

分位数回归的光滑经验似然 *

李忠桂* 何书元

(首都师范大学数学科学学院, 北京, 100048)

摘要: 对线性分位数回归模型参数的检验问题, 文献证明了用最小绝对距离法 (LAD 法) 及光滑经验似然法 (SEL 法) 构造的检验统计量在原假设下都依分布收敛到 χ_M^2 . 本文证明在局部备择假设下这两个检验统计量都依分布收敛到非中心卡方分布. 文中用随机模拟比较了两种方法在局部备择假设下的表现, 模拟结果表明在合理控制犯第一类错误的前提下, 用 SEL 法构造的检验比用 LAD 法构造的检验更有效.

关键词: 分位数回归; 光滑经验似然; 非中心卡方分布

中图分类号: O212.1

英文引用格式: Li Z G, He S Y. Smoothed empirical likelihood inference for quantile regression [J]. Chinese J. Appl. Probab. Statist., 2017, 33(5): 497–507. (in Chinese)

§1. 引言

分位数回归模型最早由 Koenker 和 Bassett 在 1978 年提出. 最简单的分位数回归模型是线性模型, 协变量 X 对因变量 Y 的线性 p ($p \in (0, 1)$) 分位数回归模型是

$$Y = X^\top \beta + U, \quad (1)$$

其中, 参数 β 是 M 维列向量, $P(U \leq 0 | X) = p$, $Y \in R$, $X \in R^M$ 的第一个分量为 1. 设 (Y_i, X_i) , $i = 1, 2, \dots, n$ 是 (Y, X) 的观测样本.

对模型 (1), 估计 β 的一般方法是 Koenker 和 D’Orey^[2] 提出的最小绝对距离法 (LAD 法). 该方法得到的估计为 $\hat{\beta}_{\text{LAD}} = \arg \min_{\beta} d_p(\beta)$, 其中

$$d_p(\beta) \triangleq p \sum_{X_i^\top \beta - Y_i < 0} |X_i^\top \beta - Y_i| + (1-p) \sum_{X_i^\top \beta - Y_i \geq 0} |X_i^\top \beta - Y_i|.$$

Otsu^[3] 与 Whang^[4] 将光滑经验似然法 (SEL 法) 引入模型 (1) 的参数估计和检验中. 已知 $K(u)$ 为核函数, 引入 $G_h(w) = \int_{u < w/h} K(u) du$, 设

$$Z_i(\beta) \triangleq [G_h(X_i^\top \beta - Y_i) - p] X_i. \quad (2)$$

*国家自然科学基金项目 (批准号: 11231010、11671274) 资助.

*通讯作者, E-mail: zhongguili@cnu.edu.cn.

本文 2016 年 1 月 14 日收到, 2016 年 10 月 13 日收到修改稿.

在检验 β 时, SEL 法所用统计量是

$$\begin{aligned} l_{\text{SEL}}(\beta) &\triangleq -2 \sup \left\{ \sum_{i=1}^n \ln(np_i) \mid p_i \geq 0, \sum_{i=1}^n p_i = 1, \sum_{i=1}^n p_i Z_i(\beta) = 0 \right\} \\ &= 2 \sum_{i=1}^n \ln[1 + \lambda^\top(\beta) Z_i(\beta)], \end{aligned} \quad (3)$$

其中, M 维列向量 $\lambda(\beta)$ 满足

$$\sum_{i=1}^n \frac{Z_i(\beta)}{1 + \lambda^\top(\beta) Z_i(\beta)} = 0, \quad \text{且 } \frac{1}{1 + \lambda^\top(\beta) Z_i(\beta)} > 0, \quad 1 \leq i \leq n. \quad (4)$$

LAD 法所用统计量是

$$l_{\text{LAD}}(\beta) \triangleq n(\hat{\beta}_{\text{LAD}} - \beta)^\top \Lambda_0^{-1} (\hat{\beta}_{\text{LAD}} - \beta), \quad (5)$$

其中, $\Lambda_0 \triangleq p(1-p)D_0^{-1}S_0D_0^{-1}$, $S_0 \triangleq \mathbb{E}[X_i X_i^\top]$, $D_0 \triangleq \mathbb{E}[f(0 \mid X_i) X_i X_i^\top]$, $f(u \mid X_i)$ 是 X_i 已知时 U_i 的条件密度函数.

Otsu^[3] 和 Whang^[4] 证明了在一定条件下, 检验统计量 $l_{\text{LAD}}(\beta_0)$ 和 $l_{\text{SEL}}(\beta_0)$ 在 $H_0 : \beta = \beta_0$ 下均依分布收敛到自由度为 M 的卡方分布 χ_M^2 .

记 $A_{\text{LAD}} = \{l_{\text{LAD}}(\beta_0) > \chi_{M,1-\alpha}^2\}$ 和 $A_{\text{SEL}} = \{l_{\text{SEL}}(\beta_0) > \chi_{M,1-\alpha}^2\}$, 其中 $\chi_{M,1-\alpha}^2$ 是 χ_M^2 的 $1 - \alpha$ 分位数, 以事件 A_{LAD} 和 A_{SEL} 为拒绝域分别得到 H_0 的两个检验. 下面称这两个检验为 l_{LAD} 检验和 l_{SEL} 检验. Otsu^[3] 和 Whang^[4] 的结论表明这两个检验犯第一类错误的极限概率均为 α .

以上成果研究的都是犯第一类错误的概率. 但在应用中, 掌握犯第二类错误的概率也是重要的. 本文主要研究待检验值 β_0 偏离真值 β 为 $a_n = O(n^{-1/2})$ 时, 上述检验接受 H_0 的概率, 即犯第二类错误的概率. 为此, 第 2 节证明在 $H_1 : \beta = \beta_0 + a_n$ 下统计量 $l_{\text{LAD}}(\beta_0)$ 和 $l_{\text{SEL}}(\beta_0)$ 都依分布收敛到非中心卡方分布. 第 3 节通过随机模拟比较 l_{LAD} 检验和 l_{SEL} 检验的功效, 结果表明在合理控制犯第一类错误的前提下, l_{SEL} 检验的功效更高.

§2. H_1 下检验统计量的渐近分布

对模型 (1), 本文检验 $H_0 : \beta = \beta_0$ vs $H_1 : \beta = \beta_0 + a_n$, 其中 β 是模型的固定真值, 变化的是待检验值 β_0 . 在给出本文的结论前, 参照文献 [4] 引入下面 6 个常用条件:

- C 1. (Y_i, X_i) , $i = 1, 2, \dots, n$, i.i.d. $\sim (Y, X)$;
- C 2. $Y_i = X_i^\top \beta + U_i$, $f(u \mid X_i)$ 是 X_i 已知时 U_i 的条件密度函数. 对 X_i 和 u 的取值, $f(u \mid X_i)$ 有界, 且在 $u = 0$ 的某个邻域内关于 u 存在 r 阶连续导数;
- C 3. $S_0 = \mathbb{E}[X_i X_i^\top]$ 和 $D_0 = \mathbb{E}[f(0 \mid X_i) X_i X_i^\top]$ 是非奇异矩阵;

C4. 核函数 $K(u)$ 满足

$$\int_R u^j K(u) du = \begin{cases} 1, & j = 0; \\ 0, & j = 1, 2, \dots, r-1; \\ c_r, & j = r, \end{cases}$$

其中, 偶数 $r \geq 2$, c_r 为正常数;

C5. h 满足 $nh^{2r} \rightarrow 0$, $n \rightarrow \infty$;

C6. a 与 a_n 是 R^M 中的常数向量.

在本文中, \xrightarrow{L} 表示依分布收敛, \xrightarrow{P} 表示依概率收敛, $\xrightarrow{\text{a.s.}}$ 表示几乎处处收敛.

定理 1 设 $l_{\text{SEL}}(\beta)$ 和 $l_{\text{LAD}}(\beta)$ 由 (3) 和 (5) 定义, 若条件 C1–C5 成立, 则在 $H_0: \beta = \beta_0$ 下, 当 $n \rightarrow \infty$ 时,

$$(I) \quad l_{\text{LAD}}(\beta_0) \xrightarrow{L} \chi_M^2,$$

$$(II) \quad l_{\text{SEL}}(\beta_0) \xrightarrow{L} \chi_M^2.$$

定理 1 的结论 (I) 来自文献 [5], 结论 (II) 来自文献 [3] 和 [4]. 下面两个定理是本文的主要结论.

定理 2 设 $\chi_M^2(\Delta)$ 是自由度为 M , 非中心参数为 $\Delta = a^\top \Lambda_0^{-1} a$ 的卡方分布, 其中 $\Lambda_0 = p(1-p)D_0^{-1}S_0D_0^{-1}$. 若条件 C1–C3 和 C6 成立, 则在 $H_1: \beta = \beta_0 + a_n$ 下,

$$(I) \quad \text{当 } a_n = o(n^{-1/2}) \text{ 时, } l_{\text{LAD}}(\beta_0) \xrightarrow{L} \chi_M^2, n \rightarrow \infty,$$

$$(II) \quad \text{当 } \sqrt{n} a_n \rightarrow a \text{ 时, } l_{\text{LAD}}(\beta_0) \xrightarrow{L} \chi_M^2(\Delta), n \rightarrow \infty.$$

证明: 结合 $\sqrt{n}(\hat{\beta}_{\text{LAD}} - \beta_0) = \sqrt{n}(\hat{\beta}_{\text{LAD}} - \beta) + \sqrt{n}a_n$ 和文献 [5] 的结论 $\sqrt{n}(\hat{\beta}_{\text{LAD}} - \beta) \xrightarrow{L} N(0, \Lambda_0)$ 便得到 (I) 和 (II). \square

定理 3 若条件 C1–C6 成立, 则在 $H_1: \beta = \beta_0 + a_n$ 下,

$$(I) \quad \text{当 } a_n = o(n^{-1/2}) \text{ 时, } l_{\text{SEL}}(\beta_0) \xrightarrow{L} \chi_M^2, n \rightarrow \infty,$$

$$(II) \quad \text{当 } \sqrt{n} a_n \rightarrow a \text{ 时, } l_{\text{SEL}}(\beta_0) \xrightarrow{L} \chi_M^2(\Delta), n \rightarrow \infty.$$

定理 2 和定理 3 表明若 $\sqrt{n} a_n \rightarrow a$, 则 $P(l_{\text{LAD}}(\beta_0) > \chi_{M,1-\alpha}^2 | H_1)$ 和 $P(l_{\text{SEL}}(\beta_0) > \chi_{M,1-\alpha}^2 | H_1)$ 具有相同的极限 $P(\chi_M^2(\Delta) > \chi_{M,1-\alpha}^2)$. 参照文献 [6] 和 [7] 得到在 H_1 下, l_{LAD} 检验和 l_{SEL} 检验的渐近功效函数都是

$$P(\chi_M^2(\Delta) > \chi_{M,1-\alpha}^2) \approx 1 - \Phi \left\{ \chi_{M,1-\alpha} - \sqrt{\Delta} - (M-1) \frac{\ln \chi_{M,1-\alpha}^2 - \ln \Delta}{4\chi_{M,1-\alpha}^2 - 4\sqrt{\Delta}} \right\}, \quad (6)$$

其中, $\chi_{M,1-\alpha} = \sqrt{\chi^2_{M,1-\alpha}}$, $\Phi(\cdot)$ 是标准正态分布的分布函数.

关于定理 3, (I) 是 (II) 的特殊情况, 故只需要证明 (II). 下面参照文献 [8] 中定理 1.1 的证明来证明 (II). 首先给出三条引理备用, 其中引理 4 来自文献 [8] 的引理 3, 引理 5 来自文献 [4] 的引理 1.

为方便, 引入标记 $U_{i,n} \triangleq Y_i - X_i^\top \beta_0$, 则在 H_1 下 $U_{i,n} = U_i + X_i^\top a_n$.

引理 4 设 $\zeta_i \geq 0$, $i = 1, 2, \dots, n$, i.i.d. 如果 $E[\zeta_1^2] < \infty$, 则 $\max_{1 \leq i \leq n} \zeta_i = o(n^{1/2})$ 且 $n^{-1} \sum_{i=1}^n \zeta_i^3 = o(n^{1/2})$, a.s.

引理 5 已知 $Z_i(\beta)$ 由 (2) 定义, 若条件 C1–C5 成立, 则 $E[Z_i(\beta)] = O(h^r)$, $n \rightarrow \infty$.

引理 6 记 I_A 为 A 的示性函数, 设 $G_0(u) \triangleq I_{\{u \geq 0\}}$, $Z_i \triangleq [G_0(X_i^\top \beta - Y_i) - p]X_i$. 若条件 C1–C6 成立, 且 $\sqrt{n} a_n \rightarrow a$, 则在 $H_1 : \beta = \beta_0 + a_n$ 下, 当 $n \rightarrow \infty$ 时,

- (I) $E[Z_i(\beta_0)] = E[Z_i(\beta)] - D_0 a n^{-1/2} + o(n^{-1/2})$,
- (II) $Z_i(\beta_0) \xrightarrow{\text{a.s.}} Z_i$,
- (III) $n^{-1} \sum_{i=1}^n \|Z_i - Z_i(\beta_0)\|^2 \xrightarrow{\text{a.s.}} 0$.

证明: (I) 在 H_1 下 $U_{i,n} = U_i + X_i^\top a_n$, 由 Taylor 展式知

$$\begin{aligned} E[Z_i(\beta_0)] &= E([G_h(-U_{i,n}) - p]X_i) \\ &= E\left(\left[G_h(-U_i) - \frac{1}{h}K\left(\frac{-U_i}{h}\right)X_i^\top a_n - p\right]X_i\right) + o(a_n) \\ &= E[Z_i(\beta)] - E\left[\frac{1}{h}K\left(\frac{-U_i}{h}\right)X_i X_i^\top a_n\right] + o(n^{-1/2}), \end{aligned} \quad (7)$$

已知 $h \rightarrow 0$, $f(u | X_i)$ 有界且在 $u = 0$ 处连续, 故根据换元法和控制收敛定理得

$$\begin{aligned} \sqrt{n} E\left[\frac{1}{h}K\left(\frac{-U_i}{h}\right)X_i X_i^\top a_n\right] &= E\left(E\left[\frac{1}{h}K\left(\frac{-U_i}{h}\right) | X_i\right]X_i X_i^\top\right)\sqrt{n} a_n \\ &= E\left(\left[\int_R K(y)f(-hy | X_i) dy\right]X_i X_i^\top\right)\sqrt{n} a_n \\ &\rightarrow D_0 a, \quad n \rightarrow \infty. \end{aligned}$$

结合 (7) 得到 $E[Z_i(\beta_0)] = E[Z_i(\beta)] - D_0 a n^{-1/2} + o(n^{-1/2})$.

(II) $\forall w \in (0, \infty)$, $n \rightarrow \infty$ 时

$$\sup_{u \in (w, \infty)} |G_h(u) - 1| = 1 - \int_{-\infty}^{w/h} K(y) dy \rightarrow 0.$$

故 $G_h(u)$ 在 (w, ∞) 上一致收敛于 1. 同理, $G_h(u)$ 在 $(-\infty, -w)$ 上一致收敛于 0.

因此 $\forall w \in (0, \infty)$, $G_h(u)$ 在 $(-\infty, -w) \cup (w, \infty)$ 上一致收敛于 $G_0(u)$. 这一结论结合 $U_{i,n} \rightarrow U_i$, 和 $\mathsf{P}(U_i \neq 0) = 1$, 得 $n \rightarrow \infty$ 时,

$$|G_h(-U_{i,n}) - G_0(-U_i)| \leq |G_h(-U_{i,n}) - G_0(-U_{i,n})| + |G_0(-U_{i,n}) - G_0(-U_i)| \xrightarrow{\text{a.s.}} 0.$$

从而 $Z_i(\beta_0) = [G_h(-U_{i,n}) - p]X_i \xrightarrow{\text{a.s.}} [G_0(-U_i) - p]X_i = Z_i$, $n \rightarrow \infty$.

(III) 显然

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \|Z_i - Z_i(\beta_0)\|^2 \\ &= \frac{1}{n} \sum_{i=1}^n [G_h(-U_{i,n}) - G_0(-U_i)]^2 X_i^\top X_i \\ &\leq \frac{2}{n} \sum_{i=1}^n \{[G_h(-U_{i,n}) - G_0(-U_{i,n})]^2 + [G_0(-U_{i,n}) - G_0(-U_i)]^2\} X_i^\top X_i, \end{aligned}$$

由引理 4 和条件 C3 得 $\max_{1 \leq i \leq n} |X_i^\top a_n| = o(1)$ a.s., 而 $U_{i,n} = U_i + X_i^\top a_n$, 故 $\forall \varepsilon > 0$, a.s. 有

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n [G_h(-U_{i,n}) - G_0(-U_{i,n})]^2 X_i^\top X_i \\ &= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n [G_h(-U_{i,n}) - G_0(-U_{i,n})]^2 [I_{\{|U_{i,n}| > \varepsilon\}} + I_{\{|U_{i,n}| \leq \varepsilon\}}] X_i^\top X_i \\ &\leq \lim_{n \rightarrow \infty} \left(\sup_{|w| > \varepsilon} [G_h(w) - G_0(w)]^2 \times \frac{1}{n} \sum_{i=1}^n X_i^\top X_i + \frac{1}{n} \sum_{i=1}^n I_{\{|U_i| \leq 2\varepsilon\}} X_i^\top X_i \right) \\ &= 0 \times \mathsf{E} X^\top X + \mathsf{E} I_{\{|U| \leq 2\varepsilon\}} X^\top X \\ &\rightarrow 0, \quad \varepsilon \rightarrow 0. \end{aligned}$$

同理可得 $\lim_{n \rightarrow \infty} n^{-1} \sum_{i=1}^n [G_0(-U_{i,n}) - G_0(-U_i)]^2 X_i^\top X_i = 0$, a.s., 随之便得 (III). \square

定理 3 (II) 的证明: 由于证明较长, 将其分为 (a), (b), (c), (d) 四个部分.

(a) 首先证明存在唯一的 $\lambda(\beta_0) \in R^M$ 使得

$$l_{\text{SEL}}(\beta_0) = 2 \sum_{i=1}^n \ln[1 + \lambda^\top(\beta_0) Z_i(\beta_0)], \quad (8)$$

其中 $\lambda(\beta_0)$ 满足 (4). 为证 $\lambda(\beta_0)$ 的存在唯一性, 只需证明 $\exists \delta > 0$ 使得 a.s.

$$\liminf_{n \rightarrow \infty} \inf_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top Z_i(\beta_0) > 0\}} \geq \delta, \quad (9)$$

其中 \mathcal{A} 是 R^M 的全体单位向量.

下面证明 (9). 参照文献 [9] 和 [10] 的方法, 设 Z_i 如引理 6 中定义, 则 Z_1, Z_2, \dots i.i.d., $\mathsf{E}[Z_1] = 0$ 且 $\mathsf{Var}(Z_1) = p(1-p)S_0$. 由文献 [11] 知存在 $\varepsilon_0 > 0$ 使得

$$\liminf_{n \rightarrow \infty} \inf_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top Z_i > \varepsilon_0\}} > \frac{1}{2} \inf_{\alpha \in \mathcal{A}} \mathsf{P}(\alpha^\top Z_1 > \varepsilon_0) > 0, \quad \text{a.s.} \quad (10)$$

根据不等式 $I_{\{\alpha^\top Z_i > \varepsilon_0\}} \leq I_{\{\alpha^\top Z_i(\beta_0) > \varepsilon_0/2\}} + I_{\{\alpha^\top(Z_i - Z_i(\beta_0)) > \varepsilon_0/2\}}$ 得

$$\inf_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top Z_i > \varepsilon_0\}} \leq \inf_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top Z_i(\beta_0) > \varepsilon_0/2\}} + \sup_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top(Z_i - Z_i(\beta_0)) > \varepsilon_0/2\}}. \quad (11)$$

结合 $\forall a \in R$, $I_{\{|a| > 1\}} \leq |a|^2$ 和 Hölder's 不等式得

$$\begin{aligned} \sup_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top(Z_i - Z_i(\beta_0)) > \varepsilon_0/2\}} &\leq \frac{4}{n\varepsilon_0^2} \sum_{i=1}^n \|Z_i - Z_i(\beta_0)\|^2 \\ &\xrightarrow{\text{a.s.}} 0, \quad n \rightarrow \infty. \quad (\text{见引理 6 (III)}) \end{aligned}$$

因此 a.s. $\lim_{n \rightarrow \infty} \sup_{\alpha \in \mathcal{A}} n^{-1} \sum_{i=1}^n I_{\{\alpha^\top(Z_i - Z_i(\beta_0)) > \varepsilon_0/2\}} = 0$, 结合 (10) 和 (11) 得

$$\liminf_{n \rightarrow \infty} \inf_{\alpha \in \mathcal{A}} \frac{1}{n} \sum_{i=1}^n I_{\{\alpha^\top Z_i(\beta_0) > \varepsilon_0/2\}} > \frac{1}{2} \inf_{\alpha \in \mathcal{A}} \mathbb{P}(\alpha^\top Z_1 > \varepsilon_0), \quad \text{a.s.}$$

显然 $\delta \triangleq \inf_{\alpha \in \mathcal{A}} \mathbb{P}(\alpha^\top Z_1 > \varepsilon_0)/2$ 满足 (9). 据此可知 0 点是由 $\{Z_1(\beta_0), Z_2(\beta_0), \dots, Z_n(\beta_0)\}$ 生成的凸包的内点, 因此存在唯一的 $\lambda(\beta_0)$ 满足 (8).

(b) 设 $\xi_n = \max_{1 \leq i \leq n} \|Z_i(\beta_0)\|$, 结合 $|G_h(X_i^\top \beta_0 - Y_i) - p| \leq 1$ 和引理 4 得

$$\xi_n \leq \max_{1 \leq i \leq n} \|X_i\| = o_p(n^{1/2}), \quad \text{和} \quad \frac{1}{n} \sum_{i=1}^n \|Z_i(\beta_0)\|^3 \leq \frac{1}{n} \sum_{i=1}^n \|X_i\|^3 = o_p(n^{1/2}). \quad (12)$$

定义 $S_n \triangleq n^{-1} \sum_{i=1}^n Z_i(\beta_0) Z_i(\beta_0)^\top$, 现在证明

$$S_n \xrightarrow{\mathbb{P}} p(1-p)S_0, \quad n \rightarrow \infty. \quad (13)$$

由强大数定理得

$$\frac{1}{n} \sum_{i=1}^n Z_i Z_i^\top \xrightarrow{\text{a.s.}} \mathbb{E}[Z_1 Z_1^\top] = p(1-p)S_0, \quad n \rightarrow \infty. \quad (14)$$

对矩阵 $B = (b_{ij})_{i,j=1}^M$, 定义 $|B| \triangleq \sum_{1 \leq i, j \leq M} |b_{ij}|$, 则 $\forall \varepsilon > 0$

$$\begin{aligned} \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n [Z_i(\beta_0) Z_i(\beta_0)^\top - Z_1 Z_1^\top]\right| \geq \varepsilon\right) &\leq \frac{1}{\varepsilon} \mathbb{E} \left| \frac{1}{n} \sum_{i=1}^n [Z_i(\beta_0) Z_i(\beta_0)^\top - Z_1 Z_1^\top] \right| \\ &\leq \frac{1}{\varepsilon} \mathbb{E} |Z_1(\beta_0) Z_1(\beta_0)^\top - Z_1 Z_1^\top|. \end{aligned}$$

由控制收敛定理和 $Z_i(\beta_0) \xrightarrow{\text{a.s.}} Z_i$ 知当 $n \rightarrow \infty$ 时, $\mathbb{E}|Z_1(\beta_0) Z_1(\beta_0)^\top - Z_1 Z_1^\top| \rightarrow 0$, 故

$$\frac{1}{n} \sum_{i=1}^n [Z_i(\beta_0) Z_i(\beta_0)^\top - Z_1 Z_1^\top] \xrightarrow{\mathbb{P}} 0. \quad (15)$$

(14) 与 (15) 保证 (13) 成立.

(c) 此处证明 $\bar{Z}(\beta_0) \triangleq n^{-1} \sum_{i=1}^n Z_i(\beta_0)$ 是渐近正态的

$$\sqrt{n} \bar{Z}(\beta_0) \xrightarrow{L} N(-D_0 a, p(1-p)S_0), \quad n \rightarrow \infty. \quad (16)$$

由中心极限定理得

$$\frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \xrightarrow{L} N(0, p(1-p)S_0), \quad n \rightarrow \infty. \quad (17)$$

由于 $\forall n \in N^+$, $(Z_i(\beta_0), Z_i)$, $i = 1, 2, \dots, n$, i.i.d., 并且当 $n \rightarrow \infty$ 时, $Z_i(\beta_0) \xrightarrow{\text{a.s.}} Z_i$. 故根据控制收敛定理可得 $n^{-2} \sum_{i=1}^n \text{Var}(\sqrt{n}[Z_i(\beta_0) - Z_i]) \rightarrow 0$, 结合 Chebyshev 大数定律推出

$$\frac{1}{n} \sum_{i=1}^n \sqrt{n}[Z_i(\beta_0) - Z_i] - \sqrt{n}E[Z_1(\beta_0) - Z_1] \xrightarrow{P} 0.$$

而由引理 5 和引理 6 (I) 知 $\sqrt{n}E[Z_1(\beta_0) - Z_1] = \sqrt{n}E[Z_1(\beta_0)] \rightarrow -D_0 a$, 故

$$\frac{1}{n} \sum_{i=1}^n \sqrt{n}[Z_i(\beta_0) - Z_i] \xrightarrow{P} -D_0 a. \quad (18)$$

结合 (17) 与 (18) 便得 (16) 如下

$$\sqrt{n} \bar{Z}(\beta_0) = \frac{1}{n} \sum_{i=1}^n \sqrt{n}[Z_i(\beta_0) - Z_i] + \frac{1}{\sqrt{n}} \sum_{i=1}^n Z_i \xrightarrow{L} N(-D_0 a, p(1-p)S_0).$$

(d) 在 (a), (b), (c) 的基础上, 参照文献 [8] 中定理 1.1 的证明得到定理 3 (II) 如下

$$l_{\text{SEL}}(\beta_0) = n \bar{Z}^\top(\beta_0) S_n^{-1} \bar{Z}(\beta_0) + o_p(1) \xrightarrow{L} \chi_M^2(\Delta), \quad n \rightarrow \infty.$$

证明完毕. \square

§3. 随机模拟

模拟采用文献 [4] 的中位数回归模型 ($p = 0.5$, $M = 2$)

$$Y_i = X_i^\top \beta + U_i, \quad i = 1, 2, \dots, n. \quad (19)$$

其中, $X_i = (1, X_{2,i})^\top$, $\beta = (1, 1)^\top$. $X_{2,1}, X_{2,2}, \dots, X_{2,n}$ 为均匀分布 $U[1, 5]$ 的样本. 误差 U_i 从以下 3 种分布中产生:

- (i) U_i , $i = 1, 2, \dots, n$, 服从自由度为 3 的 t 分布 $t(3)$;
- (ii) $U_i = 4V_i/(1 + X_{2,i})$, 其中 V_1, V_2, \dots, V_n , i.i.d. $\sim N(0, 1)$, 并且与 $X_{2,1}, X_{2,2}, \dots, X_{2,n}$ 独立;

(iii) $U_i, i = 1, 2, \dots, n$, i.i.d. $\sim \chi_3^2 - \chi_{3,0.5}^2$, 其中 $\chi_{3,0.5}^2$ 是 χ_3^2 的中位数.

在上面 3 种情况下, U_i 均满足 $P(U_i \leq 0 | X_i) = 0.5$. 在 (i) 与 (ii) 中, U_i 服从对称分布. 在 (i) 与 (iii) 中, U_i 与 X_i 独立. 在 (ii) 中, U_i 的分布依赖于 X_i . 在 (iii) 中, U_i 服从偏态分布. 本文对这 3 种情况分别进行模拟.

模拟时, 按 (3) 和 (5) 计算统计量 $l_{\text{SEL}}(\beta_0)$ 和 $l_{\text{LAD}}(\beta_0)$. 参照文献 [4], 本文用 $\widehat{\Lambda}$ 估计 (5) 中的 Λ_0 ,

$$\begin{cases} \widehat{\Lambda} = p(1-p)\widehat{D}^{-1}\widehat{S}\widehat{D}^{-1}, \\ \widehat{D} = n^{-1} \sum_{i=1}^n h^{-1} K((Y_i - X_i^\top \widehat{\beta}_{\text{LAD}})/h) X_i X_i^\top, \\ \widehat{S} = n^{-1} \sum_{i=1}^n X_i X_i^\top. \end{cases} \quad (20)$$

在 (3) 和 (20) 中都用到核函数 $K(u)$, Whang^[4] 分别将核函数 $K_1(u) = (3/4)(1-u^2)I(|u| < 1)$ 和 $K_2(u) = (15/16)(1-u^2)^2I(|u| < 1)$ 用于 (3) 和 (20) 中. 但是我们发现在模拟过程以下情况会频繁发生: 每次模拟产生的样本 $(Y_i, X_i), i = 1, 2, \dots, n$ 中的大多甚至全部会出现 $|Y_i - X_i^\top \widehat{\beta}_{\text{LAD}}|/h > 1$ 的情况. 为了防止此情况影响模拟结果, 本文在 (20) 中选用 Gauss 核函数 $K(u) = \exp(-u^2/2)/\sqrt{2\pi}$. 在 (3) 中, 本文也选用 Gauss 核函数.

检验的显著性水平为 $\alpha = 0.05$. 取 $h = n^{-\gamma}$, 参照文献 [7] 取 $a_n = -\delta(1, 1)^\top n^{-1/2}$. 取 $n \in \{20, 35, 50\}$, $\gamma \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.7, 0.9\}$ 和 $\delta \in \{0.5, 1, 1.5, 2, 3\}$, 则得 (n, γ, δ) 的 105 组取值. 对 (n, γ, δ) 的每一组取值都重复模拟 $N = 5000$ 次.

l_{LAD} 检验和 l_{SEL} 检验犯第一类错误的概率是 $\alpha_{\text{LAD}} \triangleq P(l_{\text{LAD}}(\beta_0) > \chi_{2,0.95}^2 | H_0)$ 和 $\alpha_{\text{SEL}} \triangleq P(l_{\text{SEL}}(\beta_0) > \chi_{2,0.95}^2 | H_0)$, 功效是 $p_{\text{LAD}} \triangleq P(l_{\text{LAD}}(\beta_0) > \chi_{2,0.95}^2 | H_1)$ 和 $p_{\text{SEL}} \triangleq P(l_{\text{SEL}}(\beta_0) > \chi_{2,0.95}^2 | H_1)$, 渐近功效都是 $p_0 \triangleq P(\chi_2^2(\Delta) > \chi_{2,0.95}^2)$. 根据 (6) 得到

$$p_0 \approx 1 - \Phi \left\{ \chi_{M,1-\alpha} - \sqrt{a^\top \Lambda_0^{-1} a} - (M-1) \frac{\ln \chi_{M,1-\alpha}^2 - \ln(a^\top \Lambda_0^{-1} a)}{4\chi_{M,1-\alpha} - 4\sqrt{a^\top \Lambda_0^{-1} a}} \right\}, \quad (21)$$

其中 $a = -\delta(1, 1)^\top$. 对 (i), (ii) 和 (iii), Λ_0 分别为

$$\begin{aligned} \Lambda_{(i)} &= \frac{p(1-p)}{f_t^2(0)} \begin{pmatrix} 1 & 3 \\ 3 & 31/3 \end{pmatrix}^{-1}, \\ \Lambda_{(ii)} &= \frac{p(1-p)}{f_N^2(0)} \begin{pmatrix} 1 & 10/3 \\ 10/3 & 37/3 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 3 \\ 3 & 31/3 \end{pmatrix} \begin{pmatrix} 1 & 10/3 \\ 10/3 & 37/3 \end{pmatrix}^{-1}, \\ \Lambda_{(iii)} &= \frac{p(1-p)}{f_\chi^2(\chi_{3,0.5}^2)} \begin{pmatrix} 1 & 3 \\ 3 & 31/3 \end{pmatrix}^{-1}, \end{aligned}$$

其中, f_t 是 $t(3)$ 的概率密度, f_N 是 $N(0, 1)$ 的概率密度, f_χ 是 χ_3^2 的概率密度.

随机模拟输出的是 α_{LAD} , α_{SEL} , p_{LAD} 和 p_{SEL} 的模拟值. 模拟结果见表 1 和表 2.

表 1 $H_0 : \beta = \beta_0$, $\alpha = 0.05$, $h = n^{-\gamma}$

U_i	γ	α_{LAD}			γ	α_{SEL}		
		$n = 20$	$n = 35$	$n = 50$		$n = 20$	$n = 35$	$n = 50$
(i) ¹	0.1	0.020	0.017	0.016	0.3	0.086	0.060	0.055
	0.2	0.049	0.045	0.039	0.5	0.079	0.059	0.055
	0.3	0.100	0.095	0.076	0.7	0.076	0.057	0.051
	0.4	0.158	0.145	0.130	0.9	0.070	0.059	0.055
(ii) ²	0.1	0.032	0.025	0.023	0.3	0.091	0.065	0.054
	0.2	0.060	0.059	0.055	0.5	0.067	0.060	0.052
	0.3	0.110	0.100	0.087	0.7	0.071	0.057	0.054
	0.4	0.166	0.151	0.144	0.9	0.074	0.059	0.053
(iii) ³	0.1	0.135	0.117	0.104	0.3	0.085	0.062	0.057
	0.2	0.220	0.176	0.147	0.5	0.080	0.055	0.059
	0.3	0.297	0.234	0.200	0.7	0.074	0.056	0.052
	0.4	0.375	0.301	0.263	0.9	0.072	0.057	0.050

¹ $U_i \sim t(3)$; ² $U_i = 4V_i/(1 + X_{2,i})$, $V_i \sim N(0, 1)$; ³ $U_i \sim \chi_3^2 - \chi_{3,0.5}^2$.

表 2 $H_1 : \beta = \beta_0 + a_n$, $a_n = -\delta n^{-1/2}(1, 1)^\top$, $h = n^{-\gamma}$, $\alpha = 0.05$

U_i	δ	p_0	γ	p_{LAD}			γ	p_{SEL}		
				$n = 20$	$n = 35$	$n = 50$		$n = 20$	$n = 35$	$n = 50$
(i) ¹	0.5	0.258	0.1	0.141	0.130	0.131	0.5	0.304	0.284	0.275
			0.2	0.227	0.211	0.215	0.7	0.282	0.260	0.257
	1.0	0.788	0.1	0.606	0.597	0.604	0.5	0.726	0.757	0.767
			0.2	0.713	0.714	0.728	0.7	0.717	0.745	0.755
(ii) ²	1.5	0.989	0.1	0.928	0.936	0.938	0.5	0.940	0.971	0.980
			0.2	0.956	0.957	0.969	0.7	0.937	0.964	0.973
	0.5	0.401	0.1	0.186	0.187	0.193	0.5	0.455	0.437	0.431
			0.2	0.320	0.326	0.326	0.7	0.437	0.425	0.419
(iii) ³	1.0	0.949	0.1	0.798	0.806	0.825	0.5	0.929	0.945	0.949
			0.2	0.863	0.885	0.900	0.7	0.907	0.936	0.945
	1.5	1.000	0.1	0.987	0.990	0.995	0.5	0.995	0.999	0.999
			0.2	0.995	0.996	0.998	0.7	0.995	0.998	0.999
	2.0	0.269	0.1	0.475	0.428	0.406	0.5	0.240	0.241	0.252
			0.2	0.537	0.475	0.453	0.7	0.238	0.243	0.246
	3.0	0.807	0.1	0.833	0.809	0.797	0.5	0.608	0.654	0.661
			0.2	0.853	0.835	0.817	0.7	0.582	0.648	0.666

¹ $U_i \sim t(3)$; ² $U_i = 4V_i/(1 + X_{2,i})$, $V_i \sim N(0, 1)$; ³ $U_i \sim \chi_3^2 - \chi_{3,0.5}^2$.

为了得到检验功效随 n 变化的曲线, 对模型 (19), 给定偏差 $\beta - \beta_0$, 将 $a = \sqrt{n}(\beta - \beta_0)$ 和 $\Lambda_0 = \Lambda_{(i)}$ (或 $\Lambda_{(ii)}$) 代入 (21) 便可作出 p_0 的近似值随 n 变化的曲线, 如图 1. 从图中可大致看出功效和样本量的关系, 例如当 $\beta_0 - \beta = (0.2, 0.2)^T$, $U_i \sim t(3)$ 时, 根据图 1(a) 推出为达到功效 0.75, 样本量要大于 23.

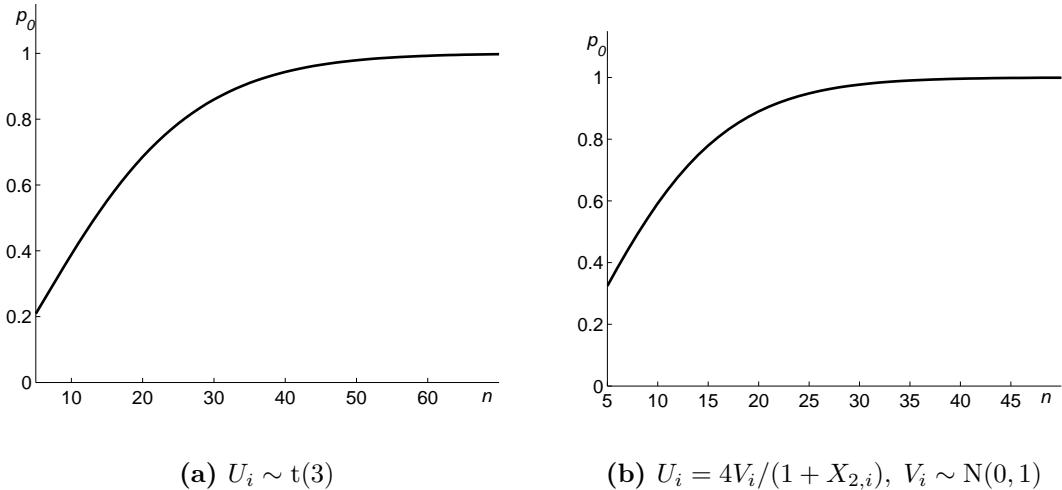


图 1 $\beta_0 - \beta = (0.2, 0.2)^T$, $\alpha = 0.05$

分析表 1 和表 2 发现:

- (a) 在表 1 中, 随着 n 增大, α_{SEL} 比 α_{LAD} 更快地接近显著水平 $\alpha = 0.05$, 表明 SEL 法的精确度更高. 这与文献 [4] 的结论一致.
- (b) 在表 1 中, 随着 γ 增大, α_{LAD} 变化比 α_{SEL} 明显, 表明 l_{LAD} 对窗宽更敏感. 原因是 (20) 用到了误差 U 的密度函数的核估计, 而密度函数的核估计对窗宽比较敏感. 对 l_{SEL} , 核函数只起光滑作用, 因此 l_{SEL} 对窗宽不如 l_{LAD} 敏感. 针对 SEL 法, Whang 推荐的窗宽是 $h = n^{-0.8}$, 因为在此窗宽下通过 Bartlett 纠偏可提高 α_{SEL} 的精确度.
- (c) 对 p_{LAD} , 表 2 只给出 $\gamma = 0.1$ 和 0.2 时的模拟值, 原因是当 $\gamma > 0.2$ 时 l_{LAD} 检验犯第一类错误的概率过大, 例如 $U_i \sim (ii)$, $n = 20$ 和 35 时 $\alpha_{LAD} \geq 0.1$, 故没必要再比较它们的功效. 表 2 显示随着 n 增大, p_{LAD} 和 p_{SEL} 向 p_0 接近, 验证了在 H_1 下 $\chi^2_2(\Delta)$ 是 $l_{LAD}(\beta_0)$ 和 $l_{SEL}(\beta_0)$ 的渐近分布.
- (d) 当 U_i 取 (i) 或 (ii) 时, 表 2 表明 l_{SEL} 检验的功效更高. 当 U_i 取 (iii) 时, 虽然 l_{LAD} 检验具有更高的功效, 但在表 1 中其犯第一类错误的概率远大于 0.05, 综合比较, 选择 l_{SEL} 检验更合适.

综上, 在合理控制犯第一类错误的前提下, SEL 法比 LAD 法更有效.

致谢 审稿人对论文所提出的宝贵意见使文章的质量有较大提高, 作者对他们表示衷心感谢. 作者也对编辑的细心工作表示衷心感谢.

参 考 文 献

- [1] Koenker R, Bassett G. Regression quantiles [J]. *Econometrica*, 1978, **46**(1): 33–50.
- [2] Koenker R, D’Orey V. Algorithm AS 229: Computing regression quantiles [J]. *J. Roy. Statist. Soc. Ser. C*, 1987, **36**(3): 383–393.
- [3] Otsu T. Empirical likelihood for quantile regression [OL]. 2003 [2003-11-01]. <http://www.cirje.e.u-tokyo.ac.jp/research/workshops/stateng/statpaper2003/otsu.pdf>.
- [4] Whang Y J. Smoothed empirical likelihood methods for quantile regression models [J]. *Econometric Theory*, 2006, **22**(2): 173–205.
- [5] Koenker R. *Quantile Regression* [M]. Cambridge, UK: Cambridge University Press, 2005.
- [6] Fraser D A S, Wu J, Wong A C M. An approximation for the noncentral chi-squared distribution [J]. *Comm. Statist. Simulation Comput.*, 1998, **27**(2): 275–287.
- [7] Wang S S, Hu T, Cui H J. Adjusted empirical likelihood inference for additive hazards regression [J]. *Comm. Statist. Theory Methods*, 2016, **45**(24): 7294–7305.
- [8] Owen A B. Empirical likelihood ratio confidence regions [J]. *Ann. Statist.*, 1990, **18**(1): 90–120.
- [9] He S Y, Liang W. Empirical likelihood for right censored data with covariables [J]. *Sci. China Math.*, 2014, **57**(6): 1275–1286.
- [10] He S Y, Liang W, Shen J S, et al. Empirical likelihood for right censored lifetime data [J]. *J. Amer. Statist. Assoc.*, 2016, **111**(514): 646–655.
- [11] Owen A B. *Empirical Likelihood* [M]. London: Chapman and Hall, 2001.

Smoothed Empirical Likelihood Inference for Quantile Regression

LI ZhongGui HE ShuYuan

(School of Mathematical Sciences, Capital Normal University, Beijing, 100048, China)

Abstract: For linear quantile regression model, this paper proves that the test statistics, based on smoothed empirical likelihood (SEL) method and least absolute deviation (LAD) method, both converge weakly to a noncentral Chi-square distribution under the local alternatives $H_1 : \beta = \beta_0 + a_n$, where β is the true parameter. Simulation results show that the SEL method is more efficient than the LAD method.

Keywords: quantile regression; smoothed empirical likelihood; noncentral Chi-square distribution

2010 Mathematics Subject Classification: 62G10; 62G20