

An application of the maximum likelihood test to the change-point problem

Edit Gombay *

Department of Statistics and Applied Probability, University of Alberta, Edmonton, Alberta, Canada

Lajos Horváth

Department of Mathematics, University of Utah, Salt Lake City, UT, USA

Received 5 October 1992

Revised 15 March 1993

A maximum-likelihood-type statistic is derived for testing a sequence of observations for no change in the parameter against a possible change. We prove that the limit distribution of the suitably normalized and centralized statistic is double exponential under the null hypothesis.

AMS 1980 Subject Classifications: Primary 62A10; Secondary 62F03.

maximum likelihood; parameter estimation; standardized partial sums; limit theorem; double exponential distribution

1. Introduction and results

Let X_1, \dots, X_n be independent random vectors with distribution functions $F(x; \theta_1), \dots, F(x; \theta_n)$, where $x \in \mathbb{R}^m$ and $\theta_i \in \Theta \subseteq \mathbb{R}^d$, $1 \leq i \leq n$. We want to test

$$H_0: \theta_1 = \dots = \theta_n$$

against the change-point alternative

$$H_A: \text{there is } k^*, 1 \leq k^* < n, \text{ such that } \theta_1 = \theta_2 = \dots = \theta_{k^*} \neq \theta_{k^*+1} = \dots = \theta_n.$$

The observations X_1, \dots, X_n have probability densities $f(x; \theta_1), \dots, f(x; \theta_n)$ with respect to ν , where ν is a σ -finite measure. If we know that the change can occur only after the k th observation, then this is a two-sample problem and we should reject for small values of

Correspondence to: Prof. Lajos Horváth, Department of Mathematics, University of Utah, Salt Lake City, UT 84112, USA.

* Research supported in part by a NSERC Canada operating grant.

$$\Lambda_k = \frac{\sup_{\theta \in \Theta} \prod_{1 \leq i \leq n} f(X_i; \theta)}{\sup_{\theta \in \Theta} \prod_{1 \leq i \leq k} f(X_i; \theta) \sup_{\theta \in \Theta} \prod_{k+1 \leq i \leq n} f(X_i; \theta)} . \quad (1.1)$$

Let

$$g_i(x; \theta) = \frac{\partial}{\partial \theta_i} \log f(x; \theta) .$$

We assume the following:

(C1) The likelihood equations $\sum_{m \leq j \leq k} g_i(X_j; \theta) = 0$, $1 \leq i \leq d$, have a unique solution for all $1 \leq m < k \leq n$.

This assumption implies immediately that we can find a unique $\hat{\theta}_k$, $\hat{\theta}_n$ and θ_{n-k}^* such that

$$\sum_{1 \leq j \leq k} g_i(X_j; \hat{\theta}_k) = 0 \quad \text{for all } 1 \leq i \leq d , \quad (1.2)$$

$$\sum_{1 \leq j \leq n} g_i(X_j; \hat{\theta}_n) = 0 \quad \text{for all } 1 \leq i \leq d , \quad (1.3)$$

and

$$\sum_{k+1 \leq j \leq n} g_i(X_j; \theta_{n-k}^*) = 0 \quad \text{for all } 1 \leq i \leq d . \quad (1.4)$$

Hence we write

$$\Lambda_k = \frac{\prod_{1 \leq j \leq n} f(X_j; \hat{\theta}_n)}{\prod_{1 \leq j \leq k} f(X_j; \hat{\theta}_k) \prod_{k+1 \leq j \leq n} f(X_j; \theta_{n-k}^*)} . \quad (1.5)$$

Since k^* is unknown we reject H_0 , if

$$Z_n = \max_{1 \leq k \leq n-1} (-2 \log \Lambda_k) \quad (1.6)$$

is large.

The main aim of this note is the computation of the limit distribution of Z_n . The true value of the parameter under H_0 is denoted by θ_0 . Let

$$g(x; \theta) = \log f(x; \theta)$$

and

$$g_{i_1 \dots i_r}(x; \theta) = \frac{\partial^r g(x; \theta)}{\partial \theta_{i_1} \dots \partial \theta_{i_r}} .$$

We assume the following regularity conditions:

(C2) There is an open interval $\Theta_0 \subseteq \mathbb{R}^d$ containing θ_0 such that $g_i(x; \theta)$, $g_{ij}(x; \theta)$ and $g_{ijk}(x; \theta)$, $1 \leq i, j, k \leq d$, exist and are continuous in θ for all $x \in \mathbb{R}^m$ and $\theta \in \Theta_0$.

(C3) There are functions $m(\mathbf{x})$ and $M(\mathbf{x})$ such that $|g_i(\mathbf{x}; \boldsymbol{\theta})| \leq m(\mathbf{x})$, $|g_{ij}(\mathbf{x}; \boldsymbol{\theta})| \leq m(\mathbf{x})$, $|g_{ijk}(\mathbf{x}; \boldsymbol{\theta})| \leq M(\mathbf{x})$ for all $\mathbf{x} \in \mathbb{R}^m$, $\boldsymbol{\theta} \in \Theta_0$, $1 \leq i, j, k \leq d$ and

$$\int_{\mathbb{R}^m} m(\mathbf{x}) \nu(d\mathbf{x}) < \infty, \quad E_{\theta_0} M(X_1) < \infty.$$

(C4) $E_{\theta} g_i(X_1; \boldsymbol{\theta}) = 0$ for all $1 \leq i \leq d$.

(C5) $I_{ij}(\boldsymbol{\theta}) = E_{\theta} g_i(X_1; \boldsymbol{\theta}) g_j(X_1; \boldsymbol{\theta}) = -E_{\theta} g_{ij}(X_1; \boldsymbol{\theta})$, $1 \leq i, j \leq d$, $I(\boldsymbol{\theta}) = \{I_{ij}(\boldsymbol{\theta}), 1 \leq i, j \leq d\}$ and $I^{-1}(\boldsymbol{\theta})$ exist and are continuous for all $\boldsymbol{\theta} \in \Theta_0$.

(C6) $\text{Var } g_{ij}(X_1; \boldsymbol{\theta}_0) < \infty$, $1 \leq i, j \leq d$.

(C7) $E_{\theta_0} |g_i(X_1; \boldsymbol{\theta}_0)|^{2+\delta} < \infty$, $1 \leq i \leq d$, for some $\delta > 0$.

Let $a(t) = (2 \log t)^{1/2}$ and $b_d(t) = 2 \log t + \frac{1}{2} d \log \log t - \log \Gamma(\frac{1}{2} d)$ where

$$\Gamma(d) = \int_0^{\infty} y^{d-1} e^{-y} dy$$

is the Gamma function.

Theorem. If H_0 and (C1)–(C7) hold, then we have

$$\lim_{n \rightarrow \infty} P\{a(\log n) Z_n^{1/2} \leq t + b_d(\log n)\} = \exp(-2e^{-t}) \quad (1.7)$$

for all t .

Remark. We note that (1.7) remains true if (C1) is replaced by the condition that there is a random variable $k_0(\omega)$ such that $\sum_{m \leq j \leq k} g_i(X_j; \boldsymbol{\theta}) = 0$, $1 \leq i \leq d$, have a unique solution for each $m, k \geq k_0$.

Several authors applied the likelihood ratio test to the change-point problem. Assuming that the observations are univariate normals with constant variance, Sen and Srivastava (1975a, b), Hawkins (1977), Yao and Davis (1986) and James et al. (1987) studied the likelihood ratio test. The limit distribution of the likelihood ratio test in the univariate normal case with possible change in the mean only was obtained by Yao and Davis (1986), Csörgő and Horváth (1988) and Gombay and Horváth (1990). Horváth (1993) obtained the limit of the likelihood ratio when changes can occur in the means and variances of univariate normal observations. Srivastava and Worsley (1986) and James et al. (1992) studied the likelihood ratio test in case of possible changes in the mean of multivariate normal random vectors. Worsley (1986a, b), Haccou et al. (1988) and Gombay and Horváth (1990) considered the exponential case. Worsley (1983) and Horváth (1989) derived the likelihood ratio test in the binomial case. Yao and Davis (1986) and Haccou et al. (1988) investigated the properties of the likelihood ratio tests under the alternative, assuming that

we have normal or exponential observations. If $k^* = [n\tau]$, where $0 < \tau < 1$, then Z_n is optimal in the Bahadur sense of efficiency in case of exponential r.v.'s (Haccou et al., 1988).

Csörgő and Horváth (1988) contains a survey of nonparametric methods which also can be used for testing H_0 against H_A . Ferger (1991) obtained asymptotic results for some nonparametric tests under the alternative hypothesis.

2. Proofs

We start with some preliminary lemmas. Let $|\mathbf{x}| = \max_{1 \leq i \leq m} |x_i|$, where $\mathbf{x} = (x_1, \dots, x_m)$, and let \mathbf{x}^T be the transpose of \mathbf{x} . We also define

$$\hat{\mathbf{Q}}_k = \left(\sum_{1 \leq j \leq k} g_1(X_j; \boldsymbol{\theta}_0), \dots, \sum_{1 \leq j \leq k} g_d(X_j; \boldsymbol{\theta}_0) \right)$$

and

$$\hat{\mathbf{Q}}_{n-k}^* = \left(\sum_{k+1 \leq j \leq n} g_1(X_j; \boldsymbol{\theta}_0), \dots, \sum_{k+1 \leq j \leq n} g_d(X_j; \boldsymbol{\theta}_0) \right).$$

Lemma 2.1. *We assume that H_0 and (C1)–(C6) hold. Then, as $n \rightarrow \infty$, we have*

$$\max_{1 \leq k \leq n} \frac{k}{\log \log k} \left| \hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0 - \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \right| = O_p(1) \quad (2.1)$$

and

$$\max_{1 \leq k < n} \frac{n-k}{\log \log(n-k)} \left| \hat{\boldsymbol{\theta}}_{n-k}^* - \boldsymbol{\theta}_0 - \frac{1}{n-k} \hat{\mathbf{Q}}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \right| = O_p(1). \quad (2.2)$$

Proof. The proof is a combination of the methods in Ibragimov and Hašminskii (1973), Serfling (1980, pp. 144–149) and Lehmann (1991, p. 415). We show that

$$\left| \hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0 - \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \right| \stackrel{\text{a.s.}}{=} O\left(\frac{1}{k} \log \log k\right), \quad \text{as } k \rightarrow \infty, \quad (2.3)$$

which immediately implies (2.1) and symmetry gives (2.2). First we prove that

$$\limsup_{k \rightarrow \infty} (k / \log \log k)^{1/2} |\hat{\boldsymbol{\theta}}_k - \boldsymbol{\theta}_0| < \infty \quad \text{a.s.} \quad (2.4)$$

It follows from Serfling (1980, p. 149) that

$$\lim_{k \rightarrow \infty} \hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta}_0 \quad \text{a.s.} \quad (2.5)$$

A three-term Taylor expansion yields for all $1 \leq i \leq d$ that

$$\begin{aligned}
& \sum_{1 \leq j \leq k} g_i(X_j; \hat{\theta}_k) - \sum_{1 \leq j \leq k} g_i(X_j; \theta_0) \\
&= \sum_{1 \leq l \leq d} (\hat{\theta}_{kl} - \theta_{0l}) A_{il}^{(k)} + \frac{1}{2} \sum_{1 \leq l \leq d} \sum_{1 \leq r \leq d} (\hat{\theta}_{kl} - \theta_{0l})(\hat{\theta}_{kr} - \theta_{0r}) B_{ilr}^{(k)}, \quad (2.6)
\end{aligned}$$

where $\hat{\theta}_k = (\theta_{k1}, \dots, \theta_{kd})$, $\theta_0 = (\theta_{01}, \dots, \theta_{0d})$,

$$A_{il}^{(k)} = \sum_{1 \leq j \leq k} g_{il}(X_j; \theta_0), \quad (2.7)$$

and by (C3) and (2.5) we have

$$|B_{ilr}^{(k)}| \leq \sum_{1 \leq j \leq k} M(X_j). \quad (2.8)$$

Now by (1.2) we can rewrite (2.6) as

$$- \sum_{1 \leq j \leq k} g_i(X_j; \theta_0) = \sum_{1 \leq l \leq d} (\hat{\theta}_{kl} - \theta_{0l}) C_{il}^{(k)}, \quad (2.9)$$

and by the strong law of large numbers and (2.5) we have

$$\lim_{k \rightarrow \infty} \frac{1}{k} C_{il}^{(k)} = E_{\theta_0} g_{il}(X_1; \theta_0) \quad \text{a.s.} \quad (2.10)$$

for all $1 \leq i, l \leq d$. By (C4) we get that $E_{\theta_0} g_i(X_k; \theta_0) = 0$, and therefore the law of the iterated logarithm yields

$$\left| \sum_{1 \leq j \leq k} g_i(X_j; \theta_0) \right| \stackrel{\text{a.s.}}{=} O((k \log \log k)^{1/2}), \quad \text{as } k \rightarrow \infty. \quad (2.11)$$

Using (C5) we can find a r.v. k_0 , such that $\{(1/k) C_{il}^{(k)}, 1 \leq i, l \leq d\}$ has an inverse, if $k \geq k_0$ and the almost sure limit of the inverse is $-I^{-1}(\theta_0)$. Hence by (2.9) and (2.11) we have (2.4).

We use again (2.6). The strong law of large numbers, (2.8) and (2.4) implies

$$\left| \sum_{1 \leq l \leq d} \sum_{1 \leq r \leq d} (\hat{\theta}_{kl} - \theta_{0l})(\hat{\theta}_{kr} - \theta_{0r}) B_{ilr}^{(k)} \right| \stackrel{\text{a.s.}}{=} O(\log \log k) \quad \text{as } k \rightarrow \infty. \quad (2.12)$$

The law of iterated logarithm yields that

$$\left| \sum_{1 \leq j \leq k} g_{il}(X_j; \theta_0) + k I_{il} \right| \stackrel{\text{a.s.}}{=} O((k \log \log k)^{1/2}) \quad \text{as } k \rightarrow \infty, \quad (2.13)$$

and therefore from (2.4) we get

$$\sum_{1 \leq l \leq d} (\hat{\theta}_{kl} - \theta_{0l})(A_{il}^{(k)} + k I_{il}) \stackrel{\text{a.s.}}{=} O(\log \log k) \quad \text{as } k \rightarrow \infty. \quad (2.14)$$

Now (2.3) follows immediately from (2.6), (2.12) and (2.14). \square

Lemma 2.2. *We assume that H_0 and (C1)–(C6) hold. Then, as $n \rightarrow \infty$, we have*

$$\max_{1 \leq k \leq n} \frac{k^{1/2}}{(\log \log k)^{3/2}} \left| \sum_{1 \leq i \leq k} (\log f(X_i; \hat{\theta}_k) - \log f(X_i; \theta_0)) - \frac{k}{2} (\hat{\theta}_k - \theta_0) I(\theta_0) (\hat{\theta}_k - \theta_0)^T \right| = O_p(1) \quad (2.15)$$

and

$$\max_{1 \leq k < n} \frac{(n-k)^{1/2}}{(\log \log(n-k))^{3/2}} \left| \sum_{k+1 \leq i \leq n} (\log f(X_i; \theta_{n-k}^*) - \log f(X_i; \theta_0)) - \frac{n-k}{2} (\theta_{n-k}^* - \theta_0) I(\theta_0) (\theta_{n-k}^* - \theta_0)^T \right| = O_p(1). \quad (2.16)$$

Proof. First we apply a three-term then a two-term Taylor expansion and get

$$\begin{aligned} & g(X_i; \theta_0) - g(X_i; \hat{\theta}_k) \\ &= \sum_{1 \leq j \leq d} g_j(X_i; \hat{\theta}_k) (\theta_{0j} - \hat{\theta}_{kj}) \\ &+ \frac{1}{2} \sum_{1 \leq j \leq d} \sum_{1 \leq l \leq d} g_{jl}(X_i; \hat{\theta}_k) (\theta_{0j} - \hat{\theta}_{kj}) (\theta_{0l} - \hat{\theta}_{kl}) \\ &+ \frac{1}{6} \sum_{1 \leq j, l, r \leq d} g_{jlr}(X_i; \eta_i^{(1)}) (\theta_{0j} - \hat{\theta}_{kj}) (\theta_{0l} - \hat{\theta}_{kl}) (\theta_{0r} - \hat{\theta}_{kr}) \\ &= \sum_{1 \leq j \leq d} g_j(X_i; \hat{\theta}_k) (\theta_{0j} - \hat{\theta}_{kj}) \\ &+ \frac{1}{2} \sum_{1 \leq j \leq d} \sum_{1 \leq l \leq d} g_{jl}(X_i; \theta_0) (\theta_{0j} - \hat{\theta}_{kj}) (\theta_{0l} - \hat{\theta}_{kl}) + U_i, \end{aligned} \quad (2.17)$$

where

$$\begin{aligned} U_i = & \sum_{1 \leq j, l, r \leq d} \{ \frac{1}{6} g_{jlr}(X_i; \eta_i^{(1)}) - \frac{1}{2} g_{jlr}(X_i; \eta_i^{(2)}) \} \\ & \times (\theta_{0j} - \hat{\theta}_{kj}) (\theta_{0l} - \hat{\theta}_{kl}) (\theta_{0r} - \hat{\theta}_{kr}), \end{aligned}$$

and $|\eta_i^{(1)} - \theta_0| \leq |\theta_0 - \hat{\theta}_k|$, $|\eta_i^{(2)} - \theta_0| \leq |\theta_0 - \hat{\theta}_k|$. Condition (C3) and the strong law of large numbers give

$$\sum_{1 \leq i \leq k} \{ |g_{jlr}(X_i; \eta_i^{(1)})| + |g_{jlr}(X_i; \eta_i^{(2)})| \} \stackrel{\text{a.s.}}{=} O(k), \quad (2.18)$$

as $k \rightarrow \infty$, and therefore (2.4) yields

$$\limsup_{k \rightarrow \infty} \frac{k^{1/2}}{(\log \log k)^{3/2}} \left| \sum_{1 \leq i \leq k} U_i \right| < \infty \quad \text{a.s.} \quad (2.19)$$

The law of iterated logarithm yields

$$\limsup_{k \rightarrow \infty} (k \log \log k)^{-1/2} \left| \sum_{1 \leq i \leq k} (g_{ji}(X_i; \theta_0) + I_{ji}(\theta_0)) \right| < \infty \quad \text{a.s.}, \quad (2.20)$$

and therefore by (2.4) we have

$$\begin{aligned} \limsup_{k \rightarrow \infty} \frac{k^{1/2}}{(\log \log k)^{3/2}} \left| \sum_{1 \leq i \leq k} \sum_{1 \leq j, l \leq d} \{g_{ji}(X_i; \theta_0) + I_{ji}(\theta_0)\} \right. \\ \left. \times (\theta_{0j} - \hat{\theta}_{kj})(\theta_{0l} - \hat{\theta}_{kl}) \right| < \infty \quad \text{a.s.} \end{aligned} \quad (2.21)$$

Now (2.15) follows from (1.2), (2.17), (2.19) and (2.21).

It is clear that (2.15) implies (2.16). \square

Putting together Lemmas 2.1, 2.2 and (2.9) we get the following result.

Lemma 2.3. *We assume that H_0 and (C1)–(C6) hold. Then, as $n \rightarrow \infty$, we have*

$$\begin{aligned} \max_{1 \leq k \leq n} \frac{k^{1/2}}{(\log \log k)^{3/2}} \left| \sum_{1 \leq i \leq k} (\log f(X_i; \hat{\theta}_k) - \log f(X_i; \theta_0)) \right. \\ \left. - \frac{1}{2k} \hat{Q}_k I^{-1}(\theta_0) \hat{Q}_k^T \right| = O_P(1) \end{aligned} \quad (2.22)$$

and

$$\begin{aligned} \max_{1 \leq k < n} \frac{(n-k)^{1/2}}{(\log \log (n-k))^{3/2}} \left| \sum_{k+1 \leq i \leq n} (\log f(X_i; \theta_{n-k}^*) - \log f(X_i; \theta_0)) \right. \\ \left. - \frac{1}{2(n-k)} Q_{n-k}^* I^{-1}(\theta_0) Q_{n-k}^{*T} \right| \\ = O_P(1). \quad \square \end{aligned} \quad (2.23)$$

The proof of Theorem also requires the following technical lemma.

Lemma 2.4. *Let $\{\xi_i = (\xi_{i1}, \dots, \xi_{id}), 1 \leq i < \infty\}$ be a sequence of i.i.d.r. vectors satisfying $E\xi_{i1} = 0$, $E\xi_{i1}^2 = 1$, $E\xi_{i1}\xi_{ij} = 0$, $1 \leq i, j \leq d$, $i \neq j$ and*

$$\max_{1 \leq j \leq d} E|\xi_{ij}|^\tau < \infty \quad \text{for some } \tau > 2.$$

Then we have

$$\lim_{n \rightarrow \infty} P \left\{ a(\log n) \max_{1 \leq k \leq n} \left(\frac{1}{k} \sum_{1 \leq j \leq d} \left(\sum_{1 \leq i \leq k} \xi_{ij} \right)^2 \right)^{1/2} - b_d(\log n) \leq t \right\} \\ = \exp(-e^{-t}) \quad (2.24)$$

for all t .

Proof. It can be found in Horváth (1993) \square .

Proof of Theorem. Let $c(n) = \log n$ and $b(n) = n/\log n$. We can write

$$Z_n = \max(A_n^{(1)}, \dots, A_n^{(6)}), \quad (2.25)$$

where

$$\begin{aligned} A_n^{(1)} &= \max_{1 \leq k \leq c(n)} (-2 \log \Lambda_k), \\ A_n^{(2)} &= \max_{c(n) \leq k \leq b(n)} (-2 \log \Lambda_k), \\ A_n^{(3)} &= \max_{b(n) \leq k \leq n/2} (-2 \log \Lambda_k), \\ A_n^{(4)} &= \max_{n/2 \leq k \leq n-b(n)} (-2 \log \Lambda_k), \\ A_n^{(5)} &= \max_{n-b(n) \leq k \leq n-c(n)} (-2 \log \Lambda_k), \end{aligned}$$

and

$$A_n^{(6)} = \max_{n-c(n) \leq k < n} (-2 \log \Lambda_k).$$

First we show that

$$a^2(\log n) A_n^{(1)} - (x + b_d(\log n))^2 \xrightarrow{P} -\infty \quad (2.26)$$

for all x . By Lemma 2.3 we have

$$\begin{aligned} A_n^{(1)} &= O_P(1) + \max_{1 \leq k \leq c(n)} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T \\ &\quad + \max_{1 \leq k \leq c(n)} \left| \frac{1}{n-k} \mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*T} - \frac{1}{n} \hat{\mathbf{Q}}_n I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_n^T \right|. \end{aligned} \quad (2.27)$$

We can find a sequence of i.i.d.r. vectors $\boldsymbol{\eta}_i = (\eta_{i1}, \dots, \eta_{id})$, $1 \leq i < \infty$, with $E\eta_{i1} = \dots = E\eta_{id} = 0$, $E\eta_{ij}^2 = 1$, $1 \leq j \leq d$, $E\eta_{ij}\eta_{ik} = 0$, $1 \leq j \neq k \leq d$, $E|\eta_{ij}|^{2+\delta} < \infty$ such that

$$\hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T = \left(\sum_{1 \leq i \leq k} \eta_{i1} \right)^2 + \dots + \left(\sum_{1 \leq i \leq k} \eta_{id} \right)^2 \quad (2.28)$$

and

$$\mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*\text{T}} = \left(\sum_{k+1 \leq i \leq n} \eta_{i1} \right)^2 + \cdots + \left(\sum_{k+1 \leq i \leq n} \eta_{id} \right)^2. \quad (2.29)$$

Darling and Erdős (1956) showed that

$$\begin{aligned} \lim_{N/L \rightarrow \infty} P \left\{ a \left(\log \frac{N}{L} \right) \max_{L \leq k \leq N} \left| k^{-1/2} \sum_{1 \leq i \leq k} \eta_{ij} \right| \leq x + b_1 \left(\log \frac{N}{L} \right) \right\} \\ = \exp(-e^{-x}), \end{aligned} \quad (2.30)$$

for all x . Also, if $N \rightarrow \infty$, $L \rightarrow \infty$, $L < N$ and $\limsup N/L < \infty$, we have

$$\max_{L \leq k \leq N} \left| k^{-1/2} \sum_{1 \leq i \leq k} \eta_{ij} \right| = O_P(1). \quad (2.31)$$

Now (2.31) implies that

$$\max_{1 \leq k \leq c(n)} \left| \frac{1}{n-k} \mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*\text{T}} - (1/n) \hat{\mathbf{Q}}_n I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_n^{\text{T}} \right| = O_P(1) \quad (2.32)$$

and (2.30) yields

$$\max_{1 \leq k \leq c(n)} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^{\text{T}} = O_P(\log \log c(n)). \quad (2.33)$$

It is clear that (2.26) follows from (2.27), (2.32) and (2.33).

Applying again Lemma 2.3 we obtain that

$$\begin{aligned} \max_{c(n) \leq k \leq n/2} \left| -2 \log A_k - \left\{ \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^{\text{T}} \right. \right. \\ \left. \left. + \frac{1}{n-k} \mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*\text{T}} - \frac{1}{n} \hat{\mathbf{Q}}_n I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_n^{\text{T}} \right\} \right| \\ = O_P \left(\frac{(\log \log n)^{3/2}}{(\log n)^{1/2}} \right). \end{aligned} \quad (2.34)$$

The invariance principle for partial sums of i.i.d.r. vectors yields

$$\begin{aligned} \max_{c(n) \leq k \leq b(n)} \left| \frac{1}{n-k} \mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*\text{T}} - \frac{1}{n} \hat{\mathbf{Q}}_n I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_n^{\text{T}} \right| \\ = O_P \left(\left(\frac{\log \log n}{\log n} \right)^{1/2} \right). \end{aligned} \quad (2.35)$$

By (2.34) and (2.30) we have

$$\max_{b(n) \leq k \leq n/2} | -2 \log A_k | = O_P(\log \log \log n), \quad (2.36)$$

and therefore

$$a^2(\log n)A_n^{(3)} - (x + b_d(\log n))^2 \xrightarrow{P} -\infty \quad (2.37)$$

for all x . Now (2.26) and (2.37) imply that

$$\max(A_n^{(1)}, A_n^{(2)}, A_n^{(3)}) \quad \text{and} \quad \max_{c(n) \leq k \leq b(n)} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T$$

are asymptotically equivalent. Using again (2.30) one can easily establish that

$$a^2(\log n) \max_{1 \leq k \leq c(n)} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T - (x + b_d(\log n))^2 \xrightarrow{P} -\infty \quad (2.38)$$

and

$$a^2(\log n) \max_{b(n) \leq k \leq n/2} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T - (x + b_d(\log n))^2 \xrightarrow{P} -\infty \quad (2.39)$$

for all x . Thus by Lemma 2.4 and (2.28) we get that

$$\max(A_n^{(1)}, A_n^{(2)}, A_n^{(3)}) \quad \text{and} \quad \max_{1 \leq k \leq n/2} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T$$

have the same limit distribution. Similar arguments show that

$$\max(A_n^{(4)}, A_n^{(5)}, A_n^{(6)}) \quad \text{and} \quad \max_{n/2 \leq k \leq n} \frac{1}{n-k} \mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*T}$$

are asymptotically equivalent. Since

$$\max_{1 \leq k \leq n/2} \frac{1}{k} \hat{\mathbf{Q}}_k I^{-1}(\boldsymbol{\theta}_0) \hat{\mathbf{Q}}_k^T \quad \text{and} \quad \max_{n/2 < k < n} \frac{1}{n-k} \mathbf{Q}_{n-k}^* I^{-1}(\boldsymbol{\theta}_0) \mathbf{Q}_{n-k}^{*T}$$

are independent, Theorem follows from Lemma 2.4. \square

References

- M. Csörgő and L. Horváth, Nonparametric methods for changepoint problems, in: Handbook of Statistics, Vol. 7 (North-Holland, Amsterdam, 1988) pp. 403–425.
- D.A. Darling and P. Erdős, A limit theorem for the maximum of normalized sums of independent random variables, Duke Math. J. 23 (1956) 143–155.
- D. Ferger, Nonparametric changepoint-tests, Preprint (1991).
- E. Gombay and L. Horváth, Asymptotic distributions of maximum likelihood tests for change in the mean, Biometrika 77 (1990) 411–414.
- P. Haccou, E. Meelis and S. van de Geer, The likelihood ratio test for the change point problem for exponentially distributed random variables, Stochastic Process. Appl. 27 (1988) 121–139.
- D.M. Hawkins, Testing a sequence of observations for a shift in location, J. Amer. Statist. Assoc. 72 (1977) 180–186.
- L. Horváth, The limit distributions of the likelihood ratio and cumulative sum tests for a change in binomial probability, J. Multivariate Anal. 31 (1989) 148–159.

- L. Horváth, The maximum likelihood method for testing changes in the parameters of normal observations, *Ann. Statist.* 21 (1993) 671–680.
- I.A. Ibragimov and R.Z. Hašminskii, On the approximation of statistical estimators by sums of independent variables, *Dokl. Akad. Nauk SSSR* 210 (1973) 883–887.
- B. James, K.L. James and D. Siegmund, Tests for a change-point, *Biometrika* 74 (1987) 71–83.
- B. James, K.L. James and D. Siegmund, Asymptotic approximations for likelihood ratio tests and confidence regions for a change-point in the mean of a multivariate normal distribution, *Statist. Sinica* 2 (1992) 69–90.
- E. Lehmann, *Theory of Point Estimation* (Wadsworth, Pacific Grove, CA, 1991).
- A. Sen and M.S. Srivastava, On tests for detecting change in mean, *Ann. Statist.* 3 (1975a) 98–108.
- A. Sen and M.S. Srivastava, Some one-sided tests for change in level, *Technometrics* 17 (1975b) 61–64.
- R.J. Serfling, *Approximation Theorems of Mathematical Statistics* (Wiley, New York, 1980).
- M.S. Srivastava and K.J. Worsley, Likelihood ratio tests for a change in the multivariate mean, *J. Amer. Statist. Assoc.* 81 (1986) 199–204.
- K.J. Worsley, The power of likelihood ratio and cumulative sum tests for a change in a binomial probability, *Biometrika* 70 (1983) 455–464.
- K.J. Worsley, Confidence regions and tests for a change-point in a sequence of exponential family random variables, *Biometrika* 73 (1986a) 91–104.
- K.J. Worsley, Confidence regions and tests for a change-point in a sequence of exponential random variables, in: *Pacific Statistical Congress* (Elsevier, Amsterdam, 1986b) pp. 266–272.
- Y.-C. Yao and R.A. Davis, The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates, *Sankhyā Ser. A* 48 (1986) 339–353.