SCIENTIA SINICA Mathematica

论 文



删失分位数回归模型中的变点检测问题

张立文1, 倪中新2, 何勇3*, 朱周帆2

- 1. 上海财经大学统计与管理学院, 上海 200433;
- 2. 上海大学经济学院, 上海 200444;
- 3. 山东财经大学统计学院, 济南 250014

E-mail: zhang.liwen@mail.shufe.edu.cn, zhongxinni@i.shu.edu.cn, heyong@sdufe.edu.cn, tylerzzf@163.com

收稿日期: 2017-12-15; 接受日期: 2018-01-31; 网络出版日期: 2018-07-11; * 通信作者 国家自然科学基金 (批准号: 11601313 和 11571218)、全国统计科学研究 (批准号: 2017LY32) 和上海市高校青年培养资助计划 (批准号: ZZSD15107) 资助项目

摘要 本文针对固定删失分位数回归模型中的变点问题提出一种新的检测方法;基于观测值的有效子集信息和分位数目标函数的次梯度提出检验统计量。在原假设下,本文得到检验统计量的渐近性质,并且通过模拟方法得到渐近分布的临界值。由于本文提出的方法仅需要在原假设下拟合模型,所以其在计算上更加有效。此外,相比较于传统的 Powell 方法,数值模拟研究发现本文提出的方法在有限样本的条件下有相近功效及更高的计算效率。最后,本文分析了一组美国居民收入数据集来展示所提方法的实际应用表现。

关键词 变点 固定删失 分位数回归 有效信息子集

MSC (2010) 主题分类 62N01, 62J99

1 引言

变点问题一直是统计学热门研究课题之一,广泛存在于经济、金融、工业质量控制和医学等各个领域. 近些年来,统计工作者在均值回归模型下的变点检测方面做了大量工作. 例如,文献 [1] 提出了基于带权重的经验分布函数的参数不变性检验. 文献 [2] 在方程组中研究了发生在未知时间的变点检测问题. 文献 [3] 考虑了在高维回归模型中由于存在协变量门限导致可能存在一个变点的问题. 文献 [4] 考虑了当回归变量和误差项非平稳时变点的估计问题. 读者可以参见文献 [5] 获得更详细的关于变点检测方面的研究. 现有的变点研究文献大多侧重于均值回归模型,而我们将在本文中考虑分位数回归模型中的变点检测问题.

自文献 [6] 提出分位数回归模型以来, 其越来越引起统计学家的兴趣和重视. 该模型在给定解释变量信息的情况下对响应变量的条件分位数进行建模. 许多文献研究过分位数回归模型中变点的检测和估计问题. 例如, 文献 [7] 提出了条件分位数回归中结构变点的两种检验方法, 一种基于次梯度统计

英文引用格式: Zhang L W, Ni Z X, He Y, et al. Testing for change points in censored quantile regression models (in Chinese).

Sci Sin Math, 2018, 48: 1159–1180, doi: 10.1360/N012017-00165

量,另一种基于 Wald 型统计量. 文献 [8] 研究了条件分位数回归模型中多个变点的估计问题. 文献 [9] 基于极大似然比方法提出了门限回归模型中门限效应的存在性检验方法,分位数回归模型是其中的一种特殊情形. 文献 [10] 提出了一种基于最大得分的方法,用于检测单个分位数回归模型或多个分位数回归模型中是否存在门限效应. 文献 [11] 在分位数回归模型下提出了一种门限效应是否为常数的一致性检验. 为了结合不同分位数水平下的信息,文献 [12] 研究了分位数回归模型中存在一个共同门限的检验和估计问题. 基于变量选择方法,文献 [13] 考虑了高维数据下分位数回归中具有一个变点的估计问题.

在现实世界中,观测数据常出现删失情形,其在医学、工业、经济和金融等各个领域中有着十分广泛的研究.越来越多的研究者对带有删失的分位数回归模型感兴趣.起先文献 [14,15] 基于分位数函数的单调不变性原理提出了删失分位数回归模型中参数估计方法 (记为 Powell 方法),并推导了所提出估计量的渐近性质.由于经典的 Powell 方法中的目标函数非凸,所以其在计算上有很大的局限性,并且模拟结果在参数维数较大或删失比例较高时表现不稳定.为此,研究者提出了一些更简单和有效的方法解决此问题,相关研究参见文献 [16-18].而关于删失分位数回归中变点检测问题的研究文献非常有限.文献 [19] 考虑了删失分位数回归模型中只存在一个变点的估计问题.文献 [20] 研究了删失数据下存在一个门限效应变点的线性变换模型.在数据存在随机删失的情形下,文献 [21] 考虑了分位数回归模型中存在协变量门限变点的检测问题.

本文研究了固定删失分位数回归模型中的变点问题并提出一种新的变点检验方法. 以往的研究 文献通常用 Powell 方法进行参数估计, 然而 Powell 方法计算困难, 并且在数据存在严重删失的情形 下表现不好 (参见文献 [15,22]). 为了弥补这个缺点, 本文提出基于观测值的有效子集信息来估计参数 (参见文献 [16,23]). 由于所提的方法仅需要在原假设下拟合模型, 故其具有易于实施、计算效率高的优点. 此外, 我们得到了原假设下检验统计量的极限分布. 通过数值模拟研究发现, 在有限样本的情形下, 本文提出的方法具有很好的功效表现. 实证研究中, 我们将本文所提的方法应用于分析 1957 到1971 年间美国南部各州白人和黑人收入数据集. 研究结果证实了白人和黑人收入差距在 1964 年民权法案颁布后发生了显著变化.

本文的内容安排如下: 第 2 节给出删失分位数回归模型中提出的变点检验方法. 原假设下检验统计量的渐近性质在第 3 节中给出. 第 4 节将渐近理论性质推广到了数据相依的情形. 同时, 第 5 节通过数值模拟研究了本文所提方法的有限样本表现. 第 6 节用美国南部各州白人和黑人收入数据集验证了本文所提方法的实际应用价值. 第 7 节给出了一个简单的小结. 所有的证明部分见附录.

2 删失分位数回归下的变点检测

2.1 模型建立与假设检验

假定 y^* 是潜在的左删失响应变量. 我们只能观测到 $y = \max(y^*, C)$, 其中 C 是固定删失点. 在给定的分位数水平 $\tau \in (0,1)$, 本文考虑以下潜在分位数回归模型:

$$y^* = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\beta}(\tau) + e,$$

其中 x 是 p 维的协变量, x 中第一维对应截距项; $\beta(\tau)$ 代表了未知的参数向量; e 是误差项, 其在给定 x 下的 τ 分位数为 0. 记 $F_y(s \mid x)$ 为给定 x 下 y 的条件分布, y 的 τ 分位数可以用 $Q_y(\tau \mid x) = F_y^{-1}(\tau \mid x) = \inf\{s: F_y(s \mid x) \geq \tau\}$ 来表示. 除此之外, 假定 y^* 和删失点 C 是条件独立的, 即对所

有的 $s \in \mathbb{R}$, 有 $P(y^* < s \mid \boldsymbol{x}, C) = P(y^* < s \mid \boldsymbol{x})$. 因此, 分位数回归模型可表示如下:

$$Q_{y^*}(\tau \mid \boldsymbol{x}, C) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{\beta}(\tau).$$

本文感兴趣的是检验协变量 \boldsymbol{x} 对 \boldsymbol{y}^* 的效应是否存在结构变化. 假设有一列随机样本 $\{(\boldsymbol{x}_t, y_t); t=1,\ldots,T\}$, 其中 t 对应的是时间指标或某个变量根据一些准则所排列的次序. 我们仅能观测到 $y_t=\max(y_t^*,C)$, 其中 C 是第 t 个观测对应的固定删失点, 且令真实误差 $e_t^0=y_t^*-\boldsymbol{x}_t^T\boldsymbol{\beta}_0(\tau)$. 不失一般性, 本文假设 C=0. 本文的研究目标在于检测出条件分位数函数中未知的变点时刻, 例如, 在第 6 节研究白人和黑人收入差距时, 关键点在于研究收入差距的条件分布是否在 1964 年民权法案颁布后发生了变化. 这里回归参数 $\boldsymbol{\beta}(\tau)$ 可能随时间变化, 因此记作 $\boldsymbol{\beta}_t(\tau)$. 此外, 响应变量 (收入的自然对数)在社会保障的最大应纳税收入水平处右缺失 (参见文献 [24]). 因此, 给定 \boldsymbol{x}_t 下 \boldsymbol{y}_t 的 τ 删失分位数回归变点模型可以记为

$$Q_{y_t}(\tau \mid \boldsymbol{x}_t) = \max\{\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_t(\tau), 0\}.$$
(2.1)

考虑的原假设和备择假设如下:

$$H_0: \boldsymbol{\beta}_t(\tau) = \boldsymbol{\beta}_0(\tau), \quad$$
対所有 $t = 1, \ldots, T,$

$$H_1: \boldsymbol{\beta}_t(\tau) = \begin{cases} \boldsymbol{\beta}_1(\tau), & \text{if } t = 1, \dots, t_0(\tau), \\ \boldsymbol{\beta}_2(\tau), & \text{if } t = t_0(\tau) + 1, \dots, T. \end{cases}$$

原假设表明回归系数中没有变点,备择假设表明存在一个变点时刻 $t_0(\tau)$. 本文第 5 节研究 1964 年民权法颁布是否对美国居民收入差距产生了影响,则该法案颁布的时间对应于可能存在的未知变点时刻 $t_0(\tau)$.

2.2 检验统计量构造

首先给出模型 (2.1) 下的参数估计. 根据分位数回归对单调变化的同变性原理, 文献 [15] 提出了参数 $\beta(\tau)$ 的 Powell 估计如下:

$$\hat{\boldsymbol{\beta}}(\tau) = \operatorname*{argmin}_{\boldsymbol{\beta}(\tau)} T^{-1} \sum_{t=1}^{T} \rho_{\tau} \{ y_{t} - \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}(\tau) \} I\{ \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{0}(\tau) > 0 \}, \tag{2.2}$$

其中 $\rho_{\tau}(u) = u\{\tau - I(u < 0)\}$ 代表文献 [6] 中提出的损失函数, $I(\cdot)$ 为示性函数. 假定 $\delta_t = I(y_t^* > 0)$ 表示删失示性函数. 在原假设下定义 $\pi_0(\boldsymbol{x}_t) = P(\delta_t = 1 \mid \boldsymbol{x}_t) = P\{e_t > -\boldsymbol{x}_t^T\boldsymbol{\beta}_0(\tau) \mid \boldsymbol{x}_t\}$, 其中 $\boldsymbol{\beta}_0(\tau)$ 为参数的真值. 参数估计 (2.2) 等价于

$$\hat{\boldsymbol{\beta}}(\tau) = \operatorname*{argmin}_{\boldsymbol{\beta}(\tau)} T^{-1} \sum_{t=1}^{T} \rho_{\tau} \{ y_{t} - \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}(\tau) \} I\{ \pi_{0}(\boldsymbol{x}_{t}) > 1 - \tau \}. \tag{2.3}$$

这表示我们可以通过对样本子集 $\{t: \pi_0(\boldsymbol{x}_t) > 1 - \tau\}$ 应用标准的分位数回归, 便可以得到一个与 Powell 估计量渐近等价的估计量. 根据文献 [16,23], 我们构造参数的两步估计量如下:

第 1 步 对 0 和 1 是否截断的指标数据用参数或非参数回归方法估计 $\pi_0(x_t)$, 记为 $\hat{\pi}(x_t)$.

第 2 步 定义最优信息子集 $J_T = \{t : \hat{\pi}(x_t) > 1 - \tau + \eta_T\}$, 其中 η_T 是事先给定的正常数, 其满足 当 $T \to +\infty$ 时, 有 $\eta_T \to 0$.

参数 η_T 的引入是为了提出所用子集的一些边界点, η_T 的收敛速度将会在第 2.3 小节中的条件 (A4) 中给出, 并且将在模拟研究中给出 η_T 的选择方法. 只要估计的删失概率 $\hat{\pi}(\cdot)$ 以 $T^{1/4}$ 的速度收敛到 $\pi_0(\cdot)$ 并且满足一定的光滑性条件, 所构造的基于有效信息子集 \mathbf{J}_T 的估计量 (2.3) 就可以达到与 Powell 估计量相同的渐近效率, 读者可参见文献 [16,23] 得到更多的细节. 因此, 基于子集 \mathbf{J}_T 下的 $\boldsymbol{\beta}(\tau)$ 的估计为

$$\hat{\boldsymbol{\beta}}(\tau) = \operatorname*{argmin}_{\boldsymbol{\beta}(\tau)} T^{-1} \sum_{t=1}^{T} \rho_{\tau} \{ y_{t} - \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}(\tau) \} I\{ \hat{\boldsymbol{\pi}}(\boldsymbol{x}_{t}) > 1 - \tau \}.$$

为得到删失分位数回归中变点的检验统计量,我们可以通过部分样本来构造,即

$$\boldsymbol{S}\{\lambda,\tau,\hat{\boldsymbol{\beta}}(\tau),\hat{\boldsymbol{\pi}}\} = \frac{1}{\sqrt{T}}\sum_{t=1}^{[T\lambda]}\boldsymbol{x}_t\psi_{\tau}\{y_t - \boldsymbol{x}_t^{\mathrm{T}}\hat{\boldsymbol{\beta}}(\tau)\}I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\},$$

其中 $\psi_{\tau}(u) = \tau - I(u < 0)$ 代表分位数目标函数的次梯度. 定义

$$\boldsymbol{H}\{\lambda,\tau,\hat{\boldsymbol{\beta}}(\tau),\hat{\boldsymbol{\pi}}\} = \hat{\boldsymbol{V}}^{-1/2}\boldsymbol{S}\{\lambda,\tau,\hat{\boldsymbol{\beta}}(\tau),\hat{\boldsymbol{\pi}}\},$$

其中 \hat{V} 代表 $S\{1,\tau,\hat{\beta}(\tau),\hat{\pi}\}$ 的渐近方差的相合估计, 具体的表达式为

$$\hat{\mathbf{V}} = \tau (1 - \tau) \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t^{\mathrm{T}} I\{\hat{\pi}(\mathbf{x}_t) > 1 - \tau + \eta_T\}.$$
(2.4)

检验统计量可以构造如下:

$$SQ_{SIS}(\tau) = \sup_{\lambda \in [0,1]} \| \boldsymbol{H}\{\lambda, \tau, \hat{\boldsymbol{\beta}}(\tau), \hat{\boldsymbol{\pi}}\} - \lambda \boldsymbol{H}\{1, \tau, \hat{\boldsymbol{\beta}}(\tau), \hat{\boldsymbol{\pi}}\} \|_{\infty},$$
(2.5)

其中 $\|\cdot\|_{\infty}$ 为向量无穷范数,即对任何向量 $\mathbf{z} = (z_1, \dots, z_k)$,有 $\|\mathbf{z}\|_{\infty} = \max(|z_1|, \dots, |z_k|)$. 在原假设下, $\hat{\boldsymbol{\beta}}(\tau)$ 为回归系数向量的收敛速度 \sqrt{T} 的估计量, $\mathbf{H}\{\lambda, \tau, \hat{\boldsymbol{\beta}}(\tau), \hat{\pi}\}$ 将收敛到一个非退化分布. 然而,备择假设下的估计量将远离真实参数,使得检验统计量取一个较大的值,从而该方法更容易将原假设和备择假设区别开来. 文献 [16] 中给出的方法并未考虑删失数据存在的情形,而本文所提方法利用观测值的子集信息去处理删失. 本文提出的方法 (今后记为 SIS (subgradient and informative subset) 方法)也是现存的 Powell 估计方法的进一步简化.

2.3 渐近性质

本节将给出检验统计量在原假设下的渐近分布. 定义 $F(\cdot \mid \boldsymbol{x})$ 和 $f(\cdot \mid \boldsymbol{x})$ 分别为响应变量 g 给定 g 不的条件分布函数和密度函数. 为简化记号,本文分别将 $f(\cdot \mid \boldsymbol{x}_t)$ 和 $f(\cdot \mid \boldsymbol{x}_t)$ 简记为 $f(\cdot \mid \boldsymbol{x}_t)$ 简记为 $f(\cdot \mid \boldsymbol{x}_t)$ 和 $f(\cdot \mid \boldsymbol{x}_t)$ 和 $f(\cdot \mid \boldsymbol{x}_t)$ 简记为 $f(\cdot \mid \boldsymbol{x}_t)$ 和 $f(\cdot \mid \boldsymbol{x}_t)$ 和

$$\|h\|_{\infty,p+\alpha} = \max_{|\boldsymbol{a}| \leqslant p} \sup_{\boldsymbol{x}} |D^{\boldsymbol{a}}h(\boldsymbol{x})| + \max_{|\boldsymbol{a}| = p} \sup_{\boldsymbol{x} \neq \boldsymbol{x}_1} \frac{|D^{\boldsymbol{a}}h(\boldsymbol{x}) - D^{\boldsymbol{a}}h(\boldsymbol{x}_1)|}{\|\boldsymbol{x} - \boldsymbol{x}_1\|^{\alpha}},$$

其中 $0<\alpha\leqslant 1$ 为某个正常数. 记 $C_l^{p+\alpha}(\mathbb{R}_x)$ 为包含所有连续函数 h $(h:\mathbb{R}_x\to\mathbb{R})$ 的集合, 这里 l 是某个正常数. 为得到检验统计量的渐近分布, 本文给出下列假设条件:

- (A1) 参数向量 $\beta_0(\tau)$ 是紧参数空间 \mathbb{B} 的一个内点.
- (A2) 下面给出一些关于分布函数和密度函数的条件:
- (A2.1) 分布函数 $F_t(\cdot)$ 绝对连续且密度函数 $f_t(\cdot)$ 连续, 并且对所有的 t 满足 $0 < L_f \le f_t\{F_t^{-1}(\tau)\}$ $\le U_f < \infty$;
- (A2.2) 对任何 $\epsilon > 0$, 存在一个 $\sigma(\epsilon) > 0$, 使得对所有的 $|s| < \sigma(\epsilon)$, 以及 $1 \le t \le T$, 有 $|f_t\{F_t^{-1}(\tau) + s\} f_t\{F_t^{-1}(\tau)\}| < \epsilon$ 成立.
 - (A3) 下面给出一些关于协变量的条件:
 - (A3.1) 协变量 x_t 定义在一个有界的凸支撑集 \mathbb{R}_x 上. 除此之外, $\mathrm{E}(x_tx_t^{\mathrm{T}})$ 是一个 $p\times p$ 正定矩阵;
 - (A3.2) 当 $T \to +\infty$ 时, 在 $\lambda \in [0,1]$ 上一致地有

$$\frac{1}{T} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \boldsymbol{x}_t^{\mathrm{T}} I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau\} \xrightarrow{p} \lambda \boldsymbol{M}_0,$$

其中 M_0 是一个 $p \times p$ 维正定矩阵;

(A3.3) 当 $T \to +\infty$ 时, 在 $\lambda \in [0,1]$ 上一致地有

$$\frac{1}{T} \sum_{t=1}^{[T\lambda]} f_t \{ \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0(\tau) \} \boldsymbol{x}_t \boldsymbol{x}_t^{\mathrm{T}} I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau \} \xrightarrow{p} \lambda \boldsymbol{J}_0,$$

其中 J_0 是一个 $p \times p$ 维正定矩阵;

- (A3.4) 对任何当 T 趋于正无穷时趋于 0 的非负序列 ϵ_T , 最小特征值 $\inf_t \lambda_{\min}(\mathbb{E}[f_t\{\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0(\tau)\}\boldsymbol{x}_t\boldsymbol{x}_t^{\mathrm{T}}]$ $I\{\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0(\tau) > \epsilon_T\}$) 大于 0;
 - (A4) η_T 满足 $\lim_{T\to\infty}\eta_T=0$ 以及 $T^{1/4}\eta_T$ 大于某个正常数 c^* .
 - (A5) 下面给出一些关于条件概率 $\pi_0(x)$ 及其估计量的条件:
- (A5.1) 对任何正数 $\epsilon_T \to 0$, 当 $\epsilon_T/\eta_T \to 1$ 和 $\boldsymbol{x} \in \mathbb{R}_{\boldsymbol{x}}$ 时, $\pi_0(\boldsymbol{x}) > 1 \tau + \epsilon_T$ 隐含 $\boldsymbol{x}^T \boldsymbol{\beta}_0(\tau) > \epsilon_T^*$, 其中 ϵ_T^* 满足 $\epsilon_T = O(\epsilon_T^*)$.
 - (A5.2) 对某些正常数 $\alpha \in (0,1]$ 和有限数 l, 有 $P(\pi_0(\boldsymbol{x}), \hat{\pi}(\boldsymbol{x}) \in C_l^{p+\alpha}(\mathbb{R}_{\boldsymbol{x}})) \to 1$.
 - (A6) 对任何 $\epsilon_T \to 0$, 当 $\|\beta(\tau) \beta_0(\tau)\| \leqslant \epsilon_T$ 时,

$$\mathbb{E}(\boldsymbol{x}_t I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \} [I \{ \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}(\tau) \leqslant 0 \} - I \{ \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0(\tau) \leqslant 0 \}]) = -\boldsymbol{Q}_T \{ \boldsymbol{\beta}(\tau) - \boldsymbol{\beta}_0(\tau) \},$$

其中 Q_T 是一个半正定矩阵且满足对充分大的 T, 有

$$0 \leqslant \lambda_{\min}(\boldsymbol{Q}_T) \leqslant \lambda_{\max}(\boldsymbol{Q}_T) < \infty.$$

(A7) 定义 V_T 为 $\frac{1}{\sqrt{T}}\sum_{t=1}^T \{\tau - I(e_t^0 \leq 0)\} x_t$ 的协方差函数, 假设极限 $V := \lim_{T \to \infty} V_T$ 存在且是非奇异的.

条件 (A1) 定义了参数向量 $\hat{\beta}(\tau)$ 的取值范围. 条件 (A2) 表示条件密度函数在 $F_t^{-1}(\tau)$ 附近光滑, 且对 $t=1,\ldots,T$ 一致成立. 条件 (A3) 给出了关于协变量的条件, 其中条件 (A3.3) 给出了模型中异方 差的约束条件, 该条件与文献 [2, 条件 3(e)] 一样. 事实上, 条件 (A3.3) 可以通过条件 (A2.1) 和 (A3.2) 以及下列条件得到:

$$\sup_{T\geq 2} \frac{1}{T} \sum_{t=1}^{T-1} t[f_{t+1}\{F_{t+1}^{-1}(\tau)\} - f_t\{F_t^{-1}(\tau)\}] < \infty,$$

这一结论易由文献 [25, 定理 1] 和 [26, 定理 2.1] 推导出,详细细节可参见文献 [2]. 本文第 5 节的模拟研究中,我们在产生模拟数据时考虑了如下线性位置尺度模型:

$$y_i^* = \boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\beta}(\tau) + \{\boldsymbol{x}_i^{\mathrm{T}} \boldsymbol{\zeta}(\tau)\} e_i,$$

其中 $\zeta(\tau) \in \mathbb{R}^p$. 模型中的异方差是线性的, 条件 (A3.3) 可以重新写为

$$\frac{1}{T} \sum_{t=1}^{[T\lambda]} \frac{\boldsymbol{x}_t \boldsymbol{x}_t^{\mathrm{T}}}{\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\zeta}(\tau)} I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau\} \xrightarrow{p} \lambda \boldsymbol{J}_0,$$

显然, 在这种情形下, 条件 (A2.1) 和 (A3.2) 能够保证上述条件成立. 条件 (A4) 给出了模型中参数的一些假定条件. 条件 (A5) 类似于文献 [16] 中的条件, 给出了关于函数 $\pi_0(x)$ 的一些假设. 条件 (A6) 类似于文献 [16] 中的假设 (A6), 其本质上是积分中值定理的一个应用. 通过将示性函数近似为一个单调光滑函数, 我们可以认为 \mathbf{Q}_T 可以被近似为一列非负定矩阵. 当 $\epsilon_T = o(T^{-1/4})$ 时, 在条件 (A1) 和 (A5.1) 下, 有

$$\mathbf{Q}_T\{\boldsymbol{\beta}(\tau) - \boldsymbol{\beta}_0(\tau)\} = o(\|\boldsymbol{\beta}(\tau) - \boldsymbol{\beta}_0(\tau)\|).$$

条件 (A7) 定义了方差协方差矩阵.

下面的定理给出了检验统计量的极限分布.

定理 2.1 假设条件 (A1)-(A7) 成立且 $\|\hat{\pi} - \pi_0\|_{\infty} = o_p(T^{-1/4})$. 在原假设下有

$$SQ_{SIS}(\tau) \Rightarrow \sup_{\lambda \in [0,1]} ||B_p(\lambda)||_{\infty},$$

其中" \Rightarrow "定义为弱收敛, B_p 是定义在 [0,1] 上的 p 维独立 Brown 桥过程.

注 2.1 检验统计量的渐近分布与文献 [7] 中的渐近理论分布相同, 而区别在于我们提出的检验统计量涉及子集信息 $I\{\hat{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}$. 渐近分布 $\sup_{\lambda \in [0,1]} \|B_p(\lambda)\|_{\infty}$ 的临界值可以通过数值模拟得到. 具体地, 过程 $\|B_p(\lambda)\|_{\infty}$ 可以通过 $\|T^{-1/2}(\sum_{t=1}^{[T\lambda]} \boldsymbol{e}_t - \lambda \sum_{t=1}^T \boldsymbol{e}_t)\|_{\infty}$ 模拟得到, 其中 $\boldsymbol{e}_t \sim N(\boldsymbol{0}, \boldsymbol{I}_p)$, 而在区间 $\lambda \in [0,1]$ 上对 $\|B_p(\lambda)\|_{\infty}$ 取上确界可以通过格子点搜索得到. 模拟研究中样本量和重复次数分别设定为 500 和 50,000. 文献 [7, 表 1] 给出了 $p \leq 10$ 时渐近分布的临界值.

注 2.2 定理 2.1 给出了检验统计量在原假设下的渐近分布, 类似于文献 [10, 推论 1], 我们给出统计量在局部备择假设模型下的渐近分布, 即

$$H_{1T}: \boldsymbol{\beta}_t(\tau) = \boldsymbol{\beta}_0(\tau) + T^{-1/2} b_T \boldsymbol{\gamma}_0(\tau) I\{t > t_0(\tau)\},$$

其中 $\gamma_0(\tau) \subset \mathbb{R}^p$. 如果 $b_T \to \infty$, 则对所有的 t > 0, 有

$$\lim_{T \to \infty} P\{\|SQ_{SIS}(\tau)\| \ge t\} = 1.$$

3 与 Powell 方法的比较

本文也可以根据 Powell 方法 [15] 提供另外一种变点存在性的检验过程. 检验统计量的构造如下:

$$\mathrm{SQ}_{\mathrm{Powell}}(\tau) = \sup_{\lambda \in [0,1]} \| \boldsymbol{H}\{\lambda, \tau, \hat{\boldsymbol{\beta}}_{\mathrm{Powell}}(\tau), \hat{\boldsymbol{\pi}}\} - \lambda \boldsymbol{H}\{1, \tau, \hat{\boldsymbol{\beta}}_{\mathrm{Powell}}(\tau), \hat{\boldsymbol{\pi}}\} \|_{\infty},$$

其中

$$\begin{split} & \boldsymbol{H}\{\lambda, \tau, \hat{\boldsymbol{\beta}}_{\text{Powell}}(\tau), \hat{\boldsymbol{\pi}}\} = \hat{\boldsymbol{V}}_{\text{Powell}}^{-1/2} \boldsymbol{S}\{\lambda, \tau, \hat{\boldsymbol{\beta}}_{\text{Powell}}(\tau), \hat{\boldsymbol{\pi}}\}, \\ & \hat{\boldsymbol{V}}_{\text{Powell}} = \tau(1-\tau) \frac{1}{T} \sum_{t=1}^{T} \boldsymbol{x}_t \boldsymbol{x}_t^{\text{T}} I\{\boldsymbol{x}_t^{\text{T}} \hat{\boldsymbol{\beta}}_{\text{Powell}}(\tau) > 0\}, \\ & \boldsymbol{S}\{\lambda, \tau, \hat{\boldsymbol{\beta}}_{\text{Powell}}(\tau), \hat{\boldsymbol{\pi}}\} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \psi_{\tau} \{y_t - \boldsymbol{x}_t^{\text{T}} \hat{\boldsymbol{\beta}}_{\text{Powell}}(\tau)\} I\{\boldsymbol{x}_t^{\text{T}} \hat{\boldsymbol{\beta}}_{\text{Powell}}(\tau) > 0\}, \end{split}$$

且 $\hat{\boldsymbol{\beta}}_{Powell}(\tau)$ 表示 Powell 估计量. 需要注意的是, 本文所提方法与 Powell 方法的主要区别在于不同的子集示性函数 $I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t)>1-\tau+\eta_T\}$ 和 $I\{\boldsymbol{x}_t^{\mathrm{T}}\hat{\boldsymbol{\beta}}_{Powell}(\tau)>0\}$. 为了给出 Powell 估计量的渐近性质, 本文给出以下条件:

- (B1) 下面给出关于参数向量的一些条件:
- (B1.1) 协变量 $\{x_t\}$ 是独立分布的, 且存在某个正数 K_0 , 对所有 t, $E||x_t|| < K_0$ 成立.
- (B1.2) 存在某些正数 ϵ_0 、 ν_0 和 T_0 , 当 $T > T_0$ 时, 有最小特征值

$$\lambda_{\min}\bigg(\mathrm{E}\bigg[\frac{1}{T}\sum_{t=1}^{T}\boldsymbol{x}_{t}\boldsymbol{x}_{t}^{\mathrm{T}}I\{\boldsymbol{x}_{t}^{\mathrm{T}}\boldsymbol{\beta}_{0}(\tau)>\epsilon_{0}\}\bigg]\bigg)>\nu_{0}.$$

(B1.3) 定义

$$G_t(z, \boldsymbol{\beta}, r) = \mathbb{E}_{\boldsymbol{x}}\{\|\boldsymbol{x}_t\|^r I(|\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}(\tau)| \leqslant \|\boldsymbol{x}_t\|z)\},$$

存在一个正数 ξ_0 , 使得

$$G_t(z, \boldsymbol{\beta}, r) \leqslant K_1 z$$

对任意的 t 一致成立, 其中 $0 \le z < \xi_0$, $\|\beta(\tau) - \beta_0(\tau)\| < \xi_0$, 并且对常数 K_1 和 ξ_0 , 有 $r \le 2$.

(B2) 对条件 (A.2.1) 进行补充, 密度函数 $f_t(\cdot)$ 是 Lipschitz 连续的.

条件 (B1) 类似于文献 [14, 条件 E.1 和 R.1], 保证了删失分位数回归模型下参数的 Powell 估计量的渐近正态性. 我们有以下定理:

定理 3.1 假设条件 (B1) 和 (B2) 成立,则在原假设下,有

$$SQ_{Powell}(\tau) \Rightarrow \sup_{\lambda \in [0,1]} ||B_p(\lambda)||_{\infty}.$$

根据定理 2.1 和 3.1, 本文提出的基于得分方法的估计量与 Powell 方法的估计量有着相同的渐近分布. 这主要是因为两种方法的参数估计有着相似的渐近性质. 但是本文提出的 SIS 方法在计算方面具有更大的优势, 这将会在第 5 节进行阐述.

4 推广

本节将提出的方法推广到允许序列相关时的情形. 为了给出该情形下 SIS 方法估计量的渐近性质, 本文给出以下条件:

(C1) $\{(y_t, x_t), t \ge 1\}$ 是平稳的且是 β - 混合序列. 定义 $\beta_1(j)$ 为 $\{(y_t, x_t), t \ge 1\}$ 的 β - 混合系数. 故存在常数 $a \in (0,1)$ 和 B > 0 使得 $\sup_{t \ge 1} \beta_t(j) \le Ba^j$ 对所有 $j \ge 1$ 成立.

(C2) 定义 $f_j(u_1, u_{1+j} \mid \boldsymbol{x}_1, \boldsymbol{x}_{1+j})$ 为给定 $(\boldsymbol{x}_{t1}, \boldsymbol{x}_{t,1+j}) = (\boldsymbol{x}_1, \boldsymbol{x}_{1+j})$ 下 (u_1, u_{1+j}) 的条件概率. 假设存在常数 $C_f^{\mathrm{T}} > 0$, 使得 $f_j(u_1, u_{1+j} \mid \boldsymbol{x}_1, \boldsymbol{x}_{1+j}) \leqslant C_f^{\mathrm{T}}$ 在 $(u_1, u_{1+j}, \boldsymbol{x}_1, \boldsymbol{x}_{1+j})$ 上对所有 $j \geqslant 1$ 一致成立.

(C3) 定义
$$V_t^*$$
 为 $\frac{1}{\sqrt{T}} \sum_{t=1}^T \{ \tau - I(e_t \leq 0) \} x_t I \{ \pi(x_t) > 1 - \tau + \eta_T \}$ 的协方差矩阵, 故

$$\boldsymbol{V}^* := \frac{1}{T} \sum_{t=1}^T \boldsymbol{V}_t^*$$

存在且非奇异.

条件 (C1) 与文献 [23,27] 中的条件相似. 条件 (C2) 对条件密度提出了一些新的约束. 条件 (C3) 定义了长期的方差 - 协方差矩阵.

本文记在时间相依情形下提出的检验统计量分别为 $SQ_{SIS}^*(\tau)$ 和 $SQ_{Powell}^*(\tau)$. 值得注意的是, 考虑到时间上的相依性, 本文在构造的检验统计量中用 V^* 替换了方差 - 协方差矩阵 V, 其中 V^* 由 \hat{V}^* 估计. 本文有以下的渐近理论结果.

定理 4.1 在原假设 H_0 下, 对 $T \to \infty$,

(i) 假设条件 (A1)-(A7) 和 (C1)-(C3) 成立, 则有

$$SQ_{SIS}^*(\tau) \Rightarrow \sup_{\lambda \in [0,1]} ||B_p(\lambda)||_{\infty}.$$

(ii) 假设条件 (B1)、(B2) 和 (C1)-(C3) 成立, 则有

$$\mathrm{SQ}^*_{\mathrm{Powell}}(\tau) \Rightarrow \sup_{\lambda \in [0,1]} \|B_p(\lambda)\|_{\infty}.$$

根据定理 4.1, 我们发现在序列相关时基于得分方法的估计量与 Powell 方法的估计量有着相同的渐近分布. 定理 4.1 与定理 2.1 和 3.1 的区别在于方差的估计上, 前者的估计只需要在独立条件下得到即可, 而后者需要在序列相关条件下得到.

5 模拟研究

本节中,为研究所提方法的有限样本表现,我们给出下面的数值模拟研究.我们的模拟数据来自下列模型:

$$y_t^* = 10x_{1,t} - 2x_{2,t} + \{1 + (x_{1,t} + x_{2,t})\varsigma\}e_t, \quad t = 1, \dots, T,$$

其中 $\mathbf{x}_t = (x_{1,t}, x_{2,t})^{\mathrm{T}}$ 服从独立的标准正态分布, 并满足截断方式 $\{\mathbf{x}_t : \|\mathbf{x}_t\|_{\infty} < 2\}$. 我们仅观测到 $y_t = \max(y_t^*, C)$. 常数 C 分别取为 -5.10 和 0, 这对应于删失比例大致为 30% 和 50%.

本文考虑四种模型. 模型 1 和 3 保持同方差性, 取 $\varsigma = 0$; 而在模型 2 和 4 中取 $\varsigma = 0.5$, 使模型 具有异方差性. 模型 1 和 2 中的误差项来自标准正态分布 N(0,1) 或者自由度为 3 的 t 分布 t(3). 类似于文献 [12], η_T 可以选取为 $T^{-1/5}\tau$. 本文考虑三个分位数水平, 即 τ 取值为 0.25、0.50 和 0.75, 分别对应于低、中和高分位数. 模拟次数设为 1,000 次. 本文在模拟研究中考虑 T 取值为 500、1,000 和 1,500 三种样本量大小, 而且 δ_t 用非参数方式来估计, 通过 R 软件中的 "gam" 函数得到.

为比较不同方法的表现,本文给出其他两种检验方法的模拟结果, Naive 方法和 Powell 方法 (记为 POW). 对 Naive 方法,其可在假定数据没有删失条件下通过文献 [7] 中提出的次梯度检验方法得到. Powell 估计可以通过 R 程序包 quantreg 中的函数 "crq" 得到.

1X I 1+1	↑里 I — 300	时们们松业及作	在佚至 1 王 4	1、113年一天相关11	500权, 40人小十	以且为 0.03
au	比例	方法	模型 1	模型 2	模型 3	模型 4
0.25	30%	SIS	0.042	0.059	0.040	0.041
		Powell	0.045	0.058	0.044	0.047
		Naive	0.044	0.036	0.035	0.038
	50%	SIS	0.044	0.048	0.035	0.043
		Powell	0.050	0.059	0.042	0.046
		Naive	0.036	0.037	0.035	0.036
0.50	30%	SIS	0.049	0.054	0.046	0.051
		Powell	0.056	0.055	0.052	0.054
		Naive	0.045	0.047	0.050	0.046
	50%	SIS	0.056	0.054	0.051	0.048
		Powell	0.050	0.056	0.049	0.056
		Naive	0.050	0.050	0.062	0.055
0.75	30%	SIS	0.062	0.064	0.054	0.051
		Powell	0.050	0.055	0.051	0.044
		Naive	0.213	0.220	0.196	0.185
	50%	SIS	0.049	0.059	0.043	0.037
		Powell	0.061	0.059	0.040	0.039
		Naive	0.219	0.225	0.210	0.213

表 1 样本量 T=500 时不同检验过程在模型 1 至 4 下的第一类错误的比较. 名义水平设置为 0.05

表 1 中给出了 T = 500 时提出方法在不同模型下的第一类错误,名义水平设定为 0.05. T = 1,000 和 T = 1,500 时本文提出方法在不同模型下的第一类错误在表 B1 和 B2 中给出. 我们可以发现所有考虑的模型中,本文所提方法的第一类错误表现良好,并且随着样本量的增加更接近名义水平. 其次,相比较带有异方差的模型 2 和 4,所提方法在同方差的模型 1 和 3 下表现得更好. 最后,总体来说, SIS 和 POW 表现类似,但比 Naive 方法好. 同时,表 2 给出了三种不同检验过程在模型 1 至 4 下一次模拟的计算时间 (秒). 我们可以发现 SIS 方法比 POW 方法计算上更加有效,特别在样本量变大时.

为了研究提出的检验方法的经验功效,本文考虑模型

$$y_t^* = 10x_{1,t} - 2x_{2,t} + d(10x_{1,t} - 2x_{2,t})I(t > t_0(\tau)) + \{1 + (x_{1,t} + x_{2,t})\varsigma\}e_t,$$

其中变点位置为 $t_0(\tau) = T/2$,参数 d 度量了备择假设远离原假设的程度, d=0 表示为模型中没有变点. 图 1 给出了三种检验方法在不同 d 下的功效图. 首先, 所有考虑的模型下随着 d 的增大检验的势逐渐增大. 其次, SIS 和 POW 功效表现特别接近, 这也证实了定理 2.1 和 3.1 中两种方法有相同的渐近分布. 此外, 它们比 Naive 方法表现都好, 这是因为后者忽视了删失. 第三, 检验方法在 C=-5.10时的功效表现比 C=0 时的功效表现更好. 这主要是因为前者对应于更小的删失比例 30%. 最后, 相比较于模型 2, 检验方法在模型 1 下表现得更好, 这是因为模型 2 中的异方差影响了参数估计的准确性. 我们也可以发现在模型 3 和 4 下有相同的结论.

为了比较三种检验过程的计算效率, 表 2 给出了在分位数水平 $\tau = 0.5$ 和固定删失点 C = 0 下, 不同检验过程一次模拟所需要的计算时间 (以秒为单位). 因为 Naive 方法没有考虑删失, 所以, 该方法的计算时间是三种方法中最少的. 因为 Powell 方法包含了最小化非凸目标函数的过程, SIS 方法比 Powell 方法具有更好的计算效率, 特别是样本量大的情形下.

表 2 在分位数水平 $\tau=0.5$ 以及固定删失点 C=0 下,不同检验过程在一次模拟中的计算时间 (以秒计算,括号中为标准差)

	T = 500			T = 1,000			T = 1,500		
	SIS	Powell	Naive	SIS	Powell	Naive	SIS	Powell	Naive
模型 1	0.15 (0.05)	0.10 (0.03)	0.05 (0.02)	0.34 (0.05)	0.38 (0.07)	0.13 (0.02)	0.52 (0.05)	0.95 (0.19)	0.27 (0.06)
模型 2	0.21 (0.05)	0.10 (0.02)	0.05 (0.02)	0.39 (0.05)	0.39 (0.07)	0.13 (0.02)	0.60 (0.09)	0.87 (0.23)	0.26 (0.06)
模型 3	0.15 (0.04)	0.10 (0.02)	0.05 (0.02)	0.29 (0.04)	0.39 (0.07)	0.13 (0.02)	0.44 (0.04)	0.94 (0.20)	0.27 (0.06)
模型 4	0.16 (0.04)	0.10 (0.02)	0.05 (0.02)	0.31 (0.05)	0.39 (0.08)	0.13 (0.02)	0.47 (0.05)	0.92 (0.21)	0.27 (0.06)

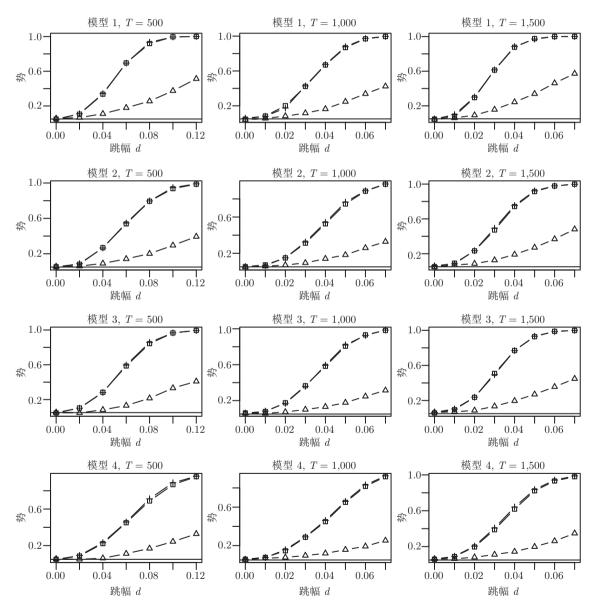


图 1 分位数水平 $\tau=0.5$ 和固定删失点 C=0 时不同检验过程在模型 1 至 4 下对 d 的势函数. 正方形、十字形和三角线分别代表 SIS、POW 和 Naive 方法

实证分析 6

美国白人和黑人收入差距一直是经济研究中的重要领域, 许多学者也都研究过这个课题. 例如, 文 献 [28] 考虑了删失数据模型下白人和黑人的收入差距. 文献 [29] 给出了劳动市场上存在种族歧视的 直接证据. 众所周知, 美国 1964 年民权法案的第七章是为了禁止种族和性别歧视而颁布的, 而实际上 这个政策是在 1965 年七月正式生效. 文献 [24] 通过删失回归模型讨论了白人和黑人收入差距. 并且 发现这种差距在政策后显著减小. 文献 [23] 用删失分位数回归模型讨论了白人和黑人的收入差距在 政策颁布后的变化,并且发现,相较于成熟群体,这个政策对年轻群体的收入差距影响更大,我们利用 提出的方法来检验白人和黑人的收入差距是否在 1964 年民权法案的颁布前后发生变化. 我们利用来 自文献 [24] 的数据, 这些数据在最高应纳税收入水平处右删失, 根据文献 [30], 我们认为年龄在 1957 年为 31 至 43 岁的群体为成熟群体. 在 1957 年到 1971 年这段时间, 我们从年轻和成熟的群体中随机 取 700 个样本, 值得注意的是, 删失机制在数值模拟和实际数据分析中是有区别的, 前者是基于左删 失, 而后者是右删失.

给定分位数水平 τ, 考虑下列带有变点的删失分位数回归模型:

$$Q_{y_t}(\tau \mid \boldsymbol{x}_t) = \begin{cases} \min\{\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_1(\tau), C\}, & t = 1, \dots, t_0(\tau), \\ \min\{\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_2(\tau), C\}, & t = t_0(\tau) + 1, \dots, T, \end{cases}$$

其中 $y_t = \min(y_t^*, C)$ 为观测变量, 即为收入的自然对数, 协变量 x_t 包括截距、种族 $(1 = \mathbb{R})$ 0 =白人)、年龄、年龄平方和教育. 参数 $\beta_i(\tau) = (a_i(\tau), b_i(\tau), c_i(\tau), d_i(\tau), e_i(\tau))^{\mathrm{T}}$ (j = 1, 2) 是第 j 段 对应的回归参数, $t_0(\tau)$ 是变点位置. 定义 $\boldsymbol{\theta}(\tau) = (\boldsymbol{\beta}_1^{\mathrm{T}}(\tau), \boldsymbol{\beta}_2^{\mathrm{T}}(\tau), t_0(\tau))^{\mathrm{T}}$. 由于在此实际数据分析中我 们更关心收入的低分位数和中位数的分布,因此考虑两个分位数水平 $\tau = 0.25$ 和 $\tau = 0.50$. 调节参数 η_T 选为 $T^{-1/5}\tau$. 我们给出了 SIS、Powell 和 Naive 方法在不同分位数水平下变点检验的 p 值. 所有的 变点检验方法都给出检验的 p 值为 0, 说明了分位数回归函数中的确存在变点. 为了更深入地研究变 点的情况, 我们提出下面的估计过程. 根据文献 [23,31], 参数可以通过下式估计:

$$\hat{\boldsymbol{\theta}}(\tau) = \operatorname*{argmin}_{\boldsymbol{\theta}(\tau)} T^{-1} \sum_{t=1}^{T} \rho_{\tau} \{ y_{t} - Q_{y_{t}}(\tau \mid \boldsymbol{x}_{t}) \} I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_{t}) < 1 - \tau + \eta_{T} \}.$$

表 3 给出了参数的估计值和对应的标准差. 我们发现在所有的分位数水平下变点位置的估计都为 1965 年左右, 这实际上表明了 1964 年民权法案颁布的影响. 同时发现在所有分位数水平下, 种族这一协变 量的影响变小, 进而说明了 1964 年民权法案的颁布有助于减少种族歧视.

表 3 个问分位数条件下,1964 年氏仪法条颁布后日人和黑人収入差距数据的分析结果												
au		第一段						第二段				变点
		$a_1(\tau)$	$b_1(au)$	$c_1(\tau)$	$d_1(au)$	$e_1(au)$	$a_2(\tau)$	$b_2(\tau)$	$c_2(\tau)$	$d_2(\tau)$	$e_2(\tau)$	
0.25	估计	9.3539	-0.3301	-0.0247	0.0004	0.0413	9.8755	-0.1745	-0.0316	0.0004	0.0193	1965
	标准差	(0.5756)	(0.0257)	(0.0294)	(0.0004)	(0.0022)	(0.3602)	(0.0063)	(0.0153)	(0.0002)	(0.0012)	
0.50	估计	8.6490	-0.0024	0.0395	-0.0004	0.0037	6.7382	-0.0102	0.1074	-0.0009	0.0077	1965
	标准差	(0.1585)	(0.0398)	(0.0082)	(0.0001)	(0.0006)	(0.3661)	(0.0450)	(0.0160)	(0.0002)	(0.0017)	

7 结论

本文研究了删失分位数回归模型中的变点检验问题,基于有效信息子集的方法提出了检验统计量并给出相应的渐近性质.同时本文将所提方法与 Powell 方法作比较.数值模拟研究发现,相比较于传统的 Powell 方法,本文提出的方法有相近功效和更高的计算效率.然而,本文尚存在一些值得改进的地方.例如,可以研究删失分位数回归下面板分位数模型中的变点问题,然而这超出了本文的研究范围,我们将在未来给出.

致谢 作者非常感谢审稿人以及副主编提出的宝贵修改意见和建议,这让文章的质量有了很大的提升.

参考文献 —

- 1 Bai J. Testing for parameter constancy in linear regressions: An empirical distribution function approach. Econometrica, 1996, 64: 597–622
- 2 Qu Z, Perron P. Estimating and testing structural changes in multivariate regressions. Econometrica, 2007, 75: 459–502
- 3 Lee S, Seo M H, Shin Y. The Lasso for high dimensional regression with a possible change point. J R Stat Soc Ser B Stat Methodol. 2016, 78: 193–210
- 4 Horváth L, Hušková M, Rice G, et al. Asymptotic properties of the cusum estimator for the time of change in linear panel data models. Econometric Theory, 2017, 33: 366–412
- 5 Csörgó M, Horvá L. Limit Theorems in Change-Point Analysis. New York: John Wiley & Sons, 1997
- 6 Koenker R, Bassett J G. Regression quantiles. Econometrica, 1978, 46: 33–50
- 7 Qu Z. Testing for structural change in regression quantiles. J Econometrics, 2008, 146: 170-184
- 8 Oka T, Qu Z. Estimating structural changes in regression quantiles. J Econometrics, 2011, 162: 248-267
- 9 Lee S, Seo M H, Shin Y. Testing for threshold effects in regression models. J Amer Statist Assoc, 2011, 106: 220–231
- 10 Zhang L W, Wang H J, Zhu Z Y. Testing for change points due to a covariate threshold in quantile regression. Statist Sinica, 2014, 24: 1859–1877
- 11 Galvao A F, Kato K, Montes-Rojas G, et al. Testing linearity against threshold effects: Uniform inference in quantile regression. Ann Inst Statist Math, 2014, 66: 413–439
- 12 Su L, Xu P. Common threshold in quantile regressions with an application to pricing for reputation. Econometric Rev, 2017, 29: 1–37
- 13 Lee S, Liao Y, Seo M H, et al. Oracle estimation of a change point in high-dimensional quantile regression. J Amer Statist Assoc. 2017, 7: 1–11
- 14 Powell J L. Least absolute deviations estimation for the censored regression model. J Econometrics, 1984, 25: 303–325
- 15 Powell J L. Censored regression quantiles. J Econometrics, 1986, 32: 143-155
- 16 Tang Y L, Wang H J, He X M, et al. An informative subset-based estimator for censored quantile regression. TEST, 2012, 21: 635-655
- 17 Volgushev S, Wagener J, Dette H. Censored quantile regression processes under dependence and penalization. Electron J Stat, 2014, 8: 2405–2447
- 18 Kobayashi G. Bayesian endogenous Tobit quantile regression. Bayesian Anal, 2017, 12: 161-191
- 19 Wang Z F, Wu Y H, Zhao L C. Change-point estimation for censored regression model. Sci China Ser A, 2007, 50: 63–72
- 20 Kosorok M R, Song R. Inference under right censoring for transformation models with a change point based on a covariate threshold. Ann Statist, 2007, 35: 967–989
- 21 Tang Y L, Song X Y, Zhu Z Y. Threshold effect test in censored quantile regression. Statist Probab Lett, 2015, 105: 149–156
- 22 Fitzenberger B. A guide to censored quantile regressions. In: Handbook of Statistics 15: Robust Inference. Amsterdam: North-Holland, 1997, 405–437
- 23 Galvao A F, Lamarche C, Lima L R. Estimation of censored quantile regression for panel data with fixed effects. J Amer Statist Assoc, 2013, 108: 1075–1089
- 24 Chay K Y, Powell J L. Semiparametric censored regression models. J Econom Perspect, 2001, 15: 29–42
- 25 Etemadi N. Convergence of weighted averages of random variables revisited. Proc Amer Math Soc, 2006, 134: 2739–2744
- 26 Newey W.K. Uniform convergence in probability and stochastic equicontinuity. Econometrica, 1991, 59: 1161-1167

- 27 Hahn J, Kuersteiner G. Bias reduction for dynamic nonlinear panel models with fixed effects. Econometric Theory, 2011, 27: 1152–1191
- 28 Brown C. Black-white earnings ratios since the civil rights act of 1964: The importance of labor market dropouts. Q J Economics, 1984, 99: 31–44
- 29 Altonji J, Blank R. Race and gender in the labor market. In: Handbook of Labor Economics, vol. 3. Amsterdam: North-Holland, 1999, 3143–3259
- 30 Levine P B, Mitchell O S. The baby boom's legacy: Relative wages in the twenty-first century. Amer Econom Rev, 1988, 72: 66–69
- 31 Li C, Wei Y, Chappell R, et al. Bent line quantile regression with application to an allometric study of land mammals' speed and mass. Biometrics, 2011, 67: 242–249
- 32 Bai J, Lumsdaine R L, Stock J H. Testing for and dating common breaks in multivariate time series. Rev Econom Stud, 1998, 65: 395–432
- 33 Merlevède F, Peligrad M. The functional central limit theorem under the strong mixing condition. Ann Probab, 2000, 28: 1336–1352

附录 A 主要理论结果的证明

在本文中, 为了表示方便, 我们在 $\beta(\tau)$ 、 $\beta_0(\tau)$ 以及 $\pi(x)$ 、 $\hat{\pi}(x)$ 的表示中省略 τ 和 x. 此外, 定义 $V_0 = \tau(1-\tau)\mathrm{E}[x_1x_1^{\mathrm{T}}I\{\hat{\pi}_0(x_1)>1-\tau\}]$. 定理 3.1 的证明与定理 2.1 的证明类似, 因此只给出定理 2.1 和 4.1 的证明.

注意到

$$\psi_{\tau}(y - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}) = \tau - I(y - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} < 0)$$
$$= \tau - I(y^* > 0, y^* - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} < 0) - I(y^* < 0, \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} > 0).$$

由此有

$$P(y < \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} \mid \boldsymbol{x}) = P(y^* < \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} \mid \boldsymbol{x})I(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} > 0),$$

$$E\{\psi_{\tau}(y - \boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta}) \mid \boldsymbol{x}\} = \tau - F(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} \mid \boldsymbol{x})I(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\beta} > 0).$$

我们给出下列定义:

$$\begin{split} \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}, \pi) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \psi_{\tau}(y_t - \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}) I\{\pi(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}, \\ \boldsymbol{S}_d(\lambda, \tau, \boldsymbol{\beta}, \pi) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ \mathrm{P}(y_t < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta} \mid \boldsymbol{x}_t) - I(y_t - \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta} < 0) \} I\{\pi(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}. \end{split}$$

因此,

$$E\{S(1,\tau,\boldsymbol{\beta},\pi)\} = E[\boldsymbol{x}I\{\pi(\boldsymbol{x}) > 1 - \tau + \eta_T\}\{\tau - P(y^* < \boldsymbol{x}^T\boldsymbol{\beta} \mid \boldsymbol{x})I(\boldsymbol{x}^T\boldsymbol{\beta} > 0)\}].$$

首先给出下面两个有用的引理.

引理 A.1 假定条件 (A1)–(A6) 成立. 对任何正的序列 $\epsilon_T \to 0$, 在原假设下, 有

$$\sup_{\lambda \in [0,1]} \sup_{\|\pi - \pi_0\|_{\infty} \leqslant \epsilon_T} \sup_{\boldsymbol{\xi} \in \boldsymbol{D}} \|\boldsymbol{S}_d(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi}}, \pi) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_0, \pi_0)\| = o_p(1),$$

其中 D 为 \mathbb{R}^p 中任何一个紧集.

证明 假定 p 维协变量为 $x_t = (x_{t,1}, \ldots, x_{t,p})^{\mathrm{T}}$. 不失一般性,假定 $x_{t,j}$ $(j = 1, \ldots, p)$ 是非负的,否则 $x_{t,j}$ 可以写为 $x_{t,j} = x_{t,j}^+ - x_{t,j}^-$, 其中 $x_{t,j}^+$ 和 $x_{t,j}^-$ 分别定义 $x_{t,j}$ 的正部和负部. 因为 \mathbf{D} 是一个紧集,对任何给定的 $\delta > 0$,我们可以把它分割成有限个集合,记为 $\{\mathbf{D}_h; h = 1, \ldots, T(\delta)\}$,并使得每个子集的直径小于或等于 δ . 对任何 $\boldsymbol{\xi} \in \mathbf{D}_h$ $(h = 1, 2, \ldots, T(\delta))$,在 \mathbf{D}_h 的边界上存在两个点 $\boldsymbol{\xi}_{h,1}$ 和 $\boldsymbol{\xi}_{h,2}$ 使得 $x_t \boldsymbol{\xi}_{h,1} \leq x_t \boldsymbol{\xi} \leq x_t \boldsymbol{\xi}_{h,2}$. 定义

$$\boldsymbol{\beta_{\xi,h,k}} = \boldsymbol{\beta_0} + \frac{1}{\sqrt{T}} \boldsymbol{\xi_{h,k}}, \quad k = 1, 2,$$

因此有 $x_t^{\mathrm{T}} \beta_{\boldsymbol{\xi},h,1} \leq x_t^{\mathrm{T}} \beta_{\boldsymbol{\xi}} \leq x_t^{\mathrm{T}} \beta_{\boldsymbol{\xi},h,2}$. 因为 $x_t I(y_t < x_t^{\mathrm{T}} \boldsymbol{\beta})$ 是 $\boldsymbol{\beta}$ 的非降函数, 所以有

$$\begin{split} & \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi}}, \pi) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_{0}, \pi_{0}) \\ & \geqslant \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ \mathrm{P}(y_{t} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_{t}) - I(y_{t} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}, h, 2}) \} \\ & \times I \{ \pi(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \} - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_{0}, \pi_{0}) \\ & = \{ \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi}, h, 2}, \pi) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_{0}, \pi_{0}) \} + \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \\ & \times \{ \mathrm{P}(y_{t} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_{t}) - \mathrm{P}(y_{t} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}, h, 2} \mid \boldsymbol{x}_{t}) \} I \{ \pi(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \}. \end{split}$$

通过将 $\beta_{\boldsymbol{\xi},h,2}(\tau)$ 替换为 $\beta_{\boldsymbol{\xi},h,1}(\tau)$, 有

$$S_d(\lambda, \tau, \boldsymbol{\beta_{\xi}}, \pi) - S(\lambda, \tau, \boldsymbol{\beta_{0}}, \pi_0) \leqslant \{S_d(\lambda, \tau, \boldsymbol{\beta_{\xi,h,1}}, \pi) - S(\lambda, \tau, \boldsymbol{\beta_{0}}, \pi_0)\} + \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{P(y_t < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta_{\xi}} \mid \boldsymbol{x}_t) - P(y_t < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta_{\xi,h,1}} \mid \boldsymbol{x}_t)\} I\{\pi(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}.$$

当 $z_1 \leqslant z_2 \leqslant z_3$ 时, $|z_2| \leqslant \max(|z_1|, |z_3|)$, 因此,

$$\sup_{\lambda \in [0,1]} \sup_{\|\pi - \pi_0\|_{\infty} \leqslant \epsilon_T} \sup_{\boldsymbol{\xi} \in \boldsymbol{D}} \|\boldsymbol{S}_d(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi}}, \pi) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_0, \pi_0)\| \leqslant (a) + (b), \tag{A.1}$$

其中

$$(a) = \sup_{\lambda \in [0,1]} \sup_{\|\boldsymbol{\pi} - \boldsymbol{\pi}_0\|_{\infty} \leqslant \epsilon_T} \max_{1 \leqslant h \leqslant T(\delta)} \max_{k=1,2} \left\| \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} \mid \boldsymbol{x}_t) - P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) \} I\{\boldsymbol{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T \} \right\|,$$

$$(b) = \sup_{\lambda \in [0,1]} \sup_{\|\boldsymbol{\pi} - \boldsymbol{\pi}_0\|_{\infty} \leqslant \epsilon_T} \max_{1 \leqslant h \leqslant T(\delta)} \max_{k=1,2} \|\boldsymbol{S}_d(\lambda, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \boldsymbol{\pi}) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_0, \boldsymbol{\pi}_0) \|.$$

(i) 对 (a) 项, 注意到 $P(y_t < \mathbf{x}_t^T \boldsymbol{\beta} \mid \mathbf{x}_t) = F_t(\mathbf{x}_t^T \boldsymbol{\beta})I(\mathbf{x}_t^T \boldsymbol{\beta} > 0)$, 通过一些简单计算, 有

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \mathbf{x}_t \{ P(y_t < \mathbf{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} \mid \mathbf{x}_t) - P(y_t < \mathbf{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \mathbf{x}_t) \} I\{\pi(\mathbf{x}_t) > 1 - \tau + \eta_T \}$$

$$= (a1) + (a2) + (a3), \tag{A.2}$$

其中

$$(a1) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ F_{t}(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}) - F_{t}(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}}) \} I(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} > 0) I\{\boldsymbol{\pi}_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \},$$

$$(a2) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} F_{t}(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}) \{ I(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} > 0) - I(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} > 0) \} I\{\boldsymbol{\pi}_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \},$$

$$(a3) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ F_{t}(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}) I(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} > 0) - F_{t}(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}}) I(\boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} > 0) \} [I\{\boldsymbol{\pi}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \}].$$

对 (a1) 项, 根据条件 (A5.1), $\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T$ 隐含 $\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0 > \epsilon_T^*$, 其中 $T^{1/4}\epsilon_T$ 大于某个正数 ϵ^* . 根据中值定理, 有

$$(a1) = \frac{1}{T} \sum_{t=1}^{[T\lambda]} f_t(x_t \tilde{\beta}) x_t x_t I\{\pi_0(x_t) > 1 - \tau + \eta_T\} (\xi_{h,k} - \xi),$$

其中 $\tilde{\boldsymbol{\beta}}$ 介于 $\boldsymbol{\beta_{\xi}}$ 与 $\boldsymbol{\beta_{\xi,h,k}}$ 之间. 因为 $\|\boldsymbol{\xi_{h,k}} - \boldsymbol{\xi}\| \le \delta$, 条件 (A3.1) 和 (A5) 成立, 所以可以通过选择一个小的 $\boldsymbol{\delta}$ 使得项 (*a*1) 任意小. 因此,

$$(a1) = o_p(1). (A.3)$$

对 (a2) 项, 有

$$\begin{aligned} \|(a2)\| &\leqslant \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t F_t(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}) |I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} > 0) - I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} > 0)| \\ &\times I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} \\ &\leqslant \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} < 0 < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}) + I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} < 0 < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}})\} \\ &\times I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}. \end{aligned}$$

根据条件 (A3.2) 和 (A6), 类似于文献 [2] 中引理 A.2 的证明, 我们有

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} < 0 < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k})$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} \{I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} > 0) - I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} > 0)\}$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} [\{I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} > 0) - I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 > 0)\}$$

$$+ \{I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 > 0) - I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} > 0)\}]$$

$$= 2\lambda \boldsymbol{Q}_T(\boldsymbol{\beta}_{\boldsymbol{\xi},h,k} - \boldsymbol{\beta}_0) = o_p(1).$$

类似地, 可以得到

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi},h,k} < 0 < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}}) = o_p(1).$$

因此,

$$(a2) = o_p(1). (A.4)$$

对 (a3) 项, 类似于对 (a1) 的证明, 有

$$|(a3)| = o_p(1). (A.5)$$

结合 (A.2)-(A.5), 可以得到

$$(a) = o_p(1). (A.6)$$

(ii) 对项 (b), 有

$$S_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \pi) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_{0}, \pi_{0})$$

$$= \{ \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \pi_{0}) - \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \pi_{0}) \} + \{ \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \pi_{0}) - \boldsymbol{S}(\lambda, \tau, \boldsymbol{\beta}_{0}, \pi_{0}) \}$$

$$= \{ \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \pi_{0}) - \boldsymbol{S}_{d}(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi},h,k}, \pi_{0}) \} + \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0}),$$
(A.7)

其中

$$\eta_t(\tau, \xi_{h,k}, \pi_0) = [\{P(y_t < x_t^T \beta_{\xi,h,k} \mid x_t) - I(y_t - x_t^T \beta_{\xi,h,k} < 0)\} \\
- \{\tau - I(y_t < x_t^T \beta_0)\}]I\{\pi_0(x_t) > 1 - \tau + \eta_T\}.$$

对 (A.7) 中第一项, 类似地有 $S_d(\lambda, \tau, \beta_{\xi,h,k}, \pi_0) - S_d(\lambda, \tau, \beta_{\xi,h,k}, \pi_0) = O_p(1)$. 通过应用文献 [32] 第 618 页中的 Doob 不等式和 Rosenthal 不等式, 对任何 T, 以及任何正的常数 ϵ , 对 (A.7) 中的第二项, 有

$$P\left(\frac{1}{\sqrt{T}} \sup_{\lambda \in [0,1]} \left\| \sum_{t=1}^{[T\lambda]} \boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0}) \right\| > \epsilon \right)$$

$$\leq \frac{C_{1}}{\epsilon^{2\gamma}} \left(\frac{1}{T^{\gamma}} E\left[\sum_{t=1}^{T} E\{\|\boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0})\|^{2} \mid \boldsymbol{x}_{t}\} \right]^{\gamma} + \frac{1}{T^{\gamma}} E\left[\sum_{t=1}^{T} E\{\|\boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0})\|^{2\gamma}\} \mid \boldsymbol{x}_{t} \right] \right), \tag{A.8}$$

对任何正的常数 $\gamma > 1$, C_1 是仅依赖于 p 和 γ 的某个常数. 根据条件 (A2.1) 和中值定理, 有

$$E(\|\boldsymbol{\eta}_{t}(\tau,\boldsymbol{\xi}_{h,k},\pi_{0})\|^{2} | \boldsymbol{x}_{t}) \leq \|\boldsymbol{x}_{t}\|^{2} \left| F_{t}\left(\boldsymbol{x}_{t}^{\mathrm{T}}\boldsymbol{\beta}_{0} + \frac{1}{\sqrt{T}}\boldsymbol{x}_{t}\boldsymbol{\xi}_{h,k}\right) - \tau \right|$$

$$= \frac{1}{\sqrt{T}} \|\boldsymbol{x}_{t}\|^{3} |f(\tilde{\boldsymbol{\xi}}_{t})| \|\boldsymbol{\xi}_{h,k}\|$$

$$\leq \frac{1}{\sqrt{T}} U_{f}\delta \|\boldsymbol{x}_{t}\|^{3},$$

其中 $\tilde{\boldsymbol{\xi}}_t$ 介于 $\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0$ 与 $(\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0 + \frac{1}{\sqrt{T}}\boldsymbol{x}_t\boldsymbol{\xi}_{h,k})$ 之间. 因此, 对 (A.8) 中的第一项, 根据条件 (A3.1), 当 $T \to \infty$ 时, 有

$$\frac{1}{T^{\gamma}} \mathbb{E} \left[\sum_{t=1}^{T} \mathbb{E} \{ \| \boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0}) \|^{2} \mid \boldsymbol{x}_{t} \} \right]^{\gamma} \leqslant \frac{1}{T^{3\gamma/2}} U_{f}^{\gamma} \delta^{\gamma} \mathbb{E} \left[\sum_{t=1}^{T} \| \boldsymbol{x}_{t} \|^{3} \right]^{\gamma} = o_{p}(1).$$
(A.9)

因为 $\|\eta_t(\tau, \xi_{h,k}, \pi_0)\| \le 4\|x_t\|$, 对 (A.8) 中的第二项, 当 $T \to \infty$ 时,

$$\frac{1}{T^{\gamma}} \mathbf{E} \left[\sum_{t=1}^{T} \mathbf{E} \{ \| \boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0}) \|^{2\gamma} \mid \boldsymbol{x}_{t} \} \right] = \frac{1}{T^{\gamma}} \mathbf{E} \left[\sum_{t=1}^{T} (4 \| \boldsymbol{x}_{t} \|)^{2\gamma - 2} \mathbf{E} \{ \| \boldsymbol{\eta}_{t}(\tau, \boldsymbol{\xi}_{h,k}, \pi_{0}) \|^{2} \mid \boldsymbol{x}_{t} \} \right]
= \frac{1}{T^{3\gamma/2}} 4^{2\gamma - 2} U_{f} \delta \sum_{t=1}^{T} \| \boldsymbol{x}_{t} \|^{2\gamma + 2}
= o_{p}(1),$$
(A.10)

进而结合 (A.8)-(A.10) 即可得到, 对任何 $\epsilon > 0$, 有

$$P\left(\frac{1}{\sqrt{T}}\sup_{\lambda\in[0,1]}\left\|\sum_{t=1}^{[T\lambda]}\boldsymbol{\eta}_t(\tau,\boldsymbol{\xi}_{h,k},\pi_0)\right\| > \epsilon\right) \to 0.$$

因此,

$$(b) = o_p(1). \tag{A.11}$$

因此, 结合 (A.1)-(A.11) 便可得到引理的结论. 这就完成了引理 A.1 的证明.

引理 A.2 假设 $\sup \|\hat{\pi} - \pi_0\|_{\infty} = o_p(T^{-1/4})$ 以及条件 (A1)–(A6) 成立. 在原假设下, 在 $\lambda \in [0,1]$ 上一致地有

$$S(\lambda, \tau, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\pi}}) = S(\lambda, \tau, \boldsymbol{\beta}_0, \boldsymbol{\pi}_0) + \lambda T^{1/2} J_0(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + o_p(1). \tag{A.12}$$

证明 通过直接计算,对任何 $\boldsymbol{\xi} \in \boldsymbol{D}$,有

$$S(\lambda, \tau, \beta_{\xi}, \hat{\pi}) = S(\lambda, \tau, \beta_0, \pi_0) + (c) + (d), \tag{A.13}$$

其中

$$(c) = S_d(\lambda, \tau, \boldsymbol{\beta}_{\boldsymbol{\xi}}, \hat{\pi}) - S(\lambda, \tau, \boldsymbol{\beta}_0, \pi_0),$$

$$(d) = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ \tau - P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) \} I\{\hat{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}.$$

对 (c) 项, 在假设 $\sup \|\hat{\pi} - \pi_0\|_{\infty} = o_p(T^{-1/4})$ 成立条件下, 根据引理 A.1, 有

$$(c) = o_p(1).$$
 (A.14)

对(d)项,其可以等价地表示为

$$(d) = (d1)_{\varepsilon} + (d2)_{\pi} + (d3)_{\varepsilon,\pi}, \tag{A.15}$$

其中

$$(d1)_{\boldsymbol{\xi}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ \tau - P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) \} I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \},$$

$$(d2)_{\pi} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ \tau - P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_0 \mid \boldsymbol{x}_t) \} [I \{ \hat{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T \}$$

$$- I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \}],$$

$$(d3)_{\boldsymbol{\xi},\pi} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) - P(y_t < \boldsymbol{x}_t^T \boldsymbol{\beta}_0 \mid \boldsymbol{x}_t) \}$$

$$\times [I \{ \hat{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T \} - I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \}].$$

$$(A.17)$$

对项 $(d1)_{\xi}$, 通过一些简单计算, 有

$$\begin{split} (d1)_{\xi} &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ \tau - \mathrm{P}(y_{t} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\xi} \mid \boldsymbol{x}_{t}) \} I \{ \pi_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \} \\ &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ \mathrm{P}(y_{t}^{*} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{0} \mid \boldsymbol{x}_{t}) - \mathrm{P}(y_{t} < \boldsymbol{x}_{t}^{\mathrm{T}} \boldsymbol{\beta}_{\xi} \mid \boldsymbol{x}_{t}) \} I \{ \pi_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T} \} \\ &= (d11)_{\xi} + (d12)_{\xi}, \end{split}$$

其中

$$(d11)_{\boldsymbol{\xi}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ P(y_t^* < \boldsymbol{x}_t^T \boldsymbol{\beta}_0 \mid \boldsymbol{x}_t) - P(y_t^* < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) \}$$

$$\times I\{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \},$$

$$(d12)_{\boldsymbol{\xi}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t P(y_t^* < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) I(\boldsymbol{x}_t^T \boldsymbol{\beta}_0 \leqslant 0)$$

$$\times I\{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \},$$

$$(A.18)$$

最后一个等式是根据 Taylor 展开式, 以及 $\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T$ 隐含 $\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 > 0$, 由此可得

$$P(y_t < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta_{\xi}} \mid \boldsymbol{x}_t) I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} = P(y_t^* < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta_{\xi}} \mid \boldsymbol{x}_t) I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta_{\xi}} > 0) I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}.$$

对 (d11), 根据均值定理有

$$(d11)_{\boldsymbol{\xi}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ F_t(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0) - F_t(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}}) \} I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \}$$
$$= -\frac{1}{T} \sum_{t=1}^{[T\lambda]} f_t \left(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 + \frac{1}{\sqrt{T}} \boldsymbol{x}_t \boldsymbol{\xi}^* \right) \boldsymbol{x}_t \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\xi} I \{ \pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \},$$

其中 $\|\boldsymbol{\xi}^*\| \leq \|\boldsymbol{\xi}\|$. 根据条件 (A2), 有

$$f_t \bigg(oldsymbol{x}_t^{\mathrm{T}} oldsymbol{eta}_0 + rac{1}{\sqrt{T}} oldsymbol{x}_t^{\mathrm{T}} oldsymbol{\xi}^* \bigg) = f_t (oldsymbol{x}_t^{\mathrm{T}} oldsymbol{eta}_0) + o_p(1).$$

根据条件 (A3.3) 有

$$(d11)_{\boldsymbol{\xi}} = -\frac{1}{T} \sum_{t=1}^{[T\lambda]} f_t(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0) \boldsymbol{x}_t \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\xi} I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} + o_p(1)$$
$$= -\lambda J_0 \boldsymbol{\xi} + o_p(1),$$

其中根据条件 (A3.2), 得到最后一个等式的根据是

$$\left\| \frac{1}{T} \sum_{t=1}^{[T\lambda]} f_t(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0) \boldsymbol{x}_t \boldsymbol{x}_t I\{1 - \tau < \pi_0(\boldsymbol{x}_t) < 1 - \tau + \eta_T\} \boldsymbol{\xi} \right\| = O_p(\eta_T \|\boldsymbol{\xi}\|) = o_p(\|\boldsymbol{\xi}\|).$$

对 (d12) , 根据条件 (A6), 有

$$(d12)_{\boldsymbol{\xi}} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t P(\boldsymbol{y}_t^* < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) I(\boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \leqslant 0) I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}$$

$$= \tau \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ I(\boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \leqslant 0) - I(\boldsymbol{x}_t^T \boldsymbol{\beta}_0 \leqslant 0) \} I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}$$

$$= -\tau \boldsymbol{Q}_{\boldsymbol{\xi}}.$$

因此,

$$(d1)_{\boldsymbol{\xi}} = (-\lambda \boldsymbol{J}_0 - \tau \boldsymbol{Q})\boldsymbol{\xi}. \tag{A.19}$$

对 $(d2)_{\pi}$ 项, 有

$$(d2)_{\pi} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ \tau - P(y_{t} < \boldsymbol{x}_{t}^{T} \boldsymbol{\beta}_{0} \mid \boldsymbol{x}_{t}) \} [I\{\hat{\pi}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T}\} - I\{\pi_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T}\}].$$

因为 $\|\hat{\pi} - \pi_0\|_{\infty} = o_p(T^{-1/4})$, 所以, $\hat{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T$ 隐含 $\pi_0(\boldsymbol{x}_t) > 1 - \tau$, 因此有 $\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0 > 0$. 同样 $\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T$ 隐含 $\boldsymbol{x}_t^{\mathrm{T}}\boldsymbol{\beta}_0 > 0$, 因此,

$$(d2)_{\pi} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \{ \tau - P(y_t < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 \mid \boldsymbol{x}_t) \} [I\{\pi(\hat{\boldsymbol{x}}_t) > 1 - \tau + \eta_T\} - I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}]$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \{ \tau - P(y_t^* < \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 \mid \boldsymbol{x}_t) \} [I\{\hat{\pi}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} - I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}]$$

$$= 0. \tag{A.20}$$

对 $(d3)_{\xi,\pi}$ 项,有

$$(d3)_{\boldsymbol{\xi},\pi} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ -P(y_t^* < \boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \mid \boldsymbol{x}_t) I(\boldsymbol{x}_t^T \boldsymbol{\beta}_{\boldsymbol{\xi}} \leqslant 0) + P(y_t^* < \boldsymbol{x}_t^T \boldsymbol{\beta}_0 \mid \boldsymbol{x}_t) I(\boldsymbol{x}_t^T \boldsymbol{\beta}_0 \leqslant 0) \}$$

$$\times [I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} - I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}]$$

$$= (d31)_{\boldsymbol{\xi},\pi} + (d32)_{\boldsymbol{\xi},\pi},$$

其中

$$(d31)_{\boldsymbol{\xi},\pi} = \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} \{ P(\boldsymbol{y}_{t}^{*} < \boldsymbol{x}_{t}^{T} \boldsymbol{\beta}_{\boldsymbol{\xi}} \boldsymbol{x}_{t} \mid \boldsymbol{x}_{t}) - P(\boldsymbol{y}_{t}^{*} < \boldsymbol{x}_{t}^{T} \boldsymbol{\beta}_{0} \boldsymbol{x}_{t} \mid \boldsymbol{x}_{t}) \} I(\boldsymbol{x}_{t}^{T} \boldsymbol{\beta}_{0} \leqslant 0)$$

$$\times [I\{\hat{\pi}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T}\} - I\{\pi_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T}\}],$$

$$(d32)_{\boldsymbol{\xi},\pi} = \tau \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_{t} [I(\boldsymbol{x}_{t}^{T} \boldsymbol{\beta}_{\boldsymbol{\xi}} \leqslant 0) - I(\boldsymbol{x}_{t}^{T} \boldsymbol{\beta}_{0} \leqslant 0)]$$

$$\times [I\{\hat{\pi}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T}\} - I\{\pi_{0}(\boldsymbol{x}_{t}) > 1 - \tau + \eta_{T}\}].$$

$$(A.21)$$

对 $(d31)_{\boldsymbol{\xi},\pi}$ 项, 有

$$(d31)_{\boldsymbol{\xi},\pi} = \frac{1}{T} \sum_{t=1}^{[T\lambda]} f_t(\boldsymbol{x}_t^{\mathrm{T}} \tilde{\boldsymbol{\beta}}) \boldsymbol{x}_t \boldsymbol{x}_t \boldsymbol{\xi} I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 \leqslant 0)$$
$$\times [I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} - I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}],$$

其中 $\tilde{\boldsymbol{\beta}}(\tau)$ 在 $\boldsymbol{\beta}_0(\tau)$ 与 $\boldsymbol{\beta}_{\boldsymbol{\xi}}(\tau)$ 之间. 因为 $\|\boldsymbol{\beta}_{\boldsymbol{\xi}}(\tau) - \boldsymbol{\beta}_0(\tau)\| \le \delta$, 通过取任意小的 δ 有 $(d31)_{\boldsymbol{\xi},\pi}$ 是 $o_p(1)$. 对 $(d32)_{\boldsymbol{\xi},\pi}$ 项, 从条件 (A5.1), 有

$$(d32)_{\boldsymbol{\xi},\pi} = \tau \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} \leqslant 0) - I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 \leqslant 0) \}$$

$$\times [I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} - I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T\}]$$

$$\leqslant \tau \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \{ I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_{\boldsymbol{\xi}} \leqslant 0) - I(\boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0 \leqslant 0) \} [I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T \geqslant \pi_0(\boldsymbol{x}_t) \} + I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau + \eta_T \geqslant \hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) \}]$$

$$= o_p(1).$$

因此,

$$(d3)_{\xi,\pi} = o_p(1). \tag{A.22}$$

根据文献 [16] 有关删失分位数中参数估计的研究, 有 $T^{1/2}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)=O_p(1)$. 通过替换 $\boldsymbol{\xi}=T^{1/2}(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}_0)$ 和条件 (A5.1), (A.19) 可以写为

$$(d1)_{\boldsymbol{\xi}} = -\lambda \sqrt{T} \boldsymbol{J}_0(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) + o_p(1). \tag{A.23}$$

通过结合 (A.13)–(A.15)、(A.20)、(A.22) 和 (A.23), 可得 (A.12). 这就完成了引理 A.2 的证明. □ **定理 2.1 的证明** 通过引理 A.2, 有

$$\begin{split} \boldsymbol{S}(\lambda, \tau, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\pi}}) - \lambda \boldsymbol{S}(1, \tau, \hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\pi}}) &= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \psi_{\tau}(y_t - \boldsymbol{x}_t^{\mathrm{T}} \hat{\boldsymbol{\beta}}) I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} \\ &- \lambda \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{x}_t \psi_{\tau}(y_t - \boldsymbol{x}_t^{\mathrm{T}} \hat{\boldsymbol{\beta}}) I\{\hat{\boldsymbol{\pi}}(\boldsymbol{x}_t) > 1 - \tau + \eta_T\} \end{split}$$

$$= \frac{1}{\sqrt{T}} \sum_{t=1}^{[T\lambda]} \boldsymbol{x}_t \psi_{\tau}(y_t - \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0) I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau\}$$

$$- \lambda \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \boldsymbol{x}_t \psi_{\tau}(y_t - \boldsymbol{x}_t^{\mathrm{T}} \boldsymbol{\beta}_0) I\{\pi_0(\boldsymbol{x}_t) > 1 - \tau\} + o_p(1)$$

$$\Rightarrow \boldsymbol{V}_0^{1/2} \{W_p(\lambda) - \lambda W_p(1)\},$$

其中 $W_p(\lambda)$ 是 p 维的独立 Wiener 过程, 最后的收敛性是根据条件 (A7), 并且通过应用函数中心极限 定理得到. 由大数定律显然可得 $\frac{1}{2}\hat{V} \stackrel{p}{\to} V_0$. 因此有

$$SQ_{SIS}(\tau) \Rightarrow \sup_{\lambda \in [0,1]} ||B_p(\lambda)||_{\infty}.$$

这就完成了定理 2.1 的证明.

定理 4.1 的证明 根据引理 A.1、A.2 和强混合序列的泛函中心极限定理 [33] 知, 当 $T \to \infty$ 时,

$$S(\lambda, \tau, \hat{\boldsymbol{\beta}}, \hat{\pi}) - \lambda S(1, \tau, \hat{\boldsymbol{\beta}}, \hat{\pi}) \Rightarrow \tilde{V}_0^{1/2}[W_p(\lambda) - \lambda W_p(1)].$$

根据 $\frac{1}{T}\tilde{V} \stackrel{p}{\to} \tilde{V}_0$, 我们可以得到 $\mathrm{SQ}^*_{\mathrm{SIS}}(\tau)$ 和 $\mathrm{SQ}^*_{\mathrm{Powell}}(\tau)$ 的渐近性质. 这就完成了定理 4.1 的证明. \square

附录 B 额外的数值模拟结果

表 B1 样本量 T=1,000 时不同检验过程在模型 1 至 4 下的第一类错误的比较. 名义水平设置为 0.05

au	比例	方法	模型 1	模型 2	模型 3	模型 4
0.25	30%	SIS	0.048	0.045	0.048	0.044
		Powell	0.045	0.041	0.049	0.047
		Naive	0.028	0.037	0.045	0.036
	50%	SIS	0.048	0.046	0.053	0.059
		Powell	0.052	0.058	0.053	0.057
		Naive	0.034	0.029	0.036	0.035
0.50	30%	SIS	0.048	0.052	0.062	0.058
		Powell	0.042	0.048	0.057	0.053
		Naive	0.058	0.048	0.044	0.058
	50%	SIS	0.037	0.043	0.063	0.059
		Powell	0.054	0.046	0.065	0.065
		Naive	0.053	0.054	0.069	0.057
0.75	30%	SIS	0.051	0.053	0.054	0.050
		Powell	0.055	0.052	0.044	0.048
		Naive	0.223	0.195	0.216	0.197
	50%	SIS	0.066	0.061	0.052	0.048
		Powell	0.064	0.052	0.052	0.044
		Naive	0.233	0.221	0.217	0.212

表 B2 样本量 T=1,500 时不同检验过程在模型 $1 \, \Xi \, 4$ 下的第一类错误的比较. 名义水平设置为 0.05

au	比例	方法	模型 1	模型 2	模型 3	模型 4
0.25	30%	SIS	0.052	0.049	0.058	0.055
		Powell	0.053	0.053	0.050	0.055
		Naive	0.037	0.041	0.034	0.037
	50%	SIS	0.036	0.038	0.043	0.053
		Powell	0.039	0.040	0.044	0.053
		Naive	0.031	0.035	0.027	0.024
0.50	30%	SIS	0.047	0.051	0.058	0.059
		Powell	0.044	0.049	0.065	0.062
		Naive	0.046	0.062	0.065	0.058
	50%	SIS	0.047	0.041	0.050	0.047
		Powell	0.041	0.050	0.050	0.051
		Naive	0.053	0.059	0.068	0.064
0.75	30%	SIS	0.050	0.052	0.042	0.044
		Powell	0.043	0.049	0.051	0.052
		Naive	0.238	0.236	0.186	0.193
	50%	SIS	0.053	0.057	0.040	0.042
		Powell	0.052	0.051	0.046	0.049
		Naive	0.250	0.234	0.216	0.201

Testing for change points in censored quantile regression models

Liwen Zhang, Zhongxin Ni, Yong He & Zhoufan Zhu

Abstract We develop a new method to test structural changes in quantile regression in the presence of fixed censoring. The testing procedure is based on the observations in an informative subset and is conducted with the subgradient of the quantile objective function, which only requires estimating the model under the null hypothesis. The proposed method is easy to understand and convenient to compute. We establish the limiting distribution of the test statistic under the null hypothesis, and show that its asymptotic critical values can be obtained via simulation methods. Simulation studies show that the proposed method has competitive performance but higher computational efficiency compared with the Powell-based method in finite samples. In addition, an empirical application to the American earnings gap is used to illustrate the performance of the proposed method.

Keywords change points, fixed censoring, quantile regression, informative subset

MSC(2010) 62N01, 62J99 doi: 10.1360/N012017-00165