 我们分别定义 $X$ 和 $Y$ 的生存函数如下:

$$G_X(t) = P(X > t) = 1 - F_X(t),$$

$$G_Y(t) = P(Y > t) = 1 - F_Y(t).$$

则式(1)可表示成

$$R_1 = - \int_0^{+\infty} G_Y(t) dG_X(t) + \theta \int_0^{+\infty} (2G_X(t) - 1)(1 - G_Y(t)) G_Y(t) dG_X(t). \quad (2)$$

考虑到变量 $X$ 和 $Y$ 删失的情况, 我们分别定义 $X$ 和 $Y$ 的右删失观测变量 $T = \min(X, C)$ 和 $Z = \min(Y, L)$ 以及删失指标函数 $\Delta = I(X \leq C)$ 和 $D = I(Y \leq L)$ . 其中, 随机变量 $C$ 和 $L$ 分别对应于 $X$ 和 $Y$ 的右删失变量. 假定 $(T_i, \Delta_i, Z_i, D_i)_{i=1}^n$ 为 $n$ 个独立同分布的随机样本. 我们可以构造如下的关于生存函数 $G_X(t)$ 和 $G_Y(t)$ 的Kaplan-Meier估计量:

$$\hat{G}_X(t) = \prod_{i=1}^n \left\{ 1 - 1 / \sum_{j=1}^n I(T_j \geq T_i) \right\}^{I(T_i \leq t) \Delta_i},$$

$$\hat{G}_Y(t) = \prod_{i=1}^n \left\{ 1 - 1 / \sum_{j=1}^n I(Z_j \geq Z_i) \right\}^{I(Z_i \leq t) D_i}.$$

将上述估计量代入式(2), 可以得到在变量右删失情况下可靠度 $R_1$ 的估计:

$$\hat{R}_1 = - \int_0^{+\infty} \hat{G}_Y(t) d\hat{G}_X(t) + \theta \int_0^{+\infty} (2\hat{G}_X(t) - 1)(1 - \hat{G}_Y(t)) \hat{G}_Y(t) d\hat{G}_X(t). \quad (3)$$

此外, Clayton copula是另一种在统计学中经常用于变量之间相关性度量的copula函数, 其函数和密度函数表达式分别为

$$\begin{aligned} C(u, v) &= (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, \quad 0 < \theta < +\infty, \\ c(u, v) &= (\theta + 1)(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta-2} u^{-(\theta+1)} v^{-(\theta+1)}, \quad 0 < \theta < +\infty. \end{aligned}$$

当变量 $X$ 和 $Y$ 之间的相关关系用Clayton copula函数度量时, 应力-强度模型可靠度为

$$\begin{aligned} R_2 &= P(X < Y) \\ &= \int_0^{+\infty} \int_x^{+\infty} (\theta + 1)(F_X^{-\theta}(x) + F_Y^{-\theta}(y) - 1)^{-1/\theta-2} \\ &\quad \cdot F_X^{-(\theta+1)}(x) F_Y^{-(\theta+1)}(y) f_X(x) f_Y(y) dy dx \\ &= \int_0^{+\infty} \{1 - F_X^{-(\theta+1)}(t)[F_X^{-\theta}(t) + F_Y^{-\theta}(t) - 1]^{-1/\theta-1}\} dF_X(t). \end{aligned} \quad (4)$$

同理, 考虑到变量右删失时的情形, 模型可靠度 $R_2$ 的估计为

$$\begin{aligned} \hat{R}_2 &= - \int_0^{+\infty} \{1 - (1 - \hat{G}_X(t))^{-(\theta+1)} \\ &\quad \cdot [(1 - \hat{G}_X(t))^{-\theta} + (1 - \hat{G}_Y(t))^{-\theta} - 1]^{-1/\theta-1}\} d\hat{G}_X(t). \end{aligned} \quad (5)$$

### §3. 估计量的渐近性质

下面我们将讨论上述可靠度的估计量的渐近性质, 为了证明的方便, 首先我们给出以下假定条件:

条件 C1 假定应力变量 $X$ 是有界的, 其取值区间为 $[0, \tau]$ .

条件 C2 存在常数 $c_0 > 0$ 使得 $P(C \geq \tau) \geq c_0$ .

条件 C3 函数 $G_X(t)$ 和 $P(C > t)$ 关于变量 $t$ 是连续的.

条件 C4 假定强度变量 $Y$ 也是有界的, 其取值区间为 $[0, \tau]$ .

条件 C5 存在常数 $l_0 > 0$ 使得 $P(L \geq \tau) \geq l_0$ .

条件 C6 函数 $G_Y(t)$ 和 $P(L > t)$ 关于变量 $t$ 连续.

首先我们证明当样本右删失时, 基于FGM copula函数的应力-强度模型可靠度的估计量 $\hat{R}_1$ 的渐近性质. 估计量 $\hat{R}_1$ 可以写成如下的形式:

$$- \int_0^{+\infty} g_2(\hat{G}_X(t), \hat{G}_Y(t)) d\hat{G}_X(t),$$

其中二元函数  $g_2(\cdot)$  是光滑的连续函数.

由上述条件, 我们可以得到如下的结论:

**引理 1** 由条件 C1–C3 可得,

$$\sup_{0 \leq t \leq \tau} |\hat{G}_X(t) - G_X(t)| \xrightarrow{P} 0.$$

由条件 C4–C6, 类似的结论对  $\hat{G}_Y(t)$  也成立. 该引理由文献[21]中引理3可得.

由引理1和连续映射定理, 可得:

$$\sup_{0 \leq t \leq \tau} |g_2(\hat{G}_X(t), \hat{G}_Y(t)) - g_2(G_X(t), G_Y(t))| \xrightarrow{P} 0.$$

**引理 2** 由条件 C1–C3 可得,  $n^{1/2}\{\hat{G}_X(t) - G_X(t)\}$  在区间  $[0, \tau]$  上弱收敛到零均值的连续路径高斯过程  $W_X(t)$ . 对于任意的  $t_1, t_2 \in [0, \tau]$ , 其协方差函数为

$$\text{Cov}\{W_X(t_1), W_X(t_2)\} = G_X(t_1)G_X(t_2) \int_0^{\min(t_1, t_2)} \frac{dH_1(u)}{\{1 - H(u)\}^2},$$

其中函数  $H(t) = P(T \leq t)$ , 函数  $H_1(t) = P(T \leq t, \Delta = 1)$ . 类似的, 引理2对估计量  $\hat{G}_Y(t)$  也成立. 该引理由文献[22]可得.

由引理2和Delta方法, 可得在区间  $[0, \tau]$  上:

$$n^{1/2}\{g_2(\hat{G}_X(t), \hat{G}_Y(t)) - g_2(G_X(t), G_Y(t))\} \xrightarrow{w} W_2(t),$$

其中  $W_2(t)$  为零均值的连续路径高斯过程.

**定理 3** 在条件 C1–C6 下, 当样本容量  $n \rightarrow \infty$  时,  $n^{1/2}(\hat{R}_1 - R_1)$  弱收敛到零均值的正态分布.

**证明:** 将  $n^{1/2}(\hat{R}_1 - R_1)$  写成如下的形式, 得:

$$\begin{aligned} & - \int_0^\tau n^{1/2} g_2(\hat{G}_X(t), \hat{G}_Y(t)) d\hat{G}_X(t) + \int_0^\tau n^{1/2} g_2(G_X(t), G_Y(t)) dG_X(t) \\ = & - \int_0^\tau n^{1/2} \{g_2(\hat{G}_X(t), \hat{G}_Y(t)) - g_2(G_X(t), G_Y(t))\} dG_X(t) \\ & - \int_0^\tau g_2(G_X(t), G_Y(t)) d[n^{1/2}\{\hat{G}_X(t) - G_X(t)\}] \\ & - n^{-1/2} \int_0^\tau n^{1/2} \{g_2(\hat{G}_X(t), \hat{G}_Y(t)) - g_2(G_X(t), G_Y(t))\} d[n^{1/2}\{\hat{G}_X(t) - G_X(t)\}]. \end{aligned}$$

由引理1和引理2, 可得:

$$\begin{aligned} & [n^{1/2}\{\hat{G}_X(t) - G_X(t)\}, n^{1/2}\{g_2(\hat{G}_X(t), \hat{G}_Y(t)) - g_2(G_X(t), G_Y(t))\}] \\ & \xrightarrow{w} \{W_X(t), W_2(t)\}. \end{aligned}$$

利用强嵌入定理, 见文献[23]. 我们可得存在一个新的概率空间, 使得:

$$\begin{aligned} & [n^{1/2}\{\widehat{G}_X(t) - G_X(t)\}, n^{1/2}\{g_2(\widehat{G}_X(t), \widehat{G}_Y(t)) - g_2(G_X(t), G_Y(t))\}] \\ & \xrightarrow{P} \{W_X(t), W_2(t)\}. \end{aligned}$$

由文献[24]中引理1, 对任意  $t \in [0, \tau]$ ,

$$\begin{aligned} & \int_0^t n^{1/2}\{g_2(\widehat{G}_X(u), \widehat{G}_Y(u)) - g_2(G_X(u), G_Y(u))\}dG_X(u) \\ & \xrightarrow{\text{a.s.}} \int_0^t W_2(u)dG_X(u), \\ & \int_0^t g_2(G_X(u), G_Y(u))d[n^{1/2}\{\widehat{G}_X(u) - G_X(u)\}] \\ & \xrightarrow{\text{a.s.}} \int_0^t g_2(G_X(u), G_Y(u))dW_X(u), \\ & \int_0^t n^{1/2}\{g_2(\widehat{G}_X(u), \widehat{G}_Y(u)) - g_2(G_X(u), G_Y(u))\}d[n^{1/2}\{\widehat{G}_X(u) - G_X(u)\}] \\ & \xrightarrow{\text{a.s.}} \int_0^t W_2(u)dW_X(u). \end{aligned}$$

因此, 在该新的概率空间下,  $n^{1/2}(\widehat{R}_1 - R_1)$  几乎处处收敛到

$$- \int_0^\tau W_2(t)dG_X(t) - \int_0^\tau g_2(G_X(t), G_Y(t))dW_X(t).$$

所以, 在原概率空间下,  $n^{1/2}(\widehat{R}_1 - R_1)$  弱收敛到零均值的正态分布. 定理得证.  $\square$

**定理 4** 在条件C1-C6下, 当样本右删失时, 基于Clayton copula函数得应力-强度模型可靠度的估计量  $\widehat{R}_2$  在样本容量  $n \rightarrow \infty$  时,  $n^{1/2}\{\widehat{R}_2 - R_2\}$  弱收敛到零均值的正态分布.

定理的证明同上, 故略.

## §4. 随机模拟

下面我们将分别研究应力变量  $X$  和强度变量  $Y$  之间的相关关系体现为FGM copula函数或者Clayton copula函数时, 在变量右删失的情形下, 模型可靠度的估计量  $\widehat{R}_1$  和  $\widehat{R}_2$  在有限样本下的统计表现.

### 4.1 基于FGM copula函数的应力-强度模型可靠度的估计量的随机模拟

为了考察估计量  $\widehat{R}_1$  在随机模拟下的表现, 我们假定应力变量  $X$  的边际分布和强度变量  $Y$  的边际分布分别服从以  $\lambda_1$  和  $\lambda_2$  为参数的指数分布. 对于参数  $\lambda_1$  和  $\lambda_2$  的取值包括相等和不等3种情况. 我们从以  $(\lambda_1, \lambda_2)$  为参数的指数联合总体中分别抽取了样本量  $n$  为50到200大

小不等的样本. 考虑到变量 $X$ 右删失, 删失变量 $C$ 由参数 $\lambda = 1$ 的指数分布生成. 类似的情形对右删失变量 $Y$ 也成立. 对于上述每种情形, 通过重复计算500次, 得到估计量 $\hat{R}_1$ 的各项统计指标, 其中包括: 估计量的估计值(Est)、估计量的方差(SD)、估计量方差的估计值(SE)、95%置信区间覆盖概率(CP). 此外, FGM Copula函数的参数 $\theta$ 取值体现变量之间相关性的差异. 因此, 通过让 $\theta$ 取较小的值 $-0.8$ 进行随机模拟, 结果体现在表1. 同时考虑 $\theta$ 取较大的值 $0.8$ , 随机模拟结果体现在表2中.

**表1** 当 $\theta = -0.8$ , 基于FGM copula函数的应力-强度模型可靠度的估计 $\hat{R}_1$ 的随机模拟结果

| $n$                             | $\lambda_1$ | $\lambda_2$ | $R$   | Est   | SD    | SE    | CP    |
|---------------------------------|-------------|-------------|-------|-------|-------|-------|-------|
| Censoring rate = (0.502, 0.500) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 1.000       | 0.500 | 0.492 | 0.069 | 0.066 | 0.924 |
| 100                             | 1.000       | 1.000       | 0.500 | 0.496 | 0.045 | 0.046 | 0.946 |
| 200                             | 1.000       | 1.000       | 0.500 | 0.497 | 0.033 | 0.033 | 0.940 |
| Censoring rate = (0.504, 0.334) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 2.000       | 0.360 | 0.358 | 0.060 | 0.059 | 0.930 |
| 100                             | 1.000       | 2.000       | 0.360 | 0.355 | 0.041 | 0.042 | 0.948 |
| 200                             | 1.000       | 2.000       | 0.360 | 0.358 | 0.029 | 0.030 | 0.944 |
| Censoring rate = (0.337, 0.497) |             |             |       |       |       |       |       |
| 50                              | 2.000       | 1.000       | 0.640 | 0.630 | 0.066 | 0.061 | 0.902 |
| 100                             | 2.000       | 1.000       | 0.640 | 0.636 | 0.043 | 0.042 | 0.938 |
| 200                             | 2.000       | 1.000       | 0.640 | 0.639 | 0.029 | 0.030 | 0.952 |

**表2** 当 $\theta = 0.8$ , 基于FGM copula函数的应力-强度模型可靠度的估计 $\hat{R}_1$ 的随机模拟结果

| $n$                             | $\lambda_1$ | $\lambda_2$ | $R$   | Est   | SD    | SE    | CP    |
|---------------------------------|-------------|-------------|-------|-------|-------|-------|-------|
| Censoring rate = (0.502, 0.500) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 1.000       | 0.500 | 0.498 | 0.076 | 0.076 | 0.946 |
| 100                             | 1.000       | 1.000       | 0.500 | 0.502 | 0.055 | 0.053 | 0.938 |
| 200                             | 1.000       | 1.000       | 0.500 | 0.500 | 0.039 | 0.038 | 0.934 |
| Censoring rate = (0.501, 0.332) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 2.000       | 0.307 | 0.313 | 0.065 | 0.062 | 0.922 |
| 100                             | 1.000       | 2.000       | 0.307 | 0.310 | 0.044 | 0.044 | 0.950 |
| 200                             | 1.000       | 2.000       | 0.307 | 0.308 | 0.031 | 0.031 | 0.944 |
| Censoring rate = (0.334, 0.504) |             |             |       |       |       |       |       |
| 50                              | 2.000       | 1.000       | 0.693 | 0.691 | 0.063 | 0.062 | 0.928 |
| 100                             | 2.000       | 1.000       | 0.693 | 0.692 | 0.043 | 0.044 | 0.958 |
| 200                             | 2.000       | 1.000       | 0.693 | 0.694 | 0.030 | 0.031 | 0.950 |

通过表1, 我们可以发现, 当样本量 $n = 50$ 较小时, 考虑到参数 $\lambda_1$ 和 $\lambda_2$ 取值不同以及变量删失率不同等情形, 估计量 $\hat{R}_1$ 各项统计指标均表现良好. 即估计值与其真值偏差较小, 标

准差的估计值接近其真值, 95%置信区间覆盖率在95%附近. 通过对比参数 $\lambda_1$ 和 $\lambda_2$ 取值的每一种情形下估计量的表现. 我们可以发现, 当样本量 $n$ 从50增加到200时, 估计值更接近其真值, 标准差的估计变小并表现出更接近其真值, 95%置信区间覆盖率一般会更接近95%.

此外, 表2随机模拟结果显示: 当变量之间的相关性在另一种情形下, 即 $\theta$ 的取值变化为0.8时. 估计量 $\hat{R}_1$ 依然表现良好.

#### 4.2 基于Clayton copula函数的应力-强度模型可靠度的估计量的随机模拟

对于估计量 $\hat{R}_2$ 的随机模拟研究, 随机模拟条件如: 变量边际分布的假定, 参数的选择, 样本量大小的选定, 删失变量的处理, 重复计算次数等等, 我们做上述相同的假定. 考虑到Clayton copula函数的参数 $\theta$ 取值大小体现了变量之间相关性的强弱. 我们分别考虑了当 $\theta = 0.8$ 时, 即变量之间相关性较弱时,  $\hat{R}_2$ 的随机模拟结果见表3. 以及当 $\theta = 2$ 时, 即变量之间相关性较强时,  $\hat{R}_2$ 的随机模拟结果见表4.

表3 当 $\theta = 0.8$ , 基于Clayton copula函数的应力-强度模型可靠度的估计 $\hat{R}_2$ 的随机模拟结果

| $n$                             | $\lambda_1$ | $\lambda_2$ | $R$   | Est   | SD    | SE    | CP    |
|---------------------------------|-------------|-------------|-------|-------|-------|-------|-------|
| Censoring rate = (0.508, 0.505) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 1.000       | 0.500 | 0.505 | 0.080 | 0.078 | 0.930 |
| 100                             | 1.000       | 1.000       | 0.500 | 0.510 | 0.055 | 0.055 | 0.932 |
| 200                             | 1.000       | 1.000       | 0.500 | 0.503 | 0.040 | 0.039 | 0.936 |
| Censoring rate = (0.496, 0.329) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 2.000       | 0.279 | 0.293 | 0.062 | 0.064 | 0.938 |
| 100                             | 1.000       | 2.000       | 0.279 | 0.294 | 0.044 | 0.045 | 0.944 |
| 200                             | 1.000       | 2.000       | 0.279 | 0.283 | 0.029 | 0.031 | 0.960 |
| Censoring rate = (0.337, 0.507) |             |             |       |       |       |       |       |
| 50                              | 2.000       | 1.000       | 0.721 | 0.722 | 0.062 | 0.062 | 0.934 |
| 100                             | 2.000       | 1.000       | 0.721 | 0.724 | 0.044 | 0.044 | 0.936 |
| 200                             | 2.000       | 1.000       | 0.721 | 0.721 | 0.030 | 0.031 | 0.952 |

通过表3, 我们发现采用Clayton copula函数来度量变量之间相关性时, 模型可靠度的估计 $\hat{R}_2$ 仍具有良好的性质. 即当样本量较小时, 估计值与其真值偏差较小, 标准差的估计值接近其真值, 95%置信区间覆盖率在95%附近. 对于参数 $\lambda_1$ 和 $\lambda_2$ 取值的每一种情形, 可以看出, 随着样本量 $n$ 的增大, 估计量的各项统计指标表现得更好.

通过表4, 我们可以得出, 即使改变变量之间的相关性, 让变量之间相关性变强时, 估计量 $\hat{R}_2$ 依然有良好的表现.

综上所述, 通过考察估计量 $\hat{R}_1$ 和 $\hat{R}_2$ 在各种不同情形下的随机模拟结果, 说明我们所提出的估计量具有良好的统计性质.



表4 当 $\theta = 2$ , 基于Clayton copula函数的应力-强度模型可靠度的估计 $\hat{R}_2$ 的随机模拟结果

| $n$                             | $\lambda_1$ | $\lambda_2$ | $R$   | Est   | SD    | SE    | CP    |
|---------------------------------|-------------|-------------|-------|-------|-------|-------|-------|
| Censoring rate = (0.500, 0.505) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 1.000       | 0.500 | 0.523 | 0.090 | 0.087 | 0.926 |
| 100                             | 1.000       | 1.000       | 0.500 | 0.515 | 0.061 | 0.061 | 0.932 |
| 200                             | 1.000       | 1.000       | 0.500 | 0.509 | 0.043 | 0.044 | 0.958 |
| Censoring rate = (0.492, 0.336) |             |             |       |       |       |       |       |
| 50                              | 1.000       | 2.000       | 0.198 | 0.227 | 0.063 | 0.065 | 0.930 |
| 100                             | 1.000       | 2.000       | 0.198 | 0.217 | 0.043 | 0.045 | 0.946 |
| 200                             | 1.000       | 2.000       | 0.198 | 0.210 | 0.031 | 0.031 | 0.934 |
| Censoring rate = (0.331, 0.499) |             |             |       |       |       |       |       |
| 50                              | 2.000       | 1.000       | 0.806 | 0.796 | 0.062 | 0.059 | 0.926 |
| 100                             | 2.000       | 1.000       | 0.806 | 0.800 | 0.040 | 0.042 | 0.958 |
| 200                             | 2.000       | 1.000       | 0.806 | 0.803 | 0.029 | 0.029 | 0.956 |

### 4.3 与已有方法的比较

Qi等<sup>[9]</sup>在变量 $X$ 和 $Y$ 相互独立的条件下提出了应力-强度模型可靠度 $R$ 的非参数估计, 然而该独立性的假定在实际应用中难以验证. 为此, 我们将本文提出的方法与之进行比较以阐明考虑变量 $X$ 和 $Y$ 相互关系的必要性. 在 $\lambda_1 = 2$ 和 $\lambda_2 = 1$ 下, 我们分别从FGM copula和Clayton copula中抽取样本量为100样本; 删失变量的产生方式如上, 并使得删失率约为33%和50%. 表5给出了在刻画相关性程度参数 $\theta$ 的不同取值下的两种方法的结果, 其中方法I为本文提出的方法, 方法II为Qi等<sup>[9]</sup>提出的方法. 从表中可以看出, 若忽略变量之间的相关性, 将得出有偏的估计, 从而进一步说明本文提出的将二者相关性考虑在内的方法的必要性与实用性.

表5 两种方法的比较结果

| Copula  | $\theta$ | Method | $R$   | Est   | SD    | SE    | CP    |
|---------|----------|--------|-------|-------|-------|-------|-------|
| FGM     | -0.8     | I      | 0.640 | 0.636 | 0.043 | 0.042 | 0.938 |
|         |          | II     | 0.640 | 0.666 | 0.049 | 0.048 | 0.886 |
|         | 0.8      | I      | 0.693 | 0.692 | 0.043 | 0.044 | 0.958 |
|         |          | II     | 0.693 | 0.663 | 0.037 | 0.039 | 0.902 |
| Clayton | 0.8      | I      | 0.721 | 0.724 | 0.044 | 0.044 | 0.936 |
|         |          | II     | 0.721 | 0.667 | 0.036 | 0.036 | 0.654 |
|         | 2        | I      | 0.806 | 0.800 | 0.040 | 0.042 | 0.958 |
|         |          | II     | 0.806 | 0.665 | 0.031 | 0.030 | 0.010 |

## §5. 结 论

综上所述, 本文研究了相关的应力变量和强度变量在右删失的情形下模型可靠度的非参数估计. 其中变量之间的相关关系采用常见的Farlie-Gumbel-Morgenstern copula函数和Clayton copula函数来度量. 采用经验过程的理论, 本文对所提出的估计量的渐近性质进行了证明. 通过考察估计量在不同情形下的随机模拟结果, 显示了估计量具有良好的统计性质. 此外, 上述方法也可以推广到当应力变量和强度变量的相关性采用其它的copula函数来度量时, 在变量右删失的情形下, 应力-强度模型可靠度的非参数估计问题.

## 参 考 文 献

- [1] Kotz S, Lumelskii Y, Pensky M. *The Stress-Strength Model and Its Generalizations: Theory and Applications* [M]. Singapore: World Scientific, 2003.
- [2] Ahmad K E, Fakhry M E, Jaheen Z F. Empirical Bayes estimation of  $P(Y < X)$  and characterizations of Burr-type  $X$  model [J]. *J. Statist. Plann. Inference*, 1997, **64**(2): 297–308.
- [3] Adimari G, Chiogna M. Partially parametric interval estimation of  $\Pr(Y > X)$  [J]. *Comput. Statist. Data Anal.*, 2006, **51**(3): 1875–1891.
- [4] Downton F. The estimation of  $\Pr(Y < X)$  in the normal case [J]. *Technometrics*, 1973, **15**(3): 551–558.
- [5] Tong H. A note on the estimation of  $\Pr\{Y < X\}$  in the exponential case [J]. *Technometrics*, 1974, **16**(4): 625–625.
- [6] Kundu D, Raqab M Z. Estimation of  $R = P(Y < X)$  for three-parameter Weibull distribution [J]. *Statist. Probab. Lett.*, 2009, **79**(17): 1839–1846.
- [7] Rezaei S, Tahmasbi R, Mahmoodi M. Estimation of  $P[Y < X]$  for generalized Pareto distribution [J]. *J. Statist. Plann. Inference*, 2010, **140**(2): 480–494.
- [8] Ghitany M E, Al-Mutairi D K, Aboukhamseen S M. Estimation of the reliability of a stress-strength system from power lindley distributions [J]. *Comm. Statist. Simulation Comput.*, 2015, **44**(1): 118–136.
- [9] Qi H, Qi F, Yu J C. Nonparametric inference for the stress-strength model under right censoring [J]. *Wuhan Univ. J. Nat. Sci.*, 2015, **20**(3): 202–206.
- [10] Domma F, Giordano S. A stress-strength model with dependent variables to measure household financial fragility [J]. *Stat. Methods Appl.*, 2012, **21**(3): 375–389.
- [11] Gupta R C, Subramanian S. Estimation of reliability in a bivariate normal distribution with equal coefficients of variation [J]. *Comm. Statist. Simulation Comput.*, 1998, **27**(3): 675–698.
- [12] Dieh W A, Dorvlo A S S, Al-Saidy O. Estimation of reliability in case of bivariate normal distribution using ranked set sampling with concomitant variable [J]. *IJCIM*, 2011, **19**(SP1): 16.1–16.4.
- [13] Hanagal D D. Note on estimation of reliability under bivariate Pareto stress-strength model [J]. *Statist. Papers*, 1997, **38**(4): 453–459.
- [14] Nadarajah S. Reliability for some bivariate beta distributions [J]. *Math. Probl. Eng.*, 2005, **2005**(1): 101–111.

- [15] Nadarajah S. Reliability for some bivariate gamma distributions [J]. *Math. Probl. Eng.*, 2005, **2005**(2): 151–163.
- [16] Nadarajah S, Kotz S. Reliability for some bivariate exponential distributions [J]. *Math. Probl. Eng.*, 2006, **2006**(Article ID 41652): 1–14.
- [17] Gupta R C, Ghitany M E, Al-Mutairi D K. Estimation of reliability from a bivariate log-normal data [J]. *J. Stat. Comput. Simul.*, 2013, **83**(6): 1068–1081.
- [18] Domma F, Giordano S. A copula-based approach to account for dependence in stress-strength models [J]. *Statist. Papers*, 2013, **54**(3): 807–826.
- [19] Nelsen R B. *An Introduction to Copulas* [M]. 2nd ed. New York: Springer, 2006.
- [20] Morgenstern D. Einfache beispiele zweidimensionaler verteilungen [J]. *Mitteilungsblatt für Mathematische Statistik*, 1956, **8**(1): 234–235.
- [21] Lo S H, Singh K. The product-limit estimator and the bootstrap: some asymptotic representations [J]. *Probab. Theory Relat. Fields*, 1986, **71**(3): 455–465.
- [22] Breslow N, Crowley J. A large sample study of the life table and product limit estimates under random censorship [J]. *Ann. Statist.*, 1974, **2**(3): 437–453.
- [23] Shorack G R, Wellner J A. *Empirical Processes with Applications to Statistics* [M]. New York: Wiley, 1986.
- [24] Lin D Y, Wei L J, Yang I, et al. Semiparametric regression for the mean and rate functions of recurrent events [J]. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 2000, **62**(4): 711–730.

## Nonparametric Estimation for Reliability of the Stress-Strength Model Based on Copula Function

QI Hui      GUAN Qiang

(Institute of Information Engineering, Sanming University, Sanming, 365004, China)

**Abstract:** In this paper, a nonparametric method for reliability of the stress-strength model is proposed when the dependent stress variable and strength variable are subject to right censoring. The dependence between variables is measured by the common Farlie-Gumbel-Morgenstern copula function and Clayton copula function. Using the empirical process theory, consistency and asymptotic normality of the proposed estimator is established in this paper. The results of numerical simulation show that the proposed method performs well in the case of finite sample. The method proposed in this paper has a wide application prospect in practice.

**Keywords:** stress-strength model; Kaplan-Meier estimator; right censoring; copula function

**2010 Mathematics Subject Classification:** 62G05; 62N86