

方差变点的加权累积和估计的极限性质

房炜杰, 金百锁

(中国科学技术大学管理学院统计与金融系, 安徽合肥 230026)

摘要: 研究了方差变点的加权累积和估计的一些极限性质. 证明了这个估计量的相合性和收敛速度, 并由此推导出了变点估计量在局部对立假设下的渐近分布的形式, 它是一个有偏移的双边布朗运动. 通过渐近分布的数值解可以构造出估计量的渐近置信区间. 计算机模拟和实际数据分析表明, 这个估计量在应用中有很好的表现.

关键词: 方差变点; 相合性; 收敛速度; 渐近分布

中图分类号: O212.7 **文献标识码:** A doi: 10.3969/j.issn.0253-2778.2020.04.001

2010 Mathematics Subject Classification: 62G20

引用格式: 房炜杰, 金百锁. 方差变点的加权累积和估计的极限性质[J]. 中国科学技术大学学报, 2020, 50(4): 389-395.

FANG Weijie, JIN Baisuo. Limit properties of weighted cumulative sum estimator of change-point in variance[J]. Journal of University of Science and Technology of China, 2020, 50(4): 389-395.

Limit properties of weighted cumulative sum estimator of change-point in variance

FANG Weijie, JIN Baisuo

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

Abstract: Some limit properties of the weighted cumulative sum (WCS) estimator for the change-point in variance was studied. The consistency and convergence rate of the estimator were proved. In addition, the asymptotic distribution of the change-point estimator under local alternative hypothesis, which was a two-side Brownian motion with a drift, was developed. By the numerical solution of asymmetric distribution, an asymptotic confidence interval was constructed. Simulation results and a real data example show that the proposed estimator has a good performance in the application.

Key words: change-point in variance; consistency; convergence rate; asymptotic distribution

0 引言

方差变点估计是数理统计、应用数学和计算机科学中非常重要的一类问题, 并广泛应用于金融学、经济学、环境学和气象科学中. Page^[1]于1954

年提出了累积和(cumulative sum, CUSUM)检验方法, 在此基础上衍生出了很多变点的检验方法. 1977年Hsu^[2]将这个方法的改进并应用于单方差变点检验. 2018年Xu等^[3]提出了方差变点的加权累积和检验方法, 并同时根据检验统计量给出了一个

收稿日期: 2019-10-25; 修回日期: 2020-02-20

基金项目: 国家自然科学基金面上项目(11571337, 71873128), 国家自然科学基金重点项目(71631006)资助.

作者简介: 房炜杰, 男, 1995年生, 硕士. 研究方向: 方差变点. E-mail: fwj007@mail.ustc.edu.cn

通讯作者: 金百锁, 博士/副教授. E-mail: jbs@ustc.edu.cn

方差变点的估计量,但没有进一步研究该估计量的性质.本文对这个估计量性质进行了研究.

Bai^[4]给出了线性过程变点的最小二乘估计的相合性、收敛速度和渐近分布的证明. Jin 等^[5]对单指标模型的变点估计也给出了这些性质的证明.这些文献中的证明方法对本文中的证明很有启发.

本文余下部分的结构如下:第 1 节主要介绍了模型和一些假设;估计的主要性质和证明分别在第 2 节和第 3 节给出;第 4 节是模拟结果;第 5 节是实例分析;最后第 6 节是结论.

1 模型和假设

考虑 $X_i, i=1, 2, \dots, n$, 是一列均值为 0, 方差为 $\sigma_i^2, i=1, 2, \dots, n$, 的独立的随机变量序列. 存在一个 $\tau^*, k_n^* = [n\tau^*]$ 为序列的变点, 使得

$$\sigma_i^2 = \begin{cases} \sigma_1^2, & 1 \leq i \leq k_n^*; \\ \sigma_n^2, & k_n^* < i \leq n \end{cases} \quad (1)$$

而且 $\sigma_1^2 \neq \sigma_n^2$, 在 k_n^* 前后分别独立同分布于不同总体.

检验统计量的构造基于 Jensen 不等式^[6]:

$$f(\alpha t_1 + \beta t_2) \leq \alpha f(t_1) + \beta f(t_2),$$

其中, f 是一个凸函数, $\alpha > 0, \beta > 0, \alpha + \beta = 1$. 对式 (1) 中的 σ_1^2 和 σ_n^2 应用上述不等式, 假设 $f(t) = t^\lambda$ 满足 $t > 0$ 且 $\lambda > 1$, 则 f 在 $t > 0$ 时是一个凸函数. 由 Jensen 不等式, 可以得到

$$\left(\frac{k}{n}\sigma_1^2 + \frac{n-k}{n}\sigma_n^2\right)^\lambda \leq \frac{k}{n}(\sigma_1^2)^\lambda + \frac{n-k}{n}(\sigma_n^2)^\lambda \quad (2)$$

记 $\sigma_0^2 = \frac{k}{n}\sigma_1^2 + \frac{n-k}{n}\sigma_n^2$. 将 $\sigma_0^2, \sigma_1^2, \sigma_n^2$ 的理论值用估计量代替, 我们就得到了估计变点的统计量. 定义方差变点的加权累积和 (weighted cumulative sum, WCS) 如下, 并记为 $v_{\lambda, k}$, 有

$$v_{\lambda, k} = k(\hat{\sigma}_{1, k}^2)^\lambda + (n-k)(\hat{\sigma}_{n, k}^2)^\lambda - n(\hat{\sigma}_0^2)^\lambda \quad (3)$$

式中,

$$\hat{\sigma}_{1, k}^2 = \frac{\sum_{i=1}^k X_i^2}{k}, \hat{\sigma}_{n, k}^2 = \frac{\sum_{i=k+1}^n X_i^2}{n-k}, \hat{\sigma}_0^2 = \frac{\sum_{i=1}^n X_i^2}{n} \quad (4)$$

根据不同 λ 的凹凸性, 变点有如下的定义:

① 当 $\lambda > 1$ 或 $\lambda < 0$ 时, 变点 \hat{k}_n 定义为

$$\hat{k}_n = \arg \max_k \{v_{\lambda, k}\};$$

② 当 $0 < \lambda < 1$ 时, 变点 \hat{k}_n 定义为

$$\hat{k}_n = \arg \min_k \{v_{\lambda, k}\}.$$

很自然的, τ^* 可以有如下估计:

$$\hat{\tau}_n = \frac{\hat{k}_n}{n} \quad (5)$$

设 $\rho_n = \sigma_n^2 - \sigma_1^2$ 是方差的变化量. 不失一般性, 假设 $\rho_n > 0$, 不然反转整个数列可以使变点前的方差更大, 也可以得到 $\rho_n > 0$. 节 2 的定理需要如下几条假设:

A1 $\lambda \neq 0, 1$.

A2 变点的理论值 k_n^* 满足 $n\delta < k_n^* < n(1-\delta)$, 对某个 $0 < \delta < 1/2$.

A3 四阶矩满足 $0 < EX_i^4 < \infty$, 对 $1 \leq i \leq n$.

A4 $\rho_n \rightarrow 0, n\rho_n^2 \rightarrow \infty$, 当 $n \rightarrow \infty$. 而且 $\sigma_n^2 \rightarrow \sigma^2, \sigma^2 > 0$.

A5 当 $i > k_n^*$ 时, $(EX_i^4 - \sigma_i^4)^{\frac{1}{2}} = \mu_1$ 为常数; 当 $i \leq k_n^*$ 时, $(EX_i^4 - \sigma_i^4)^{\frac{1}{2}} = \mu_2$ 为常数.

2 主要结论

下面的定理主要说明了统计量的收敛速度, 由此可以推导出统计量的相合性.

定理 2.1 基于模型 (1) 和假设 A1~A4 可以得到

$$|\hat{\tau}_n - \tau^*| = O_p(n^{-1}\rho_n^{-2}).$$

可以直接由定理 2.1 推导出 $\hat{\tau}_n$ 是 τ^* 的一个相合估计量. 以定理 2.1 为基础, 可以进一步地求出估计量的渐近分布, 即定理 2.2:

定理 2.2 基于假设 A1~A5 可以得到如下的渐近分布:

$$\omega_n(\hat{\tau}_n - \tau^*) \xrightarrow{d} \arg \max_s G(s) \quad (6)$$

式中,

$$G(s) = \begin{cases} -\frac{1}{2}s + \frac{\mu_1}{\mu_2} W_1(s), & s \geq 0; \\ \frac{1}{2}s - W_2(-s), & s < 0; \end{cases}$$

其中, $\omega_n = n\rho_n^2/\mu_2^2$, $W_1(s)$ 和 $W_2(s)$ 是定义在 $(0, \infty)$ 上的相互独立的标准布朗过程.

因为式 (6) 中不含参数 λ , 由定理 2.2 可以得出对不同的 λ , 它们所对应的估计量的渐近分布都是相同的. 渐近分布的具体形式可以从文献 [7] 的附录 B 中得到. 假设 $\phi = \frac{\mu_1^2}{\mu_2^2}$, 那么当 $x > 0$ 时渐近分布:

$$F(x) = 1 + \phi^{-1/2} (2\pi)^{-1/2} x^{1/2} \exp(-8^{-1}\phi^{-1}x) +$$

$$c \exp(ax) \Phi(-bx^{1/2}) + (-d + 2 - 2^{-1} \phi^{-1} x) \Phi(-2^{-1} \phi^{-1/2} x^{1/2}), x > 0,$$

其中, $a = \frac{\phi+1}{2}, b = \frac{2\phi+1}{2\sqrt{\phi}}, c = \frac{2\phi+1}{(\phi+1)\phi},$

$$d = \frac{(2\phi+1)^2}{(\phi+1)\phi}.$$

当 $x < 0$ 时

$$F(x) = -(\pi)^{-1/2} |x|^{1/2} \exp(-8^{-1} |x|) - c \exp(a|x|) \Phi(-b|x|^{1/2}) + (d - 2 + 2^{-1} |x|) \Phi(-2^{-1} |x|^{1/2}), x < 0.$$

其中, $a = 2^{-1} \frac{1}{\phi} (1 + \frac{1}{\phi}), b = \frac{1}{2} + \frac{1}{\phi}, c = \frac{\phi(\phi+2)}{\phi+1}, d = \frac{(\phi+2)^2}{\phi+1}.$

假设 c_1 表示渐近分布的 $(\alpha/2)$ 分位点, c_2 表示渐近分布的 $(1-\alpha/2)$ 分位点. 定义

$$\widehat{L} = \widehat{\rho}_n / \widehat{\mu}_2^2 = \frac{(n - \widehat{k}_n - 1)(\widehat{\sigma}_{\widehat{k}_n, n}^2 - \widehat{\sigma}_{1, \widehat{k}_n}^2)}{\sum_{i=\widehat{k}_n+1}^n (X_i^2 - \widehat{\sigma}_{\widehat{k}_n, n}^2)^2}.$$

则一个 $(1-\alpha)$ 置信区间为

$$[\widehat{k}_n - [c_2/\widehat{L}] - 1, \widehat{k}_n - [c_1/\widehat{L}] + 1].$$

这个渐近分布的形式非常复杂, 在 $\mu_1 = \mu_2$ 情况下的数值解可以在文献[8]的表 1 中查到. 比如, 当 $\alpha = 0.1$ 时, $c_1 = -7, c_2 = 7$; 当 $\alpha = 0.05$ 时, $c_1 = -11, c_2 = 11$.

3 证明

3.1 定理 2.1 的证明

对任意 $1 \leq k \leq n$, 对 $v_{\lambda, k}$ 和 v_{λ, k_n^*} 求差, 有

$$v_{\lambda, k} - v_{\lambda, k_n^*} = k(\widehat{\sigma}_{1, k}^2)^\lambda + (n - k)(\widehat{\sigma}_{n, k}^2)^\lambda - k_n^*(\widehat{\sigma}_{1, k_n^*}^2)^\lambda - (n - k_n^*)(\widehat{\sigma}_{n, k_n^*}^2)^\lambda.$$

当 $k > k_n^*$ 时, 对上式中的 $(\widehat{\sigma}_{1, k}^2)^\lambda, (\widehat{\sigma}_{n, k}^2)^\lambda, (\widehat{\sigma}_{1, k_n^*}^2)^\lambda, (\widehat{\sigma}_{n, k_n^*}^2)^\lambda$ 分别使用带有 Lagrange 余项的 Taylor 公式展开到三阶. 三阶 Taylor 展开分别得到以下四项:

$$k(\widehat{\sigma}_{1, k}^2)^\lambda = k \left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^\lambda + k\lambda \left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^{\lambda-1} \left(\widehat{\sigma}_{1, k}^2 - \frac{k_n^*}{k} \sigma_1^2 - \frac{k - k_n^*}{k} \sigma_n^2 \right) + k\lambda(\lambda - 1) \left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^{\lambda-2} \cdot \left(\widehat{\sigma}_{1, k}^2 - \frac{k_n^*}{k} \sigma_1^2 - \frac{k - k_n^*}{k} \sigma_n^2 \right)^2 +$$

$$k\lambda(\lambda - 1)(\lambda - 2) \cdot (\widehat{\sigma}^2)^{\lambda-3} \left(\widehat{\sigma}_{1, k}^2 - \frac{k_n^*}{k} \sigma_1^2 - \frac{k - k_n^*}{k} \sigma_n^2 \right)^3 \stackrel{\Delta}{=} I_1 + II_1 + III_1 + IV_1,$$

$$(n - k)(\widehat{\sigma}_{n, k}^2)^\lambda = (n - k)(\sigma_n^2)^\lambda + (n - k)\lambda(\sigma_n^2)^{\lambda-1}(\widehat{\sigma}_{n, k}^2 - \sigma_n^2) + (n - k)\lambda(\lambda - 1)(\sigma_n^2)^{\lambda-2}(\widehat{\sigma}_{n, k}^2 - \sigma_n^2)^2 + (n - k)\lambda(\lambda - 1)(\lambda - 2)(\widehat{\sigma}^2)^{\lambda-3}(\widehat{\sigma}_{n, k}^2 - \sigma_n^2)^3 \stackrel{\Delta}{=} I_2 + II_2 + III_2 + IV_2,$$

$$k_n^*(\widehat{\sigma}_{1, k_n^*}^2)^\lambda = k_n^*(\sigma_1^2)^\lambda + k_n^*\lambda(\sigma_1^2)^{\lambda-1}(\widehat{\sigma}_{n, k_n^*}^2 - \sigma_1^2) + k_n^*\lambda(\lambda - 1)(\sigma_1^2)^{\lambda-2}(\widehat{\sigma}_{1, k_n^*}^2 - \sigma_1^2)^2 + k_n^*\lambda(\lambda - 1)(\lambda - 2)(\widehat{\sigma}^2)^{\lambda-3}(\widehat{\sigma}_{1, k_n^*}^2 - \sigma_1^2)^3 \stackrel{\Delta}{=} I_3 + II_3 + III_3 + IV_3,$$

$$(n - k_n^*)(\widehat{\sigma}_{n, k_n^*}^2)^\lambda = (n - k_n^*)(\sigma_n^2)^\lambda + (n - k_n^*)\lambda(\sigma_n^2)^{\lambda-1}(\widehat{\sigma}_{n, k_n^*}^2 - \sigma_n^2) + (n - k_n^*)\lambda(\lambda - 1)(\sigma_n^2)^{\lambda-2}(\widehat{\sigma}_{n, k_n^*}^2 - \sigma_n^2)^2 + (n - k_n^*)\lambda(\lambda - 1)(\lambda - 2)(\widehat{\sigma}^2)^{\lambda-3}(\widehat{\sigma}_{n, k_n^*}^2 - \sigma_n^2)^3 \stackrel{\Delta}{=} I_4 + II_4 + III_4 + IV_4.$$

将这些项重新组合并假设:

$$I = I_1 + I_2 - I_3 - I_4, II = II_1 + II_2 - II_3 - II_4, III = III_1 + III_2 - III_3 - III_4, IV = IV_1 + IV_2 - IV_3 - IV_4.$$

因此

$$v_{\lambda, k} - v_{\lambda, k_n^*} = I + II + III + IV.$$

之后我们分别来看这四部分.

$$I = k \left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^\lambda - k_n^*(\sigma_1^2)^\lambda - (k - k_n^*)(\sigma_n^2)^\lambda = [k(\sigma_1^2 + \frac{k - k_n^*}{k}(\sigma_n^2 - \sigma_1^2))^\lambda - k(\sigma_1^2)^\lambda] + [(k - k_n^*)(\sigma_1^2)^\lambda - (k - k_n^*)(\sigma_n^2)^\lambda] = (k - k_n^*)(\sigma_n^2 - \sigma_1^2) \cdot \left[\frac{(\sigma_1^2 + \frac{k - k_n^*}{k}(\sigma_n^2 - \sigma_1^2))^\lambda - (\sigma_1^2)^\lambda}{\frac{k - k_n^*}{k}(\sigma_n^2 - \sigma_1^2)} - \frac{(\sigma_n^2)^\lambda - (\sigma_1^2)^\lambda}{\sigma_n^2 - \sigma_1^2} \right].$$

假设 $f(x) = \frac{x^\lambda - a^\lambda}{x - a}$, 其中 $a = \sigma_1^2$. 那么

$$f'(x) = \frac{(\lambda - 1)x^{\lambda-1} - \lambda ax^{\lambda-1} + a^\lambda}{(x - a)^2}.$$

当 $x \neq a$ 时, 由 Young 不等式, 当 $\lambda > 1$ 或 $\lambda < 0$ 时,

$$(\lambda - 1)x^\lambda - \lambda ax^{\lambda-1} + a^\lambda > 0;$$

当 $0 < \lambda < 1$ 时,

$$(\lambda - 1)x^\lambda - \lambda ax^{\lambda-1} + a^\lambda < 0.$$

因此有

$$f'(x) \begin{cases} > 0, \lambda > 1, \lambda < 0; \\ < 0, 0 < \lambda < 1. \end{cases}$$

当 $x \rightarrow a$ 时同样有

$$\lim_{x \rightarrow \infty} f'(x) = \frac{1}{2} \lambda (\lambda - 1) a^{\lambda-2} \begin{cases} > 0, \lambda > 1, \lambda < 0; \\ < 0, 0 < \lambda < 1. \end{cases}$$

由 Lagrange 中值定理, 有

$$I = (k - k_n^*)(\sigma_n^2 - \sigma_1^2) \left[-\frac{k_n^*}{k}(\sigma_n^2 - \sigma_1^2) \right] f'(\tilde{\sigma}^2) = - (k - k_n^*)(\sigma_n^2 - \sigma_1^2)^2 \frac{\tau^*}{\tau} f'(\tilde{\sigma}^2).$$

$\tilde{\sigma}^2$ 的值位于 $\sigma_1^2 + \frac{\tau - \tau^*}{\tau}(\sigma_n^2 - \sigma_1^2)$ 和 σ_n^2 之间. 由 $\tilde{\sigma}^2$

的取值范围, 我们知道 $\tilde{\sigma}^2 \neq \sigma_1^2$.

下面计算 II. 假设

$$A = \sum_{i=1}^{k_n^*} x_i^2 - k_n^* \sigma_1^2,$$

$$B = \sum_{i=k_n^*+1}^k x_i^2 - (k - k_n^*) \sigma_n^2,$$

$$C = \sum_{i=k+1}^n x_i^2 - (n - k) \sigma_n^2.$$

注意到由 Lagrange 中值定理有

$$(\sigma_n^2)^\lambda - (\sigma_1^2)^\lambda = \lambda (\tilde{\sigma}^2)^{\lambda-1} (\sigma_n^2 - \sigma_1^2) = O_p(\rho_n),$$

其中 $\sigma_1^2 \leq \tilde{\sigma}^2 \leq \sigma_n^2$. 由泛函中心极限定理, 有

$$\begin{aligned} II &= \lambda \left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^{\lambda-1} (A + B) + \\ &\lambda (\sigma_n^2)^{\lambda-1} C - \lambda (\sigma_1^2)^{\lambda-1} A - \lambda (\sigma_n^2)^{\lambda-1} (B + C) = \\ &\lambda \left[\left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^{\lambda-1} - (\sigma_1^2)^{\lambda-1} \right] A + \\ &\lambda \left[\left(\frac{k_n^*}{k} \sigma_1^2 + \frac{k - k_n^*}{k} \sigma_n^2 \right)^{\lambda-1} - (\sigma_n^2)^{\lambda-1} \right] B = \\ &O_p((\tau - \tau^*) \rho_n \sqrt{n}) + O_p(\rho_n \sqrt{(\tau - \tau^*) n}). \end{aligned}$$

下面计算 III, 有

$$III_1 = n O_p \left(\left(\frac{1}{\sqrt{n}} \right)^3 \right) = O_p(1),$$

同理,

$$III_2 = O_p(1), III_3 = O_p(1), III_4 = O_p(1),$$

所以有

$$III = III_1 + III_2 - III_3 - III_4 = O_p(1).$$

对于 IV, 有

$$IV_1 = n O_p(1) O_p \left(\left(\frac{1}{\sqrt{n}} \right)^3 \right) = O_p \left(\frac{1}{\sqrt{n}} \right),$$

同理,

$$IV_2 = O_p \left(\frac{1}{\sqrt{n}} \right), IV_3 = O_p \left(\frac{1}{\sqrt{n}} \right),$$

$$IV_4 = O_p \left(\frac{1}{\sqrt{n}} \right),$$

所以有

$$IV = IV_1 + IV_2 - IV_3 - IV_4 = O_p \left(\frac{1}{\sqrt{n}} \right).$$

因此将它们组合起来就可以得到

$$\begin{aligned} v_{\lambda, k} - v_{\lambda, k_n^*} &= -n(\tau - \tau^*) \rho_n^2 \frac{\tau^*}{\tau} f'(\tilde{\sigma}^2) + \\ &O_p((\tau - \tau^*) \rho_n \sqrt{n}) + \\ &O_p(\rho_n \sqrt{(\tau - \tau^*) n}) + O_p(1). \end{aligned}$$

当 $\hat{k}_n > k^*$ 时, 用 \hat{k}_n 代替上式中的 k 可以得到

$$\begin{aligned} v_{\lambda, \hat{k}_n} - v_{\lambda, k_n^*} &= -n(\hat{\tau}_n - \tau^*) \rho_n^2 \frac{\tau^*}{\tau_n} f'(\tilde{\sigma}^2) + \\ &O_p((\hat{\tau}_n - \tau^*) \rho_n \sqrt{n}) + \\ &O_p(\rho_n \sqrt{(\hat{\tau}_n - \tau^*) n}) + O_p(1). \end{aligned}$$

当 $0 < \lambda < 1$ 时, $f'(x) < 0$, 所以 $-n(\hat{\tau}_n - \tau^*) \cdot \rho_n^2 \frac{\tau^*}{\tau_n} f'(\tilde{\sigma}^2)$ 项为正. 由 \hat{k}_n 定义中 v_{λ, \hat{k}_n} 的最小性, 上式左边为负可以得到

$$\begin{aligned} n(\hat{\tau}_n - \tau^*) \rho_n^2 &\leq O_p((\hat{\tau}_n - \tau^*) \rho_n \sqrt{n}) + \\ &O_p(\rho_n \sqrt{(\hat{\tau}_n - \tau^*) n}) + O_p(1). \end{aligned}$$

当 $\lambda < 0$ 或 $\lambda > 1$ 时, $f'(x) > 0$. 同样使用 \hat{k}_n 代替式中的 k , 再由 \hat{k}_n 定义中 v_{λ, \hat{k}_n} 的最大性可以得到相同结论.

当 $k < k_n^*$ 时, 由同样的方法也可以得到

$$\begin{aligned} n(\tau^* - \hat{\tau}_n) \rho_n^2 &\leq O_p((\tau^* - \hat{\tau}_n) \rho_n \sqrt{n}) + \\ &O_p(\rho_n \sqrt{(\tau^* - \hat{\tau}_n) n}) + O_p(1). \end{aligned}$$

因此可得

$$\begin{aligned} n|\hat{\tau}_n - \tau^*| \rho_n^2 &\leq O_p(|\hat{\tau}_n - \tau^*| \rho_n \sqrt{n}) + \\ &O_p(\rho_n \sqrt{|\hat{\tau}_n - \tau^*| n}) + O_p(1). \end{aligned}$$

为使证明简洁清晰, 将后面的证明总结为如下引理:

引理 3.1 假设 X_n 是一组数列, Z_n, A_n, B_n, C_n 是满足 $Z_n^2 \leq A_n + B_n + C_n$ 的随机变量序列, 其

中 $A_n = O_p\left(\frac{Z_n^2}{X_n}\right)$, $B_n = O_p(Z_n)$, $C_n = O_p(1)$, $X_n \rightarrow \infty$. 得到 $Z_n = O_p(1)$.

证明 由 O_p 的定义对任意 $\epsilon > 0$, 存在 C_1, N_1 , 当 $n > N_1$ 时,

$$P(|A_n| > \frac{C_1 Z_n^2}{X_n}) < \frac{\epsilon}{3}.$$

对任意 $\epsilon > 0$, 存在 C_2, N_2 , 当 $n > N_2$ 时,

$$P(|B_n| > C_2 Z_n) < \frac{\epsilon}{3}.$$

对任意 $\epsilon > 0$, 存在 C_3, N_3 , 当 $n > N_3$ 时,

$$P(|C_n| > C_3) < \frac{\epsilon}{3}.$$

又因为 $X_n \rightarrow \infty$, 存在 N_4 , 当 $n > N_4$, $X_n > 2C_1$. 当 $n > \max\{N_1, N_2, N_3, N_4\}$ 时有

$$\begin{aligned} &P(|Z_n| > C_2 + \sqrt{C_2^2 + 2C_3}) = \\ &P(|Z_n| - C_2 > \sqrt{C_2^2 + 2C_3}) \leq \\ &P(|Z_n| - C_2 > \sqrt{C_2^2 + 2C_3}) + \\ &P(|Z_n| - C_2 < -\sqrt{C_2^2 + 2C_3}) = \\ &P((|Z_n| - C_2)^2 > C_2^2 + 2C_3) = \\ &P\left(\frac{1}{2} |Z_n|^2 - C_2 |Z_n| - C_3 > 0\right) \leq \end{aligned}$$

$$P(|Z_n|^2 - \frac{C_1}{X_n} |Z_n|^2 - C_2 |Z_n| - C_3 > 0) \leq$$

$$\begin{aligned} &P\left(|A_n| - \frac{C_1}{X_n} |Z_n|^2 + |B_n| - \right. \\ &C_2 |Z_n| + |C_n| - C_3 > 0) \leq \\ &P\left(|A_n| - \frac{C_1}{X_n} |Z_n|^2 > 0\right) + \\ &P(|B_n| - C_2 |Z_n| > 0) + \\ &P(|C_n| - C_3 > 0) \leq \epsilon. \end{aligned}$$

所以有 $Z_n = O_p(1)$. 证毕.

设 $X_n = \sqrt{n} \rho_n$, $Z_n = \sqrt{n} |\hat{\tau}_n - \tau^*| \rho_n$, 由引理

3.1, $\sqrt{n} |\hat{\tau}_n - \tau^*| \rho_n = O_p(1)$, 此即

$$|\hat{\tau}_n - \tau^*| = O_p(n^{-1} \rho_n^{-2}) \quad (7)$$

3.2 定理 2.2 的证明

当 $\lambda < 0$ 或 $\lambda > 1$ 时,

$$\hat{k}_n = \arg \max_k v_{\lambda, k} = \arg \max_k (v_{\lambda, k} - v_{\lambda, k_n^*}).$$

我们考察 $v_{\lambda, k} - v_{\lambda, k_n^*}$ 当 k 位于 k_n^* 邻域时的表现, 即 $k = k_n^* + O_p(\rho_n^{-2})$. 类似文献[9]中的结论, 在 $\hat{k}_n = k_n^* + \lfloor s \rho_n^{-2} \rfloor$ 的条件下会得到 $v_{\lambda, k} - v_{\lambda, k_n^*}$ 的一个弱收敛结论, 其中 $|s| \leq M$, 对任意给定的 M

> 0 . 之后应用连续映射定理^[10], 即可得到方差变点估计量的渐近分布. 定义

$$V_n(s) = v_{\lambda, k_n^* + \lfloor s \rho_n^{-2} \rfloor} - v_{\lambda, k_n^*} \quad (8)$$

当 $k > k_n^*$ 时, $s > 0$, 由定理 2.1 的证明, 有

$$V_n(s) = v_{\lambda, k_n^* + \lfloor s \rho_n^{-2} \rfloor} - v_{\lambda, k_n^*} = I + II + III + IV.$$

其中, $I = -s \frac{\tau^*}{\tau} f'(\tilde{\sigma}^2)$. 又因为 $\rho_n \rightarrow 0$, 由假设 A4

可得, $\frac{\sigma_1^2}{\sigma_n^2} \rightarrow 1$. 根据 $f'(x)$ 的表达式和 L'Hôpital 法

则: 当 $n \rightarrow \infty$ 时, $f'(\tilde{\sigma}^2) \rightarrow \frac{1}{2} \lambda (\lambda - 1) (\sigma^2)^{\lambda - 2}$.

因此

$$I \rightarrow -\frac{1}{2} \lambda (\lambda - 1) (\sigma^2)^{\lambda - 2} s.$$

$$A = \sum_{i=1}^{k_n^*} x_i^2 - k_n^* \sigma_1^2;$$

$$B = \sum_{i=k_n^*+1}^{\hat{k}_n} x_i^2 - (\hat{k}_n - k_n^*) \sigma_n^2;$$

$$C = \sum_{i=\hat{k}_n+1}^n X_i^2 - (n - \hat{k}_n) \sigma_n^2.$$

由泛函中心极限定理(Donsker 定理)可知

$$II = \lambda \left[\left(\frac{k_n^*}{\hat{k}_n} \sigma_1^2 + \frac{\lfloor s \rho_n^{-2} \rfloor}{\hat{k}_n} \sigma_n^2 \right)^{\lambda - 1} - (\sigma_1^2)^{\lambda - 1} \right] A +$$

$$\lambda \left[\left(\frac{k_n^*}{\hat{k}_n} \sigma_1^2 + \frac{\lfloor s \rho_n^{-2} \rfloor}{\hat{k}_n} \sigma_n^2 \right)^{\lambda - 1} - (\sigma_n^2)^{\lambda - 1} \right] B =$$

$$O_p((\tau - \tau^*) \rho_n \sqrt{n}) +$$

$$\left(\frac{k_n^*}{\hat{k}_n} \frac{\sigma_1^2}{\sigma_n^2} + \frac{\lfloor s \rho_n^{-2} \rfloor}{\hat{k}_n} \right)^{\lambda - 1} - 1$$

$$\rho_n \lambda \frac{\sigma_1^2}{\sigma_n^2} - 1 B =$$

$$o_p(1) + \rho_n \lambda \frac{\left(\frac{k_n^*}{\hat{k}_n} \frac{\sigma_1^2}{\sigma_n^2} + \frac{\lfloor s \rho_n^{-2} \rfloor}{\hat{k}_n} \right)^{\lambda - 1} - 1}{\frac{\sigma_1^2}{\sigma_n^2} - 1}.$$

$$\sum_{i=k_n^*+1}^{k_n^* + \lfloor s \rho_n^{-2} \rfloor} (x_i^2 - \sigma_n^2) \xrightarrow{d} \lambda (\lambda - 1) (\sigma^2)^{\lambda - 2} \mu_1 W_1(s).$$

其中, 求和项 $x_i^2 - \sigma_n^2$ 独立同分布取自随机变量序列 $X_i (i = 1, \dots, n)$ 的理论变点 k^* 之后. 这里 $W_1(s)$ 是定义在 $[0, \infty)$ 上的一个标准布朗运动.

$$III = \frac{1}{\hat{k}_n} \lambda (\lambda - 1) \left(\frac{k_n^*}{\hat{k}_n} \sigma_1^2 + \frac{\lfloor s \rho_n^{-2} \rfloor}{\hat{k}_n} \sigma_n^2 \right)^{\lambda - 2} A^2 +$$

$$\begin{aligned}
 & o_p(1) + \frac{1}{n - \hat{k}_n} \lambda(\lambda - 1)(\sigma_n^2)^{\lambda-2} C^2 - \\
 & \frac{1}{k_n^*} \lambda(\lambda - 1)(\sigma_1^2)^{\lambda-2} A^2 - \\
 & \frac{1}{n - k_n^*} \lambda(\lambda - 1)(\sigma_n^2)^{\lambda-2} C^2 - o_p(1) = o_p(1), \\
 & IV = O_p\left(\frac{1}{\sqrt{n}}\right) = o_p(1).
 \end{aligned}$$

因此,

$$\begin{aligned}
 & V_n(s) = v_{\lambda, k_n^*} + \lfloor s \rho_n^{-2} \rfloor - v_{\lambda, k_n^*} \Rightarrow \\
 & -\frac{1}{2} \lambda(\lambda - 1)(\sigma^2)^{\lambda-2} s - \lambda(\lambda - 1)(\sigma^2)^{\lambda-2} \mu_1 W_1(s).
 \end{aligned}$$

由连续映射定理可以得到 $s > 0$ 时变点估计量的渐近分布: $n\rho_n^2(\hat{\tau}_n - \tau) \xrightarrow{d} \arg \max V(s)$, 其中,

$$\begin{aligned}
 & V(s) = -\frac{1}{2} \lambda(\lambda - 1)(\sigma^2)^{\lambda-2} s + \\
 & \lambda(\lambda - 1)(\sigma^2)^{\lambda-2} \mu_1 W_1(s), \quad s > 0.
 \end{aligned}$$

当 $k < k_n^*$ 时, 由同样的方法可以得到

$$n\rho_n^2(\hat{\tau}_n - \tau) \xrightarrow{d} \arg \max V(s),$$

其中,

$$\begin{aligned}
 & V(s) = \frac{1}{2} \lambda(\lambda - 1)(\sigma^2)^{\lambda-2} s + \\
 & \lambda(\lambda - 1)(\sigma^2)^{\lambda-2} \mu_2 W_2(-s), \quad s < 0.
 \end{aligned}$$

而与 $W_1(s)$ 对称的, 收敛到 $W_2(-s)$ 的求和项 $x_i^2 - \sigma_1^2$ 独立同分布地取自随机变量序列 $X_i (i = 1, \dots, n)$ 的理论变点 k^* 之前, 所以 $W_2(s)$ 与 $W_1(s)$ 独立. 因此对任意的 s 都有

$$n\rho_n^2(\hat{\tau}_n - \tau) \xrightarrow{d} \arg \max_s V'(s).$$

其中,

$$V'(s) = \begin{cases} -\frac{1}{2}s + \mu_1 W_1(s), & s \geq 0; \\ \frac{1}{2}s + \mu_2 W_2(-s), & s < 0. \end{cases}$$

$W_1(s)$ 和 $W_2(s)$ 是定义在 $(0, \infty)$ 上的独立标准布朗运动过程. 当 $0 < \lambda < 1$ 时, 可以得到同样的结论. 取 $s = \mu_2^2 v$, 可以得到

$$\arg \max_s V'(s) = \mu_2^2 \arg \max_v G(v),$$

其中,

$$G(v) = \begin{cases} -\frac{1}{2}v + \frac{\mu_1}{\mu_2} W_1(v), & v \geq 0; \\ \frac{1}{2}v + W_2(-v), & v < 0. \end{cases}$$

这是一个有偏移的双边布朗运动.

4 计算机模拟

为了验证 WCS 估计在实际应用中也有很好的表现, 将它同 PELT 估计^[11] 和传统 CUSUM 估计, 即 WCS 估计在 $\lambda = 2$ 时的情况, 通过计算机模拟进行对比.

按照模型(1)随机生成一组样本, X_1, X_2, \dots, X_n , 来验证几种估计量的好坏. 记变点位置为 τ^* , 设变点之前的样本点服从正态分布 $N(0, 1)$, 变点之后的样本点服从另一个不同方差的正态分布 $N(0, 1.5^2)$, 这里固定 $n = 1\,000$. 这样每个方差变点估计量对这组样本都会产生一个估计值. 将上面的步骤重复 1 000 次, 每个估计量会产生 1 000 个估计值. 文献[3]的模拟结果表明 WCS 在 $\lambda = 0.1$ 时有更好的表现, 因此我们使用带有参数 $\lambda = 0.1$ 的 WCS 估计. $\tau^* = 0.3, 0.5, 0.8$, 三种估计方法各自的直方图如图 1 所示.

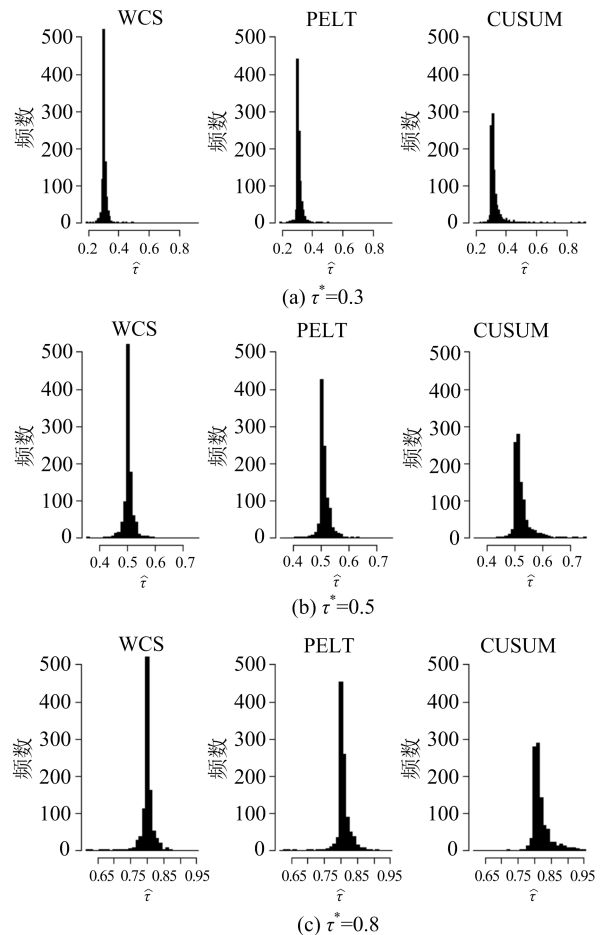


图 1 不同变点理论值下变点估计的直方图

Fig. 1 Histograms of change-point estimations under different theoretical values of change-points

从图 1 中可以看出, WCS 估计相对于其他两种估计方法估计值更加集中于理论值附近. 图中直方图的带宽为 0.01, 所以只有 WCS 估计可以做到有一半以上的估计值 $\hat{\tau}$ 与变点理论值 τ^* 的差的绝对值小于 0.005. 这说明我们的方法在估计方差变点位置时会得到更好的结果.

5 实证分析

我们选择深证 A 股海康威视(002415)2014 年 3 月 3 日到 2016 年 3 月 1 日每日收盘价的对数收益率数据, 研究其方差变点. 这里认为对数收益率的均值为 0. 海康威视每日收盘价的对数收益率如图 2 所示, 整体来看, 在第 200 天之前点的分布更密集一些. 同样使用 $\lambda=0.1$ 的 WCS 估计, 得到的变点估计值为 185, 也就是在 2014 年 11 月 28 日. 对以这个时间分开的两段继续使用 WCS 检验, 没有发现其他变点. 而在几乎相同时间, 海康威视的产品被爆出有安全漏洞, 虽然公司及时出面澄清, 但有关部门仍然进行了长时间的调查. 这导致了事件发生后海康威视股价对数收益率的大幅度波动, 事件前后对数收益率的分布有着不同的方差. 实证说明我们的方法可以有效地估计出真实数据中变点的位置.

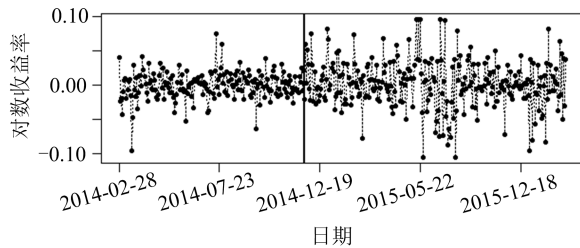


图 2 海康威视每日收盘价的对数收益率

Fig. 2 Daily logarithmic rate of return of Hikvision

6 结论

本文主要证明了 WCS 估计的三个重要极限性质: 相合性、收敛速度和渐近分布. 由渐近分布的具体形式可知, 不同的 λ 有着相同的渐近分布, 所以无法据此得到一个最优的 λ , 这也是后续需要解决的问题. 之后又构造了渐近分布的渐近置信区间. 计算机模拟表明 WCS 估计相对传统方差变点估计有更

好的表现. 实证分析显示, WCS 估计在实际数据分析中也是可行的. 但是渐近置信区间在处理真实数据时存在较大误差, 今后需要研究构造更合理的置信区间. 而且因为对不同的 λ , WCS 估计有相同的渐近分布, 所以无法根据渐近分布判断最佳的 λ , 因此采用其他方法研究最佳的 λ 值得进一步研究.

参考文献(References)

- [1] PAGE E S. Continuous inspection schemes [J]. *Biometrika*, 1954, 41: 100-115.
- [2] HSU D A. Tests for variance shift at an unknown time point [J]. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 1977, 26(3): 279-284.
- [3] XU M, WU Y, JIN B. Detection of a change-point in variance by a weighted sum of powers of variances test [J]. *Journal of Applied Statistics*, 2018, 46(4): 664-679.
- [4] BAI J. Least squares estimation of a shift in linear processes [J]. *Journal of Time Series Analysis*, 1994, 15(5): 453-472.
- [5] JIN B, DONG C, TAN C, et al. Estimator of a change point in single index models [J]. *Science China Mathematics*, 2014, 57(8): 1701-1712.
- [6] JENSEN J L W V. Sur les fonctions convexes et les inégalités entre les valeurs moyennes [J]. *Acta Mathematica*, 1906, 30: 175-193.
- [7] BAI J. Estimation of a change point in multiple regression models [J]. *Review of Economics and Statistics*, 1997, 79(4): 551-563.
- [8] YAO Y C. Approximating the distribution of the maximum likelihood estimate of the change-point in a sequence of independent random variables [J]. *The Annals of Statistics*, 1987, 15(3): 1321-1328.
- [9] PICARD D. Testing and estimating change-points in time series [J]. *Advances in Applied Probability*, 1985, 17(4): 841-867.
- [10] MANN H B, WALD A. On stochastic limit and order relationships [J]. *The Annals of Mathematical Statistics*, 1943, 14(3): 217-226.
- [11] KILLICK R, FEARNHEAD P, ECKLEY I A. Optimal detection of changepoints with a linear computational cost [J]. *Journal of the American Statistical Association*, 2012, 107(500): 1590-1598.