

# 线性模型中 M 估计分布的 随机加权方法逼近\*

吴小燕 赵林城 杨亚宁

(中国科学技术大学统计与金融系, 合肥 230026)

**摘要** 在线性模型中, M 估计的渐近分布通常都涉及到不易估计的未知误差分布的某些量, 如果要估计渐近方差, 就需对这些冗余参数进行估计. 利用随机加权方法可以避免先对误差分布中的冗余参数进行估计. 给出了当自变量是随机变量时, M 估计分布的随机加权逼近, 证明了 M 估计分布的随机加权逼近是一致相合的. 当取不同的凸函数, 样本大小和随机权时, 进一步利用蒙特卡洛方法研究估计分布. 研究表明随机权取泊松权时, 不仅达到同样的效果而且可以减小计算量.

**关键词** 线性模型, M 估计, 随机加权.

MR(2000) 主题分类号 62J05

## 1 引言

在线性回归分析中, M 方法可说是最受重视、研究成果最多的一种, 但对 M 估计的方差估计时, 要涉及到冗余参数的估计, 而这些冗余参数往往不能够被精确估计. 这是因为这些冗余参数通常都与误差分布的密度函数有关, 而未知密度函数或与其有关的量的估计一般要涉及到非参数函数估计方法, 通常较困难而且不是十分有效的. 随机加权是一种有效解决此问题的方法之一 (参见文献 [1]). 随机加权方法与传统的 Bootstrap 方法有些类似, 但它在应用时则更加灵活, 因为可以较自由地选取随机权 (参见文献 [2–4]).

文献 [1] 证明了线性模型中当自变量非随机时, M 估计分布的随机加权逼近是一致相合的. 自变量随机的情况大多出现于观察研究中, 特别是横断抽样, 这种研究方式多出现于社会调查, 心理学研究, 经济学等领域. 在本文中我们将讨论当自变量随机时, M 估计分布的随机加权逼近的一致相合性问题. 当取不同的凸函数, 样本大小和随机权时, 我们利用蒙特卡洛方法研究了 M 估计分布的随机加权逼近. 同时将随机加权方法与传统的 Bootstrap 方法作了比较.

考虑线性模型

$$y_i = x_i' \beta + e_i, \quad i = 1, 2, \dots, n, \quad (1.1)$$

\* 国家自然科学基金 (10471136) 资助课题.

收稿日期: 2006-06-12.

其中  $\beta$  是未知的  $p$ -维回归参数向量,  $y_i, x_i$  和  $\{e_i\}$  分别是第  $i$  次观察或试验时的目标变量(因变量)、解释变量(自变量)和随机误差.  $y_i$  和  $x_i$  分别为 1 维和  $p$  维向量. 在本文中我们假定自变量是随机的, 并且  $(x, e), (x_i, e_i), i = 1, 2, \dots, n$  独立同分布. 假定  $\rho(\cdot)$  是定义  $R^1$  上的非单调凸函数, 使得  $\inf_u E[\rho(e_1 + u) - \rho(e_1)] = 0$ . 在模型 (1.1) 中参数  $\beta$  的 M 估计记为  $\hat{\beta}_n$ , 即满足

$$\sum_{i=1}^n \rho(y_i - x_i' \hat{\beta}_n) = \inf_{\beta \in R^p} \sum_{i=1}^n \rho(y_i - x_i' \beta). \quad (1.2)$$

随机权满足的条件如同文献 [1], 假定随机权  $w_i, i = 1, 2, \dots, n$ , 满足

(A<sub>0</sub>)  $w_1, w_2, \dots$  独立同分布.  $P(w_1 \geq 0) = 1, Ew_1 = 1, Ew_1^2 = \tau \geq 1$ , 并且  $\{w_i\}$  和  $\{(x_i, e_i)\}$  是相互独立的.

满足 (A<sub>0</sub>) 的一些简单的随机权如

i) 泊松分布:  $w \sim \text{Poisson}(1)$ ;

ii) 指数分布:  $w \sim \exp(1)$ ;

iii) 对数正态分布:  $w \sim \log N(-\log(\sqrt{2}), \log(2))$  等等.

线性模型 (1.1) 中回归参数  $\beta$  的随机加权 M 估计  $\beta_n^*$  满足

$$\sum_{i=1}^n w_i \rho(y_i - x_i' \beta_n^*) = \inf_{\beta \in R^p} \sum_{i=1}^n w_i \rho(y_i - x_i' \beta). \quad (1.3)$$

特别地我们取  $\tau = 2$ , 在给定的  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  时, 我们将用  $\sqrt{n}(\beta_n^* - \hat{\beta}_n)$  的条件分布去逼近  $\sqrt{n}(\hat{\beta}_n - \beta)$  的分布.

本文中  $\mathcal{L}^*, P^*, E^*$  和  $\text{Var}^*$  指的是给定的  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  时的条件分布, 条件概率, 条件期望和条件方差.  $\|b\|$  为  $b$  的欧氏范数.  $c$  是正常数, 在不同的式子中是不同的,  $I(B)$  是  $B$  的示性函数. 真实参数记  $\beta_0$ .  $\lim_{n \rightarrow \infty} X_n = X$  in pr. 表示  $X_n$  依概率收敛到  $X$ .

## 2 主要结果

假定  $\rho(\cdot)$  是定义  $R^1$  上的非单调凸函数,  $\psi_-$  和  $\psi_+$  是  $\rho$  的左右导数,  $\psi$  是介于  $\psi_-$  和  $\psi_+$  之间的一个函数, 即

$$\psi_-(u) \leq \psi(u) \leq \psi_+(u) \quad (\text{对任意的 } u \in R).$$

我们进一步引入以下假定

(A<sub>1</sub>)  $G(u|x) \equiv E[\psi(e+u)|x], g(u|x) \equiv \frac{\partial G(u|x)}{\partial u}$ . 给定  $x, G(0|x) = 0, g(0|x) > 0, G(u|x)$  在  $u = 0$  附近存在且是有限的,  $g(\cdot|x)$  在 0 点连续, 且  $|g(u|x)| \leq A(x)$ , 其中  $A(x)$  满足  $E A(x) \|x\|^2 < \infty$ ;

(A<sub>2</sub>)  $0 < \sigma^2(x) \equiv E[\psi^2(e)|x] < \infty$ , 对于给定  $x$ , 当  $u \rightarrow 0$  时,  $E[(\psi(e+u) - \psi(e))^2|x] \rightarrow 0$ , a.s. 且  $E[(\psi(e+u) - \psi(e))^2|x] \leq A(x)$ ;

(A<sub>3</sub>)  $J_0 \equiv E[g(0|x)xx']$  非奇异,  $V \equiv E[\sigma^2(x)xx'] < \infty$ , 且  $E\|x\|^2 < \infty$ .

以上条件在实际中容易满足, 比方说, 如果  $x$  与  $e$  相互独立, 对于 LAD 估计来说,  $\rho(u) = |u|, \psi(u) = 2I(u \geq 0) - 1, G(0|x) = 0$  等价于  $\text{med}(e) = 0$ , 而  $g(0|x) > 0$  意味着  $e$  在 0 点的概率密度函数大于 0.

**定理 2.1** 在模型 (1.1) 之下, 假定 (A<sub>0</sub>)–(A<sub>3</sub>) 成立,  $\beta_n^*$  见 (1.3) 中的定义, 则当  $n \rightarrow \infty$  时,

$$\sqrt{n}(\beta_n^* - \beta_0) = \frac{1}{\sqrt{n}} J_0^{-1} \sum_{i=1}^n w_i \psi(e_i) x_i + o_p(1). \quad (2.1)$$

特别地, 如果取  $w_i \equiv 1$ , 则得到当  $n \rightarrow \infty$  时,

$$\sqrt{n}(\hat{\beta}_n - \beta_0) = \frac{1}{\sqrt{n}} J_0^{-1} \sum_{i=1}^n \psi(e_i) x_i + o_p(1) \xrightarrow{L} N(0, J_0^{-1} V J_0^{-1}), \quad (2.2)$$

其中  $N$  为  $p \times 1$  维随机向量, 服从  $N(0, J_0^{-1} V J_0^{-1})$  分布.

**定理 2.2** 在模型 (1.1) 之下, 假定 (A<sub>0</sub>)–(A<sub>3</sub>) 成立, 在 (A<sub>0</sub>) 中  $\tau = 2$ .  $\hat{\beta}_n$  和  $\beta_n^*$  如 (1.2) 和 (1.3) 中的定义. 则当  $n \rightarrow \infty$  时,

$$\mathcal{L}^*(\sqrt{n}(\beta_n^* - \hat{\beta}_n) \rightarrow N(0, J_0^{-1} V J_0^{-1}) \quad \text{in pr.}, \quad (2.3)$$

进一步,

$$\sup_{u \in R^p} |P^*(\sqrt{n}(\beta_n^* - \hat{\beta}_n) \leq u) - P(N \leq u)| \rightarrow 0 \quad \text{in pr.}. \quad (2.4)$$

对比 (2.2) 和 (2.4), 当  $n \rightarrow \infty$  时,  $\sqrt{n}(\beta_n^* - \hat{\beta}_n)$  和  $\sqrt{n}(\hat{\beta}_n - \beta_0)$  之间的多维 Kolmogorow-Smirnov 距离为

$$\sup_u |P^*(\sqrt{n}(\beta_n^* - \hat{\beta}_n) \leq u) - P(\sqrt{n}(\hat{\beta}_n - \beta_0) \leq u)| \rightarrow 0, \quad \text{in pr.}, \quad (2.5)$$

那么在这种意义下  $\sqrt{n}(\hat{\beta}_n - \beta_0)$  的分布可以用被给  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  的条件下  $\sqrt{n}(\beta_n^* - \hat{\beta}_n)$  的条件分布来逼近.

### 3 主要结果的证明

在证明定理之前, 我们需要以下引理.

**引理 3.1**  $D$  是  $R^p$  上的开凸集,  $f_1, f_2, \dots$  是  $D$  上的一系列随机凸函数,  $f$  是  $D$  上的某一实函数, 对任意的  $x \in D$ , 当  $n \rightarrow \infty$  时,  $f_n(x) \rightarrow f(x)$  a.s. (或 in pr.). 则  $f$  是凸函数, 并且对  $D$  上任一紧集  $H$ , 当  $n \rightarrow \infty$  时,

$$\sup_{x \in H} |f_n(x) - f(x)| \rightarrow 0, \quad \text{a.s. (或 in pr.)}$$

参见文献 [5] 中的定理 II.1.

**引理 3.2** 在定理 2.1 的条件下, 当  $n \rightarrow \infty$  时,

$$\sup_{\|\gamma\| \leq c} \left| \sum_{i=1}^n w_i \left\{ \rho\left(e_i - \frac{1}{\sqrt{n}} x_i' \gamma\right) - \rho(e_i) + \frac{1}{\sqrt{n}} \psi(e_i) x_i' \gamma \right\} - \frac{1}{2} \gamma' J_0 \gamma \right| \rightarrow 0, \quad \text{in pr.} \quad (3.1)$$

证 令

$$\begin{aligned} f_n(\gamma) &= \sum_{i=1}^n w_i \left\{ \rho\left(e_i - \frac{1}{\sqrt{n}} x_i' \gamma\right) - \rho(e_i) + \frac{1}{\sqrt{n}} \psi(e_i) x_i' \gamma \right\} \\ &= \sum_{i=1}^n w_i \int_0^{-\frac{1}{\sqrt{n}} x_i' \gamma} \{\psi(e_i + u) - \psi(e_i)\} du. \end{aligned}$$

由 (A<sub>1</sub>) 及控制收敛定理, 得到

$$\begin{aligned} Ef_n(\gamma) &= nE\left\{\int_0^{-\frac{1}{\sqrt{n}}x'\gamma} E[\psi(e+u) - \psi(e)|x]du\right\} = nE\left\{\int_0^{-\frac{1}{\sqrt{n}}x'\gamma} G(u|x)du\right\} \\ &= nE\left\{\int_0^{-\frac{1}{\sqrt{n}}x'\gamma} g(0|x)u(1+o(1))du\right\} = \frac{1}{2}\gamma'E[g(0|x)xx']\gamma(1+o(1)). \end{aligned}$$

因此当  $n \rightarrow \infty$  时,

$$Ef_n(\gamma) \rightarrow \frac{1}{2}\gamma'J_0\gamma. \quad (3.2)$$

由 Schwarz 不等式、(A<sub>2</sub>) 及控制收敛定理, 可得

$$\begin{aligned} \text{Var}[f_n(\gamma)] &= \sum_{i=1}^n \text{Var}\left\{w_i \int_0^{-\frac{1}{\sqrt{n}}x'_i\gamma} [\psi(e_i+u) - \psi(e_i)]du\right\} \\ &\leq nE\left\{w_i \int_0^{-\frac{1}{\sqrt{n}}x'\gamma} [\psi(e+u) - \psi(e)]du\right\}^2 \\ &\leq \tau nE\left|\frac{1}{\sqrt{n}}x'\gamma \cdot \int_0^{-\frac{1}{\sqrt{n}}x'\gamma} E[(\psi(e+u) - \psi(e))^2|x]du\right| \\ &= \tau E(x'\gamma)^2 o(1). \end{aligned}$$

因此当  $n \rightarrow \infty$  时,

$$\text{Var}[f_n(\gamma)] \rightarrow 0. \quad (3.3)$$

从 (3.2) 式和 (3.3) 式, 可以得到当  $n \rightarrow \infty$  时,  $f_n(\gamma) \rightarrow \frac{1}{2}\gamma'J_0\gamma$ , in pr. 因为  $f_n(\gamma)$  是凸函数, 由引理 3.1 可知 (3.1) 式成立.

**定理 2.1 的证明** 不丢失一般性, 我们假定真实参数  $\beta_0 = 0$ , 记

$$\bar{\beta}_n = J_0^{-1} \sum_{i=1}^n \frac{1}{\sqrt{n}} w_i \psi(e_i) x_i. \quad (3.4)$$

容易看出在定理 2.1 的条件下, Lindeberg 条件成立, 则当  $n \rightarrow \infty$  时,

$$\bar{\beta}_n \xrightarrow{\mathcal{L}} N(0, \tau J_0^{-1} V J_0^{-1}), \quad (3.5)$$

其中  $J_0 = E[g(0|x)xx']$ ,  $V = E[\sigma^2(x)xx']$ . 这就包含了对任意给定  $\varepsilon > 0$ , 总存在一个常数  $c > 0$  使得当  $n \geq n_0$  时,

$$P(\|\bar{\beta}_n\| > c) < \frac{\varepsilon}{2}. \quad (3.6)$$

由引理 3.2 知, 对任意  $\delta > 0$ , 当  $n \rightarrow \infty$  时,

$$I(\|\bar{\beta}_n\| \leq c) \sup_{\|\gamma - \bar{\beta}_n\| = \delta} \left| \sum_{i=1}^n w_i \left\{ \rho\left(y_i - \frac{x'_i}{\sqrt{n}}\gamma\right) - \rho\left(y_i - \frac{x'_i}{\sqrt{n}}\bar{\beta}_n\right) \right\} - \frac{1}{2}(\gamma - \bar{\beta}_n)' J_0 (\gamma - \bar{\beta}_n) \right| \rightarrow 0.$$

记  $\lambda$  为  $J_0$  的最小特征值, 通过以上两式, 对足够大的  $n$  有

$$P\left\{ \inf_{\|\gamma - \bar{\beta}_n\| = \delta} \sum_{i=1}^n w_i \rho\left(y_i - \frac{x'_i}{\sqrt{n}}\gamma\right) \geq \sum_{i=1}^n w_i \rho\left(y_i - \frac{x'_i}{\sqrt{n}}\bar{\beta}_n\right) + \frac{1}{4}\lambda\delta^2 \right\} \geq 1 - \varepsilon. \quad (3.7)$$

由 (3.7) 式和  $\rho(\cdot)$  的凸性, 得到对足够大的  $n$

$$P\{\|\sqrt{n}\beta_n^* - \bar{\beta}_n\| < \delta\} \geq 1 - \varepsilon.$$

因为  $\varepsilon$  和  $\delta$  是任意的, 故由 (3.7) 式可知定理 2.1 的证明完成.

### 定理 2.2 的证明 记

$$Z_n = J_0^{-1} \sum_{i=1}^n \frac{1}{\sqrt{n}} (w_i - 1) \psi(e_i) x_i, \quad \sqrt{n}(\beta_n^* - \hat{\beta}_n) = Z_n + \zeta_n.$$

由定理 2.1, 当  $n \rightarrow \infty$  时,

$$E^*(\|\zeta_n\| \wedge 1) \rightarrow 0, \quad \text{in pr.}$$

其中  $a \wedge b = \min\{a, b\}$ .

令  $U_0$  为  $R^p$  中单位球面  $U = \{\theta : \|\theta\| = 1\}$  上的可数稠子集. 首先我们证对任意  $\theta_0 \in U_0$  和任意  $-\infty < v < \infty$ , 当  $n \rightarrow \infty$  时,

$$P^*(\theta_0' Z_n \leq v) \rightarrow P(\theta_0' N \leq v), \quad \text{in pr.}$$

由 Lindeberg 中心极限定理, 只需证明当  $n \rightarrow \infty$  时,

$$\text{Var}^*(\theta_0' Z_n) = \frac{1}{n} \theta_0' J_0^{-1} \sum_{i=1}^n \psi^2(e_i) x_i x_i' J_0^{-1} \theta_0 \rightarrow \theta_0' J_0^{-1} V J_0^{-1} \theta_0, \quad \text{in pr.} \quad (3.8)$$

和  $\forall \varepsilon > 0$ , 当  $n \rightarrow \infty$  时,

$$[\text{Var}^*(\theta_0' Z_n)]^{-1} \sum_{i=1}^n E^* \eta_{in}^2 I(\|\eta_{in}\| \geq \varepsilon [\text{Var}^*(\theta_0' Z_n)]^{\frac{1}{2}}) \rightarrow 0, \quad \text{in pr.} \quad (3.9)$$

其中  $\eta_{in} = \frac{1}{\sqrt{n}} \theta_0' J_0^{-1} (w_i - 1) \psi(e_i) x_i$ .

(3.8) 式成立是显然的. 下面证明 (3.9) 式成立.

令

$$T_n = \max_{1 \leq i \leq n} \psi^2(e_i) \theta_0' J_0^{-1} x_i x_i' J_0^{-1} \theta_0.$$

因为

$$E \psi^2(e_i) \theta_0' J_0^{-1} x_i x_i' J_0^{-1} \theta_0 = \theta_0' J_0^{-1} V J_0^{-1} \theta_0 < \infty,$$

因此当  $n \rightarrow \infty$  时,

$$\frac{T_n}{n} \rightarrow 0 \quad \text{a.s.}$$

所以有

$$\begin{aligned} & \sum_{i=1}^n E^* \eta_{in}^2 I(\|\eta_{in}\|^2 \geq \varepsilon) \\ &= \frac{1}{n} \sum_{i=1}^n \psi^2(e_i) \theta_0' J_0^{-1} x_i x_i' J_0^{-1} \theta_0 E^*(w_i - 1)^2 I\left(\frac{1}{n} (w_i - 1)^2 \psi^2(e_i) \theta_0' J_0^{-1} x_i x_i' J_0^{-1} \theta_0 \geq \varepsilon\right) \\ &\leq \frac{1}{n} \sum_{i=1}^n \psi^2(e_i) \theta_0' J_0^{-1} x_i x_i' J_0^{-1} \theta_0 E^*(w_i - 1)^2 I\left(\frac{T_n}{n} (w_i - 1)^2 \geq \varepsilon\right) \rightarrow 0, \quad \text{in pr.} \end{aligned}$$

因此 (3.9) 式得证.

$\{n_1\}$  是  $\{n\}$  的任一子列, 利用对角线方法可以取到一子列  $\{n_2\} \subset \{n_1\}$ , 对任意  $\theta_0 \in U_0$  以概率 1 成立

$$\sup_{v \in R^1} |P^*(\theta'_0 Z_{n_2} \leq v) - P(\theta'_0 N \leq v)| \rightarrow 0.$$

那么对任意  $\theta \in U$

$$\sup_{v \in R^1} |P^*(\theta'(Z_{n_2} + \zeta_{n_2}) \leq v) - P(\theta'N \leq v)| \rightarrow 0$$

也成立, 因此当  $n \rightarrow \infty$  时,

$$\sup_{u \in R^p} |P^*(\sqrt{n}(\beta_{n_2}^* - \hat{\beta}_{n_2}) \leq u) - P(N \leq u)| \rightarrow 0, \quad \text{in pr.} \quad (3.10)$$

对  $\{n\}$  的任一子列  $\{n_1\}$ , 总存在  $\{n_2\} \subset \{n_1\}$  使得 (3.10) 式成立, 这就包含了当  $n \rightarrow \infty$  时,

$$\sup_{u \in R^p} |P^*(\sqrt{n}(\beta_n^* - \hat{\beta}_n) \leq u) - P(N \leq u)| \rightarrow 0, \quad \text{in pr.}$$

定理 2.2 的证明完成.

## 4 模拟结果

这节我们利用蒙特卡洛方法来研究随机加权方法逼近 M 估计分布的好坏, 我们取了不同大小的样本, 不同的随机权, 不同的误差, 凸函数分别取绝对值函数 (最小一乘估计) 和 Huber 函数, 当取 Huber 函数时

$$\psi(x) = \begin{cases} x, & \text{如果 } |x| \leq 1.345, \\ 1.345, & \text{如果 } x > 1.345, \\ -1.345, & \text{如果 } x < -1.345. \end{cases}$$

我们假定模型为

$$y_i = \alpha + x_i \beta + e_i, \quad i = 1, 2, \dots, n.$$

随机误差为标准柯西分布 (即  $t_1$  分布) 和  $t_2$  分布. 假定  $x_i$  和  $e_i$  相互独立. 真实参数  $\alpha = 0$ ,  $\beta = 1$ . 我们的重点是对参数  $\beta$  的估计.

通过计算, 渐近方差如下

$$\begin{aligned} \sigma_{l_1 u t_1}^2 &= 7.402, & \sigma_{l_1 u t_2}^2 &= 6, & \sigma_{l_1 n t_1}^2 &= 2.46, & \sigma_{l_1 n t_2}^2 &= 2, \\ \sigma_{h u t_1}^2 &= 8.528, & \sigma_{h u t_2}^2 &= 5.538, & \sigma_{h n t_1}^2 &= 2.843, & \sigma_{h n t_2}^2 &= 1.846, \end{aligned}$$

其中下标  $l_1$  指的是选取凸函数为绝对值函数, 下标  $h$  指的是选取凸函数为 Huber 函数, 下标  $u$  指的是  $x \sim U(-1, 1)$ , 下标  $n$  指的是  $x \sim N(0, 1)$ , 下标  $t_1$  指的是  $e \sim t_1$ , 下标  $t_2$  指的是  $e \sim t_2$ .

随机权选为

指数分布:  $w \sim \exp(1)$ ,

泊松分布:  $w \sim \text{Poisson}(1)$ .

首先模拟 M-估计和随机加权 M-估计分布情况,并且将随机加权方法与一般的 Bootstrap 方法进行了比较.随机产生 10000 组  $(y_i, x_i), 1 \leq i \leq n$  的样本,同时产生 10000 组  $w_i, 1 \leq i \leq n$ , 从而计算相应回归参数的随机加权 M 估计. 因为回归变量是独立同分布的, 在计算参数的 bootstrap 估计时, 从  $\{(y_i, x_i), 1 \leq i \leq n\}$  中再随机抽样  $n$  次, 同样也产生 10000 组. 图 1 是凸函数取为 Huber 函数及自变量服从  $U(-1, 1)$  分布时,  $M$  估计, Bootstrap 估计和随机加权 M 估计分布的情况. 其中第 1、2 列图的误差服从柯西分布, 第 3、4 两列图的误差服从  $t_2$  分布, 样本大小  $n = 500$ . 不出所料, 当样本达到 500 时用各种估计逼近效果都非常好. 另外我们也看到随机加权方法与一般的 Bootstrap 的方法没有明显的不同. 当凸函数取为绝对值函数及自变量服从  $N(0, 1)$  时的情况类似.

接下来我们研究各种估计的方差估计.在模拟中我们选取不同的样本大小 200, 400, 1000. 随机产生 1,000 组  $(y_i, x_i), 1 \leq i \leq n$  和  $w_i, 1 \leq i \leq n$  分别计算出 M 估计方差、随机加权 M 估计的方差和 Bootstrap 估计的方差. 将上述过程再重复 1000 次, 进一步分别计算出它们的平均方差. 我们把所得平均方差结果写在表 1 和表 2 中, 其中表 1 为当  $x \sim U(-1, 1), e \sim t_1$  或  $e \sim t_2$  和凸函数取绝对值函数时,  $\text{Var}(\sqrt{n}\hat{\beta}_n)$  不同的估计. 表 2 为凸函数取 Huber 函数时,  $\text{Var}(\sqrt{n}\hat{\beta}_n)$  不同的估计. 表 1 和表 2 中的第 1 栏是样本大小; 第 2 栏是真实方差; 第 3 栏是 Bootstrap 方差估计; 第 4、5 栏是随机权分别为指数分布、泊松分布时的方差估计. 可以看出当样本很大时, 随机加权方法与 Bootstrap 方法能较好的逼近方差  $\sigma_n^2$ . 但是随机加权方法可以较自由地选到随机权. 其中在计算过程中泊松权只用了 62.3% 的样本, 但同样达到了其它权一样的效果.

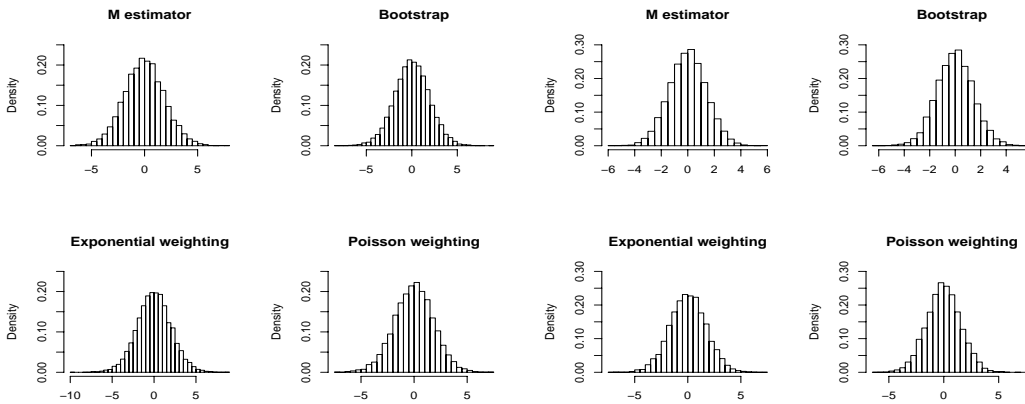


图 1

表 1

$\sqrt{n}\hat{\beta}_n$ 的渐近方差 $\sigma_{t_1 ut_1}^2 = 7.402$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	7.911	9.723	9.556	9.762
400	7.681	8.849	8.751	8.947
1000	7.542	8.279	8.175	8.162

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{l_1ut_2}^2 = 6$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	6.166	7.390	7.352	7.438
400	6.135	6.866	6.847	7.032
1000	6.068	6.636	6.619	6.529

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{l_1nt_1}^2 = 2.467$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	2.700	3.519	3.485	3.508
400	2.591	3.080	3.051	3.057
1000	2.466	2.847	2.763	2.773

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{l_1nt_2}^2 = 2$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	2.045	2.620	2.569	2.683
400	2.059	2.389	2.381	2.340
1000	2.027	2.226	2.242	2.217

表 2

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{hut_1}^2 = 8.528$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	10.997	11.911	13.174	10.601
400	10.879	11.288	12.722	10.225
1000	10.755	10.947	12.361	9.929

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{hut_2}^2 = 5.538$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	5.978	6.215	8.657	7.373
400	5.939	6.031	8.443	7.062
1000	5.753	5.926	8.389	6.859

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{hnt_1}^2 = 2.842$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	3.725	4.112	4.648	3.702
400	3.629	3.796	4.292	3.471
1000	3.553	3.602	4.202	3.285

$\sqrt{n}\widehat{\beta}_n$ 的渐近方差 $\sigma_{hnt_2}^2 = 1.846$				
$n$	$\sigma_n^2$	Bootstrap	Exponential	Poisson
200	2.035	2.126	3.054	2.504
400	1.988	2.006	2.947	2.380
1000	1.959	1.959	2.887	2.311



## 参 考 文 献

- [1] Rao C R and Zhao L C. Approximation to the distribution of M-estimates in linear models by randomly weighted bootstrap. *Sankhyā Ser. A*, 1992, **54**: 323–331.
- [2] Rubin D B. The Bayesian bootstrap. *Ann. Statist.*, 1981, **9**: 130–134.
- [3] Lo A Y. A large sample study of the Bayesian bootstrap. *Ann. Statist.*, 1987, **15**: 360–375.
- [4] Weng C S. On a second-order asymptotic property of the Bayesian bootstrap. *Ann. Statist.*, 1989, **17**: 705–710.
- [5] Anderson P K and Gill R D. Cox's regression model for counting processes: A large sample study. *Ann. Statist.*, 1982, **10**: 1100–1120.

**APPROXIMATING THE DISTRIBUTION  
OF M-ESTIMATORS IN LINEAR MODELS BY RANDOMLY  
WEIGHTED BOOTSTRAP**

WU Xiaoyan      ZHAO Lincheng      YANG Yaning

*(Department of Statistics and Finance, University of Science and Technology of China, Hefei 230026)*

**Abstract** The asymptotic distribution of the M-estimators are generally related to quantities of the error distribution that can not be conveniently estimated. The randomly weighted bootstrap method provides a way of assessing the distribution of the M-estimators without estimating the nuisance quantities of the error distributions. In this paper, the distribution of M-estimators is approximated by the randomly weighted bootstrap method in linear models when the covariates are random. It is shown that the randomly weighted bootstrapping estimation of the distribution of the M-estimator is uniformly consistent. Also, the variance estimates is investigated by Monte Carlo simulations for different choices of the convex function, sample size and random weights. Poisson weighting is recommended for reducing the computational burden in the randomly weighted bootstrapping M-estimators.

**Key words** Linear model, M-estimation, randomly weighted bootstrap.