第 35 卷 第 3 期 电 子 科 技 大 学 学 报 Vol.35 No.3
2006 年 6 月 Journal of University of Electronic Science and Technology of China Jun. 2006

删失数据下的Fisher信息量

黄继伟[1]，李云飞[2]，朱 宏[2]

(1. 江苏科技大学数理学院 江苏 镇江 212003; 2. 电子科技大学应用教学学院 成都 610054)

【摘要】根据现有的有关Fisher信息量的结论，通过逆危险率函数，推出了Fisher信息量的一种表达式，给出了一个被截尾的随机变量所保留的Fisher信息量和丢失的Fisher信息量的表达式，通过这种表达式可看到删失数据下Fisher信息量的组成成分。最后，得到了删失数据下Fisher信息量的一个性质。

关 键 词 逆危险率函数; Fisher信息量; 删失数据; 示性函数

中图分类号 O212.1 文献标识码 A

Fisher Information in Randomly Censored Data

HUANG Ji-wei[1]，LI Yun-fei[2]，ZHU Hong[2]

(1. School of Mathematic and Physics, Jianysu University of Science and Technology Zhenjiang Jangsu 212003;

2. School of Applied Mathematics, Univ. of Elec. Sci. & Tech. of China Chengdu 610054)

Abstract Through using the reversed hazard rate of random variables, this paper derives an expression of Fisher information based on the knowledge of Fisher information in hand. The expression of the retained and lost Fisher information of a random variable due to censorship is produced. So the component of Fisher information under censorship is detected. Then, this paper gives a property of Fisher information in censored data.

Key words reversed hazard rate; fisher information; censored data; indicator function

由于删失数据常用在可靠性分析，生存分析，医学研究中等，文献[1-3]为研究随机右删失模型的。而现在有关逆危险率函数的性质引起了广大研究者们的兴趣，本文则用逆危险率函数研究删失数据模型。

设随机变量 X 和 Y 的密度函数为 $f(x;\theta)$ 和 $g(x;\theta)$，分布函数为 $F(x;\theta)$ 和 $G(x;\theta)$，其中 $\theta\in\Theta$，Θ 是 R 中一开集， $x\in E=(a,b)$， E 是 R 中一个有限或无限的区间。在随机左删失模型中，不能完全观测到 X ，而仅能观测到 (Z,δ)，其中 $Z=\max(X,Y)$ ， $\delta=I(X>Y)$，这里 $I(\cdot)$ 表示某事件的示性函数。众所周知， Z 的密度函数为 $h(x;\theta)=fG+gF$ ， Z 中含有的关于 θ 的Fisher信息量表示为 $I^{Z}(\theta)$，则：

$$I^{Z}(\theta)=\int_E[\frac{\partial}{\partial\theta}\ln(fG+gF)]^2(fG+gF)\,\mathrm{d}x \tag{1}$$

式中 (Z,δ) 的似然函数为 $L(z,\delta)=[f(z;\theta)G(z;\theta)]^{\delta}[g(z;\theta)F(z;\theta)]^{1-\delta}$ ；(Z,δ) 中含有的Fisher信息量为 $I^{Z,\delta}(\theta)$ ；文献[1]给出了：

$$I^{Z,\delta}(\theta)=\int_E[\frac{\partial}{\partial\theta}\ln(fG)]^2 fG\,\mathrm{d}x+\int_E[\frac{\partial}{\partial\theta}\ln(gF)]^2 gF\,\mathrm{d}x \tag{2}$$

式中 X的逆危险率函数为 $r_X(x;\theta)$，定义为 $r_X(x;\theta)=f(x;\theta)/F(x;\theta)$。文献[2]给出了$X$的Fisher信息量的一种表示：

$$I^{X}(\theta)=\int_E(\frac{\partial}{\partial\theta}\ln f)^2 f\,\mathrm{d}x=\int_E(\frac{\partial}{\partial\theta}\ln r_X)^2 f\,\mathrm{d}x \tag{3}$$

文献[3]通过危险率函数得到了在随机右删失下Fisher信息量的一种表示，本文则通过逆危险率函数，得到了随机左删失下Fisher信息量的表示及相关性质。

收稿日期：2004 – 03 – 29

作者简介：黄继伟(1976 –)，男，硕士，主要从事数理统计等方面的研究.

1 主要结果

设集合$\{x; f(x;\theta)>0\}$和$\{x; g(x;\theta)>0\}$与参数θ无关，对任意$\theta\in\Theta$，$\partial/\partial\theta\, f(x;\theta)$和$\partial/\partial\theta\, g(x;\theta)$在$E$上可积，对任意$x\in E$和$\theta\in\Theta$，$\dfrac{\partial}{\partial\theta}f(x;\theta)=\dfrac{\partial^2}{\partial\theta\partial x}F(x;\theta)=\dfrac{\partial^2}{\partial x\partial\theta}F(x;\theta)$，$\dfrac{\partial}{\partial\theta}g(x;\theta)=\dfrac{\partial^2}{\partial\theta\partial x}G(x;\theta)=\dfrac{\partial^2}{\partial x\partial\theta}G(x;\theta)$，上述的导数都存在，所涉及到的极限，导数，积分运算可以交换次序，且当$x\to b$或$x\to a$时都有$(\partial/\partial\theta\ln F)^2F\to 0$和$(\partial/\partial\theta\ln G)^2G\to 0$。

定理 1

$$I^{Z,\delta}(\theta)=I^X(\theta)-\int_E(\frac{\partial}{\partial\theta}\ln r_X)^2 f\overline{G}\,\mathrm{d}x+I^Y(\theta)-\int_E(\frac{\partial}{\partial\theta}\ln r_Y)^2 g\overline{F}\,\mathrm{d}x \tag{4}$$

证明 用f'_θ表示$\partial f/\partial\theta$，$F'_\theta$表示$\partial F/\partial\theta$，$g'_\theta$表示$\partial g/\partial\theta$，$G'_\theta$表示$\partial G/\partial\theta$，首先有：

$$\int_E[\frac{\partial}{\partial\theta}\ln(fG)]^2 fG\,\mathrm{d}x=I^X(\theta)-\int_E(\frac{\partial}{\partial\theta}\ln f)^2 f\overline{G}\,\mathrm{d}x+2\int_E f'_\theta G'_\theta\,\mathrm{d}x+\int_E\frac{(G'_\theta)^2}{G}f\,\mathrm{d}x \tag{5}$$

由分部积分可得：

$$\int_E[\frac{\partial}{\partial\theta}(fG)]^2 fG\,\mathrm{d}x=I^X(\theta)-\int_E(\frac{\partial}{\partial\theta}\ln f)^2 f\overline{G}\,\mathrm{d}x+2\int_E f'_\theta G'_\theta\,\mathrm{d}x+\int_E[\frac{2g'_\theta G'_\theta}{gG}-\frac{(G'_\theta)^2}{G^2}]g\overline{F}\,\mathrm{d}x \tag{6}$$

类似的可得到：

$$\int_E[\frac{\partial}{\partial\theta}(gF)]^2 gF\,\mathrm{d}x=I^Y(\theta)-\int_E(\frac{\partial}{\partial\theta}\ln f)^2 g\overline{F}\,\mathrm{d}x+2\int_E g'_\theta F'_\theta\,\mathrm{d}x+\int_E[\frac{2f'_\theta F'_\theta}{fF}-\frac{(F'_\theta)^2}{F^2}]f\overline{G}\,\mathrm{d}x \tag{7}$$

把式(6)和式(7)代入式(2)得：

$$\begin{aligned}I^{Z,\delta}(\theta)=I^X(\theta)+\int_E[\frac{2f'_\theta F'_\theta}{fF}-\frac{(F'_\theta)^2}{F^2}-(\frac{\partial}{\partial\theta}\ln f)^2]f\overline{G}\,\mathrm{d}x+I^Y(\theta)+2\int_E f'_\theta G'_\theta\,\mathrm{d}x+\\2\int_E g'_\theta F'_\theta\,\mathrm{d}x+\int_E[\frac{2g'_\theta G'_\theta}{gG}-\frac{(G'_\theta)^2}{G^2}-(\frac{\partial}{\partial\theta}\ln g)^2]g\overline{F}\,\mathrm{d}x\end{aligned} \tag{8}$$

由分部积分有：

$$2\int_E f'_\theta G'_\theta\,\mathrm{d}x=-2\int_E g'_\theta F'_\theta\,\mathrm{d}x \tag{9}$$

而

$$\begin{aligned}\int_E[\frac{2f'_\theta F'_\theta}{fF}-\frac{(F'_\theta)^2}{F^2}-(\frac{\partial}{\partial\theta}\ln f)]f\overline{G}\,\mathrm{d}x=\int_E(\frac{\partial}{\partial\theta}\ln F)(\frac{\partial}{\partial\theta}\ln\frac{f}{F})f\overline{G}\,\mathrm{d}x-\int_E(\frac{\partial}{\partial\theta}\ln f)(\frac{\partial}{\partial\theta}\ln\frac{f}{F})f\overline{G}\quad\mathrm{d}x=\\-\int_E(\frac{\partial}{\partial\theta}\ln r_X)^2 f\overline{G}\,\mathrm{d}x\end{aligned} \tag{10}$$

同理可得：

$$\int_E[\frac{2g'_\theta G'_\theta}{gG}-\frac{(G'_\theta)^2}{G^2}-(\frac{\partial}{\partial\theta}\ln g)]g\overline{F}\,\mathrm{d}x=-\int_E(\frac{\partial}{\partial\theta}\ln r_Y)^2 g\overline{F}\,\mathrm{d}x \tag{11}$$

把式(9)～(11)代入式(8)可得式(4)。 证毕

在随机左删失下，X由于被Y截尾所保留的和丢失的Fisher信息量可定义为：

$$I_L^X(\theta)=\int_E(\frac{\partial}{\partial\theta}\ln r_X)^2 fG\,\mathrm{d}x,\quad I_R^X(\theta)=\int_E(\frac{\partial}{\partial\theta}\ln r_Y)^2 f\overline{G}\,\mathrm{d}x$$

由式(3)和(4)得：

$$\begin{cases}I^X(\theta)=\int_E(\frac{\partial}{\partial\theta}\ln r_X)^2 fG\mathrm{d}x+\int_E(\frac{\partial}{\partial\theta}\ln r_X)^2 f\overline{G}\mathrm{d}x=I_L^X(\theta)+I_R^X(\theta)\\ I^{Z,\delta}(\theta)=\int_E[\frac{\partial}{\partial\theta}\ln(fG)]^2 fG\mathrm{d}x+\int_E[\frac{\partial}{\partial\theta}\ln(gF)]^2 gF\mathrm{d}x=I_L^X(\theta)+I_L^Y(\theta)\end{cases} \tag{12}$$

式(12)则定量地给出，在随机左删失下，(Z,δ)中含有的Fisher信息量与X和Y所保留的Fisher信息量的关系。(Z,δ)中含有的Fisher信息量与Z中含有的Fisher信息量的关系由下面的定理2给出。

定理 2 $I^{Z,\delta}(\theta)\geq I^Z(\theta)$，等号成立的充要条件是$r_X/r_Y$与$\theta$无关。

证明　由式(1)和(2)有：

$$I^{Z,\delta}(\theta)-I^{Z}(\theta)=\int_E[\frac{\partial}{\partial\theta}\ln(fG)]^2 fG\,\mathrm{d}x-\int_E\frac{fG}{fG+gF}[\frac{\partial}{\partial\theta}\ln(fG)]^2 fG\,\mathrm{d}x+$$
$$\int_E[\frac{\partial}{\partial\theta}\ln(gF)]^2 gF\,\mathrm{d}x-\int_E\frac{gF}{fG+gF}[\frac{\partial}{\partial\theta}\ln(gF)]^2 gF\,\mathrm{d}x-$$
$$2\int_E\frac{fGgF[\ln(fG)][\ln(gF)]}{fG+gF}\mathrm{d}x=$$
$$\int_E[\frac{\partial}{\partial\theta}\ln(fG)]^2\frac{fGgF}{fG+gF}\mathrm{d}x+\int_E[\frac{\partial}{\partial\theta}\ln(gF)]^2\frac{fGgF}{fG+gF}\mathrm{d}x-$$
$$\int_E\frac{fGgF[\ln(fG)][\ln(gF)]}{fG+gF}\mathrm{d}x=\int_E\frac{fGgF}{fG+gF}(\frac{\partial}{\partial\theta}\frac{r_X}{r_Y})^2\mathrm{d}x\geq 0$$

等号成立的充要条件是 r_X/r_Y 与 θ 无关。　证毕

2　结 束 语

本文得到的仅仅是一维数据左删失下Fisher信息量的一种表示及相关性质，那么在多维数据左删失下，Fisher信息量的表示及相应的性质是否成立有待进一步研究。

参　考　文　献

[1] Prakasa R. Remarks on cramér-rao type integral inequalities for randomly censored data[C]// In: Koul H L Analysis of Censored Data, Hayward, CA, 1995: 68-79.

[2] Efron B, Johnstone L. Fisher information in terms of the hazard rate[J]. Ann. Statist, 1990, 18: 38-62.

[3] Zheng G, Joseph L G. On the fisher information in randomly censored data[J].Statist. Probab. Lett, 2001, 52: 421-426.

编　辑　刘文珍