应用概率统计 第二十三卷
第二期 2007 年 5 月

Chinese Journal of Applied Probability
and Statistics Vol. 23 No. 2 May 2007

尺度参数变点的非参数检验 *

王黎明
(上海财经大学统计学系，上海， 200433)

王静龙
(华东师范大学统计系，上海， 200062)

摘 要

本文讨论了尺度参数模型参数变点的假设检验问题．基于两样本 U - 统计量，我们给出了两个检验，并且研究了检验统计量分布的极限性质．我们证明了这两个检验统计量的极限分布分别是 $\sup\limits_{0<t<1}|B(t)|$ 和极值分布，其中 $\{B(t), 0 \le t \le 1\}$ 是一个 Brown 桥．

关 键 词: 变点， U - 统计量，位置 – 尺度参数模型， Brown 桥，极值分布.

学 科 分 类 号: O212.1.(请补中图分类号)

§1. 引 言

设 $X_1, X_2, \cdots, X_n$ 是相互独立的连续型随机变量， $X_1, \cdots, X_k$ 来自总体 $F((x-\mu)/\sigma_1)$, $X_{k+1}, \cdots, X_n$ 来自总体 $F((x-\mu)/\sigma_2)$, 其中 k 未知正整数， μ_1, μ_2 为位置参数， $\sigma_1, \sigma_2 > 0$ 为尺度参数．其中上述模型称为位置 – 尺度参数模型．当 $\mu_1 \ne \mu_2$ 或 $\sigma_1 \ne \sigma_2$, k 称为变点．关于模型中位置参数和尺度参数变点的统计推断问题，已有不少文献进行了讨论，如 Wolfe and Schechtmen (1984) 提出了位置参数变点的非参数检验, 并且给出了变点估计的置信区间．Csorgo and Horvath (1988) 利用 U - 统计量讨论了位置参数变点的检测问题．Krishnaiah and Miao (1988), Miao (1993) 也讨论了位置参数的变点问题．王黎明和王静龙 (2002) 利用两样本 U - 统计量讨论了位置参数的变点问题，提出了检验统计量，并且给出了其极限性质．以及缪柏其和魏登云 (1994) 利用局部比较的方法讨论了尺度参数的变点问题．本文将基于 U - 统计量，研究位置 – 尺度参数模型中尺度参数变点的检验问题，并讨论其渐近性质．

我们考虑考虑尺度参数变点的检验问题，并假设根据验前信息知道 σ_1 和 σ_2 的大小关系，例如 $\sigma_1 > \sigma_2$. 由此可见，本文考虑的尺度参数变点的检验问题是

$$H_0 : \sigma_1 = \sigma_2 \leftrightarrow H_1 : \sigma_1 > \sigma_2. \tag{1}$$

至于备择假设为 $H_1 : \sigma_1 < \sigma_2$ 的尺度参数变点的检验问题类似地求解．根据文 [8], 我们引核函数

$$\phi(x_1, x_2; y_1, y_2) = I_{(|x_1-x_2|>|y_1-y_2|)} - 1/2, \tag{2}$$

* 本文获得上海财经大学 “211” 第二期标志性工程资助.

本文 2003 年 2 月 14 日收到， 2006 年 1 月 10 日收到修改稿.

其中 $I_{(|x_1-x_2|>|y_1-y_2|)}$ 为示性函数:

$$I_{(|x_1-x_2|>|y_1-y_2|)}=\begin{cases}1, & |x_1-x_2|>|y_1-y_2|;\\ 0, & |x_1-x_2|\le|y_1-y_2|.\end{cases}$$

据此得到两样本 U - 统计量

$$V_{k,n-k}=\left\{\binom{k}{2}\binom{n-k}{2}\right\}^{-1}\sum_{1\le i<j\le k}\ \sum_{k+1\le r<m\le n}\phi(X_i,X_j;X_r,X_m). \tag{3}$$

基于这个两样本 U - 统计量，我们给出了检验问题 (1) 的两个检验统计量，并讨论了当没有变点时检验统计量分布的渐近性质.

§2. 变点的非参数检验

基于 U - 统计量 (3), 我们构造如下两个的检验统计量:

$$L_{1n}=\frac{1}{2\cdot n^{3/2}}\cdot\max_{2\le k\le n-2}k(n-k)V_{k,n-k}, \tag{4}$$

$$L_{2n}=\frac{1}{2}\cdot\max_{2\le k\le n-2}\sqrt{\frac{k(n-k)}{n}}V_{k,n-k}. \tag{5}$$

关于检验问题 (1), 当 L_{1n} 或 L_{2n} 的值比较大时，拒绝原假设 H_0, 认为变点 k 存在, $\sigma_1>\sigma_2$; 否则，不能拒绝原假设 H_0, 认为没有变点. 至于备择假设为 $H_1:\sigma_1<\sigma_2$ 的尺度参数变点的检验问题，我们在 L_{1n} 或 L_{2n} 的值比较小时，拒绝原假设 H_0.

下面研究当原假设 H_0 为真，即没有变点时检验统计量 L_{1n} 或 L_{2n} 的渐近分布 $(n\to\infty)$. 我们有如下的定理.

定理 设 $X_1,X_2,\cdots,X_n$ i.i.d., 其分布函数为 $F(x)$, $F(x)$ 为连续函数，记 $0<\sigma_{10}^2=\mathrm{Var}\,(\phi_{10}(X_1))$, 其中

$$\phi_{10}(x_1)=\mathrm{E}[\phi(X_1,X_2;Y_1,Y_2)|X_1=x_1].$$

那么

(i) 在 $x>0$ 时,

$$\lim_{n\to\infty}\mathrm{P}(L_{1n}/\sigma_{10}\le x)=\mathrm{P}\Big(\sup_{0<t<1}B(t)\le x\Big), \tag{6}$$

$$\lim_{n\to\infty}\mathrm{P}(|L_{1n}|/\sigma_{10}\le x)=\mathrm{P}\Big(\sup_{0<t<1}|B(t)|\le x\Big), \tag{7}$$

其中 $\{B(t),0\le t\le 1\}$ 是一个 Brown 桥.

(ii) 对于一切的实数 x,

$$\lim_{n\to\infty}\mathrm{P}(L_{2n}/\sigma_{10}\le a(x,n))=\exp\{-e^{-x}\}, \tag{8}$$

$$\lim_{n\to\infty}\mathrm{P}(|L_{2n}|/\sigma_{10}\le a(x,n))=\exp\{-2e^{-x}\}, \tag{9}$$

其中 $a(x,n)=(2\log n)^{-1/2}(x+2\log n+(1/2)\cdot\log\log n-(1/2)\cdot\log\pi)$, 即检验统计量 L_{2n} 的极限分布为第一类极值分布.

在定理中, $\sup\limits_{0<t<1}|B(t)|$ 的分布函数为

$$\mathsf{P}\Big(\sup_{0<t<1}|B(t)|\le x\Big)=\sum_{k=-\infty}^{\infty}(-1)^k\exp(-2k^2x^2),\qquad x>0.$$

为了证明这个定理, 我们需要给出下列引理:

引理 1　设 $X_1,X_2,\cdots,X_n$ 是相互独立的随机变量, 同分布于分布函数 $F(x)$, $\mathsf{E}X_1=\mu$, $0<\mathsf{Var}\,X_1=\sigma^2<\infty$, 记 $S(k)=\sum\limits_{1\le i\le k}X_i$, 则

(i) $(1/\sigma)\cdot n^{-1/2}\max\limits_{1\le k\le n-1}|S(k)-(k/n)\cdot S(n)|\xrightarrow{L}\sup\limits_{0<t<1}|B(t)|$, 其中 $\{B(t),0\le t\le 1\}$ 是一个 Brown 桥.

(ii) $(1/\sigma)\cdot n^{-1/2}\max\limits_{1\le k\le n-1}(S(k)-(k/n)\cdot S(n))\xrightarrow{L}\sup\limits_{0<t<1}B(t)$.

(iii) 如果 $\mathsf{E}X_1^2\log\log(|X_1^2|+1)<\infty$, 对于任意的实数 x, 则

$$\lim_{n\to\infty}\mathsf{P}\Big(\max_{1\le k\le n}\Big(\frac{n}{k(n-k)}\Big)^{1/2}\Big(S(k)-\frac{k}{n}S(n)\Big)\le a(x,n)\Big)=\exp\{-e^{-x}\},$$

$$\lim_{n\to\infty}\mathsf{P}\Big(\max_{1\le k\le n}\Big(\frac{n}{k(n-k)}\Big)^{1/2}\Big|S(k)-\frac{k}{n}S(n)\Big|\le a(x,n)\Big)=\exp\{-2e^{-x}\}.$$

引理 1 的证明可以参见 Csorgo and Horvath (1997).

引理 2 (Slutsky 定理)　设随机变量序列 $\{X_n\}$ 和 $\{Y_n\}$ 满足

$$Y_n\xrightarrow{L}F,\qquad (X_n-Y_n)\xrightarrow{p}0,$$

其中 F 为连续函数, 则有 $X_n\xrightarrow{L}F$.

引理 2 的证明参见 Sen, P.K. and Singer, J.M. (1993).

引理 3　$\{X_n\}$, $\{Y_n\}$ 为任意序列, 则 $\Big|\max\limits_{1\le k\le n}|X_k|-\max\limits_{1\le k\le n}|Y_k|\Big|\le\max\limits_{1\le k\le n}|X_k-Y_k|$.

下面我们证明这个定理, 首先做一些准备工作. 为简化起见, 记核函数为

$$\phi(x_1,x_2;y_1,y_2)=\Phi(x;y),\qquad \text{其中 } x=(x_1,x_2),\ y=(y_1,y_2).$$

将对偶 $\{(X_i,X_j),1\le i<j\le k\}$ 按字典顺序排列, 即在 $i_1<i_2$ 时, (X_{i_1},X_{j_1}) 排在 (X_{i_2},X_{j_2}) 的前面, 而在 $j_1<j_2$ 时, (X_i,X_{j_1}) 排在 (X_i,X_{j_2}) 的前面. 我们用 $Z_1,\cdots,Z_K$ 表示以字典顺序排列的对偶 $\{(X_i,X_j),1\le i<j\le k\}$, 用 $Z_{K+1},\cdots,Z_N$ 表示以字典顺序排列的对偶 $\{(X_r,X_m),k+1\le r<m\le n\}$. 显然

$$K=\binom{k}{2},\qquad N-K=\binom{n-k}{2}.$$

设 $Z_i=(X_{i_1},X_{i_2})$, $Z_j=(X_{j_1},X_{j_2})$. 由于 $X_1,X_2,\cdots,X_n$ 是 i.i.d. 的连续型的随机变量序列, 并且

$$\mathsf{E}(\Phi(Z_i;Z_j)|Z_j=z_j)=\mathsf{P}(|X_{i_1}-X_{i_2}|>|X_{j_1}-X_{j_2}||X_{j_1}=x_{j_1},X_{j_2}=x_{j_2})-1/2,$$

$$\mathsf{E}(\Phi(Z_i;Z_j)|Z_i=z_i)=\mathsf{P}(|X_{i_1}-X_{i_2}|>|X_{j_1}-X_{j_2}||X_{i_1}=x_{i_1},X_{i_2}=x_{i_2})-1/2.$$

所以

$$\mathsf{E}(\Phi(Z_i;Z_j)|Z_j=z_j)+\mathsf{E}(\Phi(Z_j;Z_i)|Z_j=z_j)=0. \tag{10}$$

记 $\Phi^*(Z_i;Z_j)=\Phi(Z_i;Z_j)+\mathsf{E}(\Phi(Z_j;Z_i)|Z_j)+\mathsf{E}(\Phi(Z_j;Z_i)|Z_i)$. 则由 (10) 式知,

$$\mathsf{E}(\Phi^*(Z_i;Z_j|Z_i))=0,\qquad \mathsf{E}(\Phi^*(Z_i;Z_j|Z_j))=0. \tag{11}$$

从而在 i_1,i_2 和 j 互不相同时，有

$$\begin{aligned}\mathsf{E}[\Phi^*(Z_{i_1};Z_j)\Phi^*(Z_{i_2};Z_j)] &= \mathsf{E}[\mathsf{E}(\Phi^*(Z_{i_1};Z_j)\Phi^*(Z_{i_2};Z_j)|Z_j)]\\ &= \mathsf{E}[\mathsf{E}(\Phi^*(Z_{i_1};Z_j)|Z_j)\mathsf{E}(\Phi^*(Z_{i_2};Z_j)|Z_j)]=0.\end{aligned} \tag{12}$$

记

$$A_{k,n-k}=\sum_{1\le i<j\le k}\ \sum_{k+1\le r<m\le n}\phi(X_i,X_j;X_r,X_m),$$

则

$$\begin{aligned}A_{k,n-k} &= \sum_{1\le i\le K}\ \sum_{K+1\le j\le N}\Phi(Z_i;Z_j)\\ &= \sum_{1\le i<j\le N}\Phi(Z_i;Z_j)-\sum_{1\le i<j\le K}\Phi(Z_i;Z_j)-\sum_{K+1\le i<j\le N}\Phi(Z_i;Z_j)\\ &= B_N-B_K-B_{N-K}.\end{aligned} \tag{13}$$

B_N, B_K 和 B_{N-K} 可表示为如下形式

$$B_N=\sum_{1\le i<j\le N}\Phi^*(Z_i;Z_j)-\sum_{1\le i<j\le N}(\mathsf{E}(\Phi(Z_j;Z_i)|Z_j)+\mathsf{E}(\Phi(Z_j;Z_i)|Z_i)), \tag{14}$$

$$B_K=\sum_{1\le i<j\le K}\Phi^*(Z_i;Z_j)-\sum_{1\le i<j\le K}(\mathsf{E}(\Phi(Z_j;Z_i)|Z_j)+\mathsf{E}(\Phi(Z_j;Z_i)|Z_i)), \tag{15}$$

$$B_{N-K}=\sum_{K+1\le i<j\le N}\Phi^*(Z_i;Z_j)-\sum_{K+1\le i<j\le N}(\mathsf{E}(\Phi(Z_j;Z_i)|Z_j)+\mathsf{E}(\Phi(Z_j;Z_i)|Z_i)). \tag{16}$$

记 $B_K^*=\sum_{1\le i<j\le K}\Phi^*(Z_i;Z_j)$, $B_{N-K}^*=\sum_{K+1\le i<j\le N}\Phi^*(Z_i;Z_j)$, $1\le K\le N$. 根据 (11) 式

$$\begin{aligned}\mathsf{E}(B_{K+1}^*|\sigma(Z_1,Z_2,\cdots,Z_K)) &= \mathsf{E}\Big(\sum_{1\le i<j\le K+1}\Phi^*(Z_i;Z_j)\Big|\sigma(Z_1,Z_2,\cdots,Z_K)\Big)\\ &= B_K^*+\sum_{i=1}^{K}\mathsf{E}(\Phi^*(Z_i;Z_{K+1})|Z_i)=B_K^*.\end{aligned}$$

因此，$(B_K^*,\sigma(Z_1,Z_2,\cdots,Z_K),1\le K\le N)$ 是鞅，根据 Chebyshev 不等式，对于 $\varepsilon>0$, $\delta>0$, $K\le N$, 都有

$$\mathsf{P}\Big(\frac{|B_K^*|}{K\sqrt{N^\delta}}\ge\varepsilon\Big)\le\frac{\mathsf{E}(B_K^*)^2}{\varepsilon^2K^2N^\delta},$$

$$\begin{aligned}\mathsf{E}(B_K^*)^2 &= \mathsf{E}[B_{K-1}^*+(B_K^*-B_{K-1}^*)]\\ &= \mathsf{E}(B_{K-1}^*)^2+\mathsf{E}(B_K^*-B_{K-1}^*)^2+2\mathsf{E}[B_{K-1}^*(B_K^*-B_{K-1}^*)].\end{aligned}$$

因为 $(B_K^*, \sigma(Z_1, Z_2, \cdots, Z_K), 1 \le K \le N)$ 是鞅和 (12) 式，则

$$\mathsf{E}[B_{K-1}^*(B_K^* - B_{K-1}^*)] = \mathsf{E}\{\mathsf{E}[B_{K-1}^*(B_K^* - B_{K-1}^*)|\sigma(Z_1, Z_2, \cdots, Z_{K-1})]\} = 0.$$

所以

$$\mathsf{E}(B_K^*)^2 = \mathsf{E}(B_{K-1}^*)^2 + \mathsf{E}(B_K^* - B_{K-1}^*)^2 = \mathsf{E}(B_{K-1}^*)^2 + \mathsf{E}\Big(\sum_{1\le i\le K} \Phi^*(Z_i, Z_K)\Big)^2.$$

根据上述递推公式，则

$$\mathsf{E}(B_K^*)^2 = \sum_{L=1}^{K} \mathsf{E}\Big(\sum_{1\le i\le L} \Phi^*(Z_i, Z_L)\Big)^2.$$

因此，

$$\begin{aligned} \mathsf{P}\Big(\frac{|B_K^*|}{K\sqrt{N^\delta}} \ge \varepsilon\Big) &\le \frac{1}{\varepsilon^2 N^\delta} \sum_{L=1}^{K} \frac{1}{K^2} \mathsf{E}\Big(\sum_{1\le i<L} \Phi^*(Z_i, Z_L)\Big)^2 \\ &\le \frac{C_1}{\varepsilon^2 N^\delta} \sum_{L=1}^{K} \frac{L-1}{L^2} = O\Big(\frac{\log K}{N^\delta}\Big), \end{aligned}$$

其中 $C_1 > 0$ 是常数．所以， $|B_N^*| \overset{\mathsf{P}}{=} O(n^{2+\delta})$， $\max_{K\le N} |B_K^*| \overset{\mathsf{P}}{=} O(n^{2+\delta})$, 对所有的 $\delta > 0$ 成立．类似地可证， $\max_{K<N} |B_{N-K}^*| \overset{\mathsf{P}}{=} O(n^{2+\delta})$, 对所有的 $\delta > 0$ 也成立．由 (13) 至 (16) 诸式可见，

$$\max_{2\le k\le n-2} \Big|A_{k,n-k} - \sum_{K+1\le j\le N} \sum_{1\le i\le K} (\mathsf{E}(\Phi(Z_j; Z_i)|Z_j) + \mathsf{E}(\Phi(Z_j; Z_i)|Z_i))\Big| \overset{\mathsf{P}}{=} O(n^{2+\delta}). \tag{17}$$

在 $1 \le i \le K$, $K+1 \le j \le N$ 时， $Z_i = (X_{i_1}, X_{i_2})$ 和 $Z_j = (X_{j_1}, X_{j_2})$ 没有相同的元素 $(i_1 < i_2 < j_1 < j_2)$. 则可记

$$\mathsf{E}(\Phi(Z_i; Z_j)|Z_j = z_j) = \mathsf{P}(|X_{i_1} - X_{i_2}| > |x_{j_1} - x_{j_2}|) - 1/2 = \Phi_1(z_j),$$
$$\mathsf{E}(\Phi(Z_i; Z_j)|Z_i = z_i) = \mathsf{P}(|x_{i_1} - x_{i_2}| > |X_{j_1} - X_{j_2}|) - 1/2 = \Phi_2(z_i).$$

则由 (10) 式知， $\Phi_1(z_i) + \Phi_2(z_i) = 0$. 从而由 (14) 式得到

引理 4 在定理的条件下，对于所有的 $\delta > 0$,

$$\max_{2\le k\le n-2} \Big|A_{k,n-k} - \Big\{(N-K) \sum_{1\le i\le K} \Phi_1(Z_i) - K \sum_{K+1\le i\le N} \Phi_1(Z_i)\Big\}\Big| \overset{\mathsf{P}}{=} O(n^{2+\delta}), \tag{18}$$

而

$$\sum_{1\le i\le K} \Phi_1(Z_i) = \sum_{1\le i<j\le k} (\Phi_1(X_i, X_j)), \qquad \sum_{K+1\le i\le N} \Phi_1(Z_i) = \sum_{k+1\le i<j\le n} (\Phi_1(X_i, X_j)).$$

令

$$\phi_1(x) = \mathsf{E}(\Phi_1(X_1, X_2)|X_1 = x),$$

显然有， $\mathsf{E}(\Phi_1(X_1, X_2)|X_2 = x) = \phi_1(x)$. 则在上述条件下，下列引理成立．

引理 5 在定理的条件下,

$$\max_{2\le k\le n-2}\Big|\sum_{1\le i\le K}\Phi_1(Z_i)-(k-1)\sum_{1\le i\le k}\phi_1(X_i)\Big| \overset{\mathsf{P}}{=} O(n), \tag{19}$$

$$\max_{2\le k\le n-2}\Big|\sum_{K+1\le i\le N}\Phi_1(Z_i)-(n-k-1)\sum_{k+1\le i\le n}\phi_1(X_i)\Big| \overset{\mathsf{P}}{=} O(n). \tag{20}$$

证明:

$$\begin{aligned}&\sum_{1\le i\le K}\Phi_1(Z_i)=\sum_{1\le i<j\le k}(\Phi_1(X_i,X_j))\\ =&\sum_{1\le i<j\le k}(\Phi_1(X_i,X_j)-\phi_1(X_i)-\phi_1(X_j))+\sum_{1\le i<j\le k}\phi_1(X_i)+\sum_{1\le i<j\le k}\phi_1(X_j)\\ =&\ \Delta_k+(k-1)\sum_{1\le i\le k}\phi_1(X_i).\end{aligned}$$

同理

$$\sum_{K+1\le i\le N}\Phi_1(Z_i)=\sum_{k+1\le i<j\le n}(\Phi_1(X_i,X_j))=\Delta_{n-k}+(n-k-1)\sum_{k+1\le i\le n}\phi_1(X_i).$$

对于 Δ_k, 因为

$$\begin{aligned}&\mathsf{E}(\Delta_{k+1}|\sigma(X_1,X_2,\cdots,X_k))\\ =&\ \mathsf{E}\Big(\sum_{1\le i<j\le k}(\Phi_1(X_i,X_j)-\phi_1(X_i)-\phi_1(X_j))\Big|\sigma(X_1,X_2,\cdots,X_k)\Big)\\ =&\ \Delta_k+\sum_{i=1}^{k}\mathsf{E}((\Phi_1(X_i,X_{k+1})-\phi_1(X_i)-\phi_1(X_{k+1}))|X_i)=\Delta_k,\end{aligned}$$

所以, $(\Delta_k,\sigma(X_1,X_2,\cdots,X_k),2\le k\le n-2)$ 是鞅, 根据鞅的 Kolmogrov 不等式, 对于 $\varepsilon>0$,

$$\mathsf{P}\Big(\max_{2\le k\le n-2}|\Delta_k|>n\varepsilon\Big)\le n^{-2}\varepsilon^{-2}\mathsf{E}(\Delta_n^2)=Cn^{-2}\varepsilon^{-2}\binom{n}{2}=O(1),$$

其中 $C>0$ 为常数. 因此, (19) 式成立. 同理, 对于 (20) 式也可以类似得到. 引理 5 证毕. #

进一步地, 由引理 5, 容易得到

$$\max_{2\le k\le n-2}k^{-1}\Big|\sum_{1\le i\le K}\Phi_1(X_i)-(k-1)\sum_{1\le i\le k}\phi_1(X_i)\Big| \overset{\mathsf{P}}{=} O(1), \tag{21}$$

$$\max_{2\le k\le n-2}(n-k)^{-1}\Big|\sum_{K+1\le i\le N}\Phi_1(X_i)-(n-k-1)\sum_{k+1\le i\le n}\phi_1(X_i)\Big| \overset{\mathsf{P}}{=} O(1). \tag{22}$$

我们再考察 $2^{-1}n^{-3/2}k(n-k)V_{k,n-k}$ 的收敛情况

$$\begin{aligned}&\max_{2\le k\le n-2}\Big|2^{-1}n^{-3/2}k(n-k)V_{k,n-k}-n^{-1/2}\Big\{\sum_{1\le i\le k}\phi_1(X_i)-\frac{k}{n}\sum_{k+1\le i\le n}\phi_1(X_i)\Big\}\Big|\\ =&\max_{2\le k\le n-2}\Big|2^{-1}n^{-3/2}k(n-k)V_{k,n-k}-2^{-1}n^{-3/2}k(n-k)\Big\{N\sum_{1\le i\le K}\Phi_1(X_i)-K\sum_{K+1\le i\le N}\Phi_1(X_i)\Big\}\\ &+2^{-1}n^{-3/2}k(n-k)\Big\{N\sum_{1\le i\le K}\Phi_1(X_i)-K\sum_{K+1\le i\le N}\Phi_1(X_i)\Big\}\\ &-n^{-1/2}\Big\{\sum_{1\le i\le k}\phi_1(X_i)-\frac{k}{n}\sum_{k+1\le i\le n}\phi_1(X_i)\Big\}\Big|\\ \le&\ V_1+V_2+V_3,\end{aligned}$$

其中

$$\begin{aligned} V_1 &= \max_{2\le k\le n-2}\Big\{2n^{-3/2}(k-1)^{-1}(n-k-1)^{-1} \\ &\quad\times\Big|A_{k,n-k}-\Big\{(N-K)\sum_{1\le i\le K}\Phi_1(Z_i)-K\sum_{K+1\le i\le N}\Phi_1(Z_i)\Big\}\Big|\Big\}, \\ V_2 &= \max_{2\le k\le n-2} n^{-3/2}(k-1)^{-1}(n-k)\Big|\sum_{1\le i\le K}\Phi_1(Z_i)-(k-1)\sum_{1\le i\le k}\phi_1(Z_i)\Big|, \\ V_3 &= \max_{2\le k\le n-2} n^{-3/2}k(n-k-1)^{-1}\Big|\sum_{K+1\le i\le N}\Phi_1(Z_i)-(n-k-1)\sum_{k+1\le i\le n}\phi_1(Z_i)\Big|. \end{aligned}$$

对于 V_1，根据引理 4，易得

$$V_1 \overset{\mathsf{P}}{=} O(n^{\delta-1/2}). \tag{23}$$

关于 V_2，由 (21) 式

$$\begin{aligned} V_2 &= \max_{2\le k\le n-2} n^{-3/2}(k-1)^{-1}(n-k)\Big|\sum_{1\le i\le K}\Phi_1(Z_i)-(k-1)\sum_{1\le i\le k}\phi_1(Z_i)\Big| \\ &\le n^{-3/2}\max_{2\le k\le n-2}\Big(\frac{k}{k-1}\Big)(n-k)\max_{2\le k\le n-2}k^{-1}\Big|\sum_{1\le i\le K}\Phi_1(Z_i)-(k-1)\sum_{1\le i\le k}\phi_1(Z_i)\Big| \\ &= n^{-3/2}\max_{2\le k\le n-2}k^{-1}\Big|\sum_{1\le i\le K}\Phi_1(Z_i)-(k-1)\sum_{1\le i\le k}\phi_1(Z_i)\Big| \\ &\overset{\mathsf{P}}{=} n^{-3/2}(n-2)\overset{\mathsf{P}}{=} O(n^{-1/2}). \end{aligned} \tag{24}$$

同理，对于 V_3，我们可以类似得到

$$V_3 \overset{\mathsf{P}}{=} O(n^{-1/2}). \tag{25}$$

综上所述，结合 (23), (24) 和 (25), 我们证明了下列引理：

引理 6 在定理的条件下，对于所有的 $\delta>0$,

$$\max_{2\le k\le n-2}\Big|L_{1n}-n^{1/2}\Big\{\sum_{1\le i\le k}\phi_1(X_i)-\frac{k}{n}\sum_{1\le i\le n}\phi_1(X_i)\Big\}\Big|\overset{\mathsf{P}}{=} O(n^{\delta-1/2}), \tag{26}$$

$$\max_{2\le k\le n-2}\Big|L_{2n}-n^{1/2}\Big\{\sum_{1\le i\le k}\phi_1(X_i)-\frac{k}{n}\sum_{1\le i\le n}\phi_1(X_i)\Big\}\Big|\overset{\mathsf{P}}{=} O(n^{\delta-1/2}). \tag{27}$$

这是因为

$$L_{1n}=\frac{1}{2n^{3/2}}\max_{2\le k\le n-2}k(n-k)V_{k,n-k}, \qquad L_{2n}=\frac{1}{2}\max_{2\le k\le n-2}\sqrt{k(n-k)}\cdot V_{k,n-k},$$

所以引理 6 显然成立.

定理的证明： 在定理的条件下，随机变量 $\phi_1(X_1),\phi_1(X_2),\cdots,\phi_1(X_n)$ 相互独立，并且服从同一分布，记

$$\begin{aligned} T_{1n} &= \frac{1}{\sigma_{10}\sqrt{n}}\Big(\sum_{1\le i\le k}\phi_1(X_i)-\frac{k}{n}\sum_{1\le i\le n}\phi_1(X_i)\Big), \\ T_{2n} &= \frac{1}{\sigma_{10}}\sqrt{\frac{n}{k(n-k)}}\Big(\sum_{1\le i\le k}\phi_1(X_i)-\frac{k}{n}\sum_{1\le i\le n}\phi_1(X_i)\Big). \end{aligned}$$

则 T_{1n} 满足引理 1, 所以 $T_{1n} \xrightarrow{L} \sup\limits_{0<t<1} B(t)$, $|T_{1n}| \xrightarrow{L} \sup\limits_{0<t<1} |B(t)|$. 由 (26) 式和引理 3 可知，取 $0<\delta<1/2$, 对任意的 $\varepsilon>0$,

$$\lim_{n\to\infty} \mathsf{P}\Big(\Big|\max_{2\le k\le n-2} \frac{1}{2\sigma_{10} n^{3/2}} k(n-k)V_{k,n-k} - \max_{2\le k\le n-2} T_{1n}\Big| > \varepsilon\Big) = 0, \tag{28}$$

$$\lim_{n\to\infty} \mathsf{P}\Big(\Big|\max_{2\le k\le n-2} \Big|\frac{1}{2\sigma_{10} n^{3/2}} k(n-k)V_{k,n-k}\Big| - \max_{2\le k\le n-2} |T_{1n}|\Big| > \varepsilon\Big) = 0. \tag{29}$$

由引理 1, 对任意的 $x>0$,

$$\lim_{n\to\infty} \mathsf{P}(L_{1n}/\sigma_{10} \le x) = \mathsf{P}\Big(\sup_{0<t<1} B(t) \le x\Big), \tag{30}$$

$$\lim_{n\to\infty} \mathsf{P}(|L_{1n}|/\sigma_{10} \le x) = \mathsf{P}\Big(\sup_{0<t<1} |B(t)| \le x\Big). \tag{31}$$

同理，因为核函数 ϕ 满足条件 (3), 则 T_{2n} 满足引理 1, 即对于一切的实数 x,

$$\lim_{n\to\infty} \mathsf{P}\Big(\max_{2\le k\le n-2} T_{2n} \le a(x,n)\Big) = \exp(-e^{-x}),$$

$$\lim_{n\to\infty} \mathsf{P}\Big(\max_{2\le k\le n-2} |T_{2n}| \le a(x,n)\Big) = \exp(-2e^{-x}).$$

由 (25) 式，取 $0<\delta<1/2$, 并且根据引理 3 可知，对任意的 $\varepsilon>0$,

$$\lim_{n\to\infty} \mathsf{P}\Big(\Big|\max_{2\le k\le n-2} \frac{1}{2\sigma_{10}} \sqrt{\frac{k(n-k)}{n}} V_{k,n-k} - \max_{2\le k\le n-2} T_{2n}\Big| > \varepsilon\Big) = 0, \tag{32}$$

$$\lim_{n\to\infty} \mathsf{P}\Big(\Big|\max_{2\le k\le n-2} \Big|\frac{1}{2\sigma_{10}} \sqrt{\frac{k(n-k)}{n}} V_{k,n-k}\Big| - \max_{2\le k\le n-2} |T_{2n}|\Big| > \varepsilon\Big) = 0. \tag{33}$$

由引理 1, 对任意的实数 x,

$$\lim_{n\to\infty} \mathsf{P}(L_{2n}/\sigma_{10} \le a(x,n)) = \exp(-e^{-x}), \tag{34}$$

$$\lim_{n\to\infty} \mathsf{P}(|L_{2n}|/\sigma_{10} \le a(x,n)) = \exp(-2e^{-x}). \tag{35}$$

定理证毕. #

作者对审稿专家提出的修改建议表示衷心的感谢.

参 考 文 献

[1] Csorgo, M. and Horvath, L., *Limit Theorems in Change-Point Analysis*, John Wiley & Sons, 1997.

[2] Aly, E.-E. and Kochar, S.C., Change point tests based on U-statistics with applications in reliability, *Metrika*, **45**(1997), 259–269.

[3] Hall, P., On the invariance principle for U-statistics, *Stochastic Processes and their Application*, **9**(1979), 163–174.

[4] Janson, S. and Wichura, M.J., Invariance principle for stochastic area and related stochastic integrals, *Stochastic Processes and their Application*, **16**(1983), 71–84.

[5] Sen, P.K. and Singer, J.M., *Large Sample Methods in Statistics*, Chapmen & Hall, 1993.

[6] Wang, J.L. and Bhatti, M.I., Tests and confidence interval for change-point in scale of two-parameter exponential distribution, *Statistica*, **58(1)**(1998), 127–136.

[7] Wang, J.L. and Bhatti, M.I., Three tests for a change-point in variance of normal distribution, *Chinese Journal of Applied Probability and Statistics*, **14(2)**(1998), 113–121.

[8] 陈希孺，方兆本，李国英，陶波，非参数统计, 上海科学技术出版社，1989.

[9] 缪柏其，关于只有一个转变点模型的非参数推断，系统科学与数学, **13(2)**(1993), 132–140.

[10] 缪柏其，魏登云，关于刻度参数变点的非参数统计推断，中国科技大学学报, **24(3)**(1994), 263–270.

[11] 王黎明，王静龙，位置参数变点的非参数检验及其渐近性质，数学年刊 (A 辑), **23(2)**(2002), 229–234.

Non-Parameter Test for Change-Point of Scale Parameter

Wang Liming

(*Department of Statistics, Shanghai University of Finance & Economics, Shanghai, 200433*)

Wang Jinglong

(*Department of Statistics, East China Normal University, Shanghai, 200062*)

The change-point problem about the scale parameter is discussed in this paper. Based on two-sample U-statistic two tests are proposed, and their approximate distributions, which are $\sup_{0<t<1}|B(t)|$ and extreme valve distribution respectively, where $\{B(t), 0 \le t \le 1\}$ is a Brownian bridge, are obtained.