Journal of Multivariate Analysis **68**, 96–119 (1999)
Article ID jmva.1998.1780, available online at http://www.idealibrary.com on IDEAL®

Testing for Changes in Multivariate Dependent Observations with an Application to Temperature Changes*

Lajos Horváth

University of Utah

Piotr Kokoszka

University of Liverpool, Liverpool, United Kingdom

and

Josef Steinebach

Philipps-Universität, Marburg, Germany

Received February 13, 1997; revised April 22, 1998

We develop procedures for testing for changes in the mean of multivariate m-dependent stationary processes. Several test statistics are considered and corresponding limit theorems are derived. These include functional and Darling–Erdős type limit theorems. The tests are shown to be consistent under alternatives of abrupt and gradual changes in the mean. Finite sample performance is examined by means of a simulation study, and the procedures are applied to the analysis of the average monthly temperatures in Prague. © 1999 Academic Press

AMS 1990 subject classifications: 62G10, 62M07, 62M30, 62P99.
Key words and phrases: change-point tests; multivariate stationary processes; temperature changes.

1. INTRODUCTION

Our observations consist of n random vectors, $\mathbf{X}_1 = (X_{1,1}, ..., X_{1,d})$, $\mathbf{X}_2 = (X_{2,1}, ..., X_{2,d})$, ..., $\mathbf{X}_n = (X_{n,1}, ..., X_{n,d})$. Let $\boldsymbol{\mu}_i = E\mathbf{X}_i$. We assume that the observations $\mathbf{X}_1, ..., \mathbf{X}_n$ can be written as

$$\mathbf{X}_i = \boldsymbol{\mu}_i + \mathbf{e}_i, \qquad 1 \leqslant i \leqslant n.$$

* This work was partially supported by NATO Scientific and Environmental Division Grant CRG 960503.

0047-259X/99 $30.00

The errors $\mathbf{e}_1 = (e_{1,1}, ..., e_{1,d})$, $\mathbf{e}_2 = (e_{2,1}, ..., e_{2,d})$, ... satisfy the following conditions:

$$Ee_{i,j} = 0, \qquad 1 \leqslant i < \infty \quad \text{and} \quad 1 \leqslant j \leqslant d, \tag{1.1}$$

$$\{\mathbf{e}_i, 1 \leqslant i < \infty\} \qquad \text{is a stationary sequence,} \tag{1.2}$$

$$E\,\|\mathbf{e}_i\|^{\nu} < \infty \qquad \text{for some} \quad \nu > 2 \tag{1.3}$$

and

$$\begin{array}{l}\text{there is an integer } m \text{ such that } \sigma\{\mathbf{e}_i, 1 \leqslant i \leqslant k\} \text{ and } \sigma\{\mathbf{e}_i, l \leqslant i < \infty\} \\ \text{are independent for each } l \text{ and } k \text{ satisfying } l - k \geqslant m.\end{array} \tag{1.4}$$

We wish to test the null hypothesis

$$H_0\colon \boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots = \boldsymbol{\mu}_n$$

against the alternative that the means changed over time. We divide the data into two subsets, before and after $\mathbf{X}_k$, and define the sample means

$$\hat{\boldsymbol{\mu}}(k) = \frac{1}{k} \sum_{1 \leqslant i \leqslant k} \mathbf{X}_i, \qquad \hat{\boldsymbol{\mu}}(k) = (\hat{\mu}_1(k), ..., \hat{\mu}_d(k))$$

and

$$\tilde{\boldsymbol{\mu}}(k) = \frac{1}{n-k} \sum_{k < i \leqslant n} \mathbf{X}_i, \qquad \tilde{\boldsymbol{\mu}}(k) = (\tilde{\mu}_1(k), ..., \tilde{\mu}_d(k)).$$

If H_0 holds, then $\hat{\boldsymbol{\mu}}(k)$ as well as $\tilde{\boldsymbol{\mu}}(k)$ are unbiased estimators for the common mean, so the differences

$$\boldsymbol{\Delta}(k) = \hat{\boldsymbol{\mu}}(k) - \tilde{\boldsymbol{\mu}}(k)$$

are near $\mathbf{0}$ if H_0 holds, and are away from $\mathbf{0}$, if the mean changed. It is easy to see that if conditions (1.1)–(1.4) hold, then there is a matrix $\mathbf{D}$ such that

$$\lim_{n \to \infty} \frac{1}{n} E\left(\sum_{1 \leqslant i \leqslant n} \mathbf{e}_i\right)' \left(\sum_{1 \leqslant i \leqslant n} \mathbf{e}_i\right) = \mathbf{D}. \tag{1.5}$$

We assume that

$$\operatorname{rank} \mathbf{D} = d. \tag{1.6}$$

The central limit theorem yields that $\hat{\boldsymbol{\mu}}(k)$ is approximately normal $N(\boldsymbol{\mu}, (1/k)\,\mathbf{D})$ under H_0. Similarly, the distribution of $\tilde{\boldsymbol{\mu}}(k)$ can be approximated with normal $N(\boldsymbol{\mu}, (1/(n-k))\,\mathbf{D})$ distribution, where $\boldsymbol{\mu}$ stands for the common mean. By (1.4), the correlation between $\hat{\boldsymbol{\mu}}(k)$ and $\tilde{\boldsymbol{\mu}}(k)$ is negligible, and we get that $\boldsymbol{\Delta}(k)$ is approximately normal $N(\mathbf{0}, (1/k)\,\mathbf{D}+(1/(n-k))\,\mathbf{D}$ under H_0. Thus the functionals of

$$\begin{aligned} Z(k) &= \boldsymbol{\Delta}(k)\left(\frac{1}{k}\,\mathbf{D}+\frac{1}{n-k}\,\mathbf{D}\right)^{-1}\boldsymbol{\Delta}'(k) \\ &= \frac{k(n-k)}{n}\,\boldsymbol{\Delta}(k)\,\mathbf{D}^{-1}\,\boldsymbol{\Delta}'(k). \end{aligned} \tag{1.7}$$

can be used to test H_0. Let

$$T(k)=\frac{(k(n-k))^2}{n^3}\,\boldsymbol{\Delta}(k)\,\mathbf{D}^{-1}\,\boldsymbol{\Delta}'(k).$$

Our first result is the weak convergence of $\{T(nt), 0\leqslant t\leqslant 1\}$.

THEOREM 1.1. *We assume that H_0 holds. If* (1.1)–(1.4) *and* (1.6) *are satisfied, then*

$$T(nt) \xrightarrow{\mathscr{D}[0,1]} \sum_{1\leqslant i\leqslant d} B_i^2(t),$$

where $B_1, B_2, ..., B_d$ are independent Brownian bridges.

It follows that under the conditions of Theorem 1.1,

$$\max_{1\leqslant k<n} \frac{(k(n-k))^2}{n^3}\,\boldsymbol{\Delta}(k)\,\mathbf{D}^{-1}\,\boldsymbol{\Delta}'(k) \xrightarrow{\mathscr{D}} \sup_{0\leqslant t\leqslant 1}\sum_{1\leqslant i\leqslant d} B_i^2(t) \tag{1.8}$$

and

$$\frac{1}{n^4}\sum_{1\leqslant k<n}(k(n-k))^2\,\boldsymbol{\Delta}(k)\,\mathbf{D}^{-1}\,\boldsymbol{\Delta}'(k) \xrightarrow{\mathscr{D}} \sum_{1\leqslant i\leqslant d}\int_0^1 B_i^2(t)\,dt. \tag{1.9}$$

Kiefer (1959a, b) obtained the distribution function of the limit in (1.9).

Introducing weight functions we can increase the power of tests based on (1.8) and (1.9) against early or late changes. Let

$$Q=\{q\text{: } q \text{ is non-decreasing in a neighbourhood of 0, non-increasing in a neighbourhood of 1 and } \inf_{\varepsilon\leqslant t\leqslant 1-\varepsilon} q(t)>0 \text{ for all } 0<\varepsilon<1/2\}.$$

THEOREM 1.2. *We assume that H_0 holds and* (1.1)–(1.4) *and* (1.6) *are satisfied.*

(i) *If*

$$I_{0,1}(q, c) = \int_0^1 \frac{1}{t(1-t)} \exp\left(-\frac{cq^2(t)}{t(1-t)}\right) dt < \infty$$

for some $c > 0$, then

$$\frac{1}{n^2} \max_{1 \leqslant k < n} \frac{k(n-k)}{q^2(k/n)} Z(k) \xrightarrow{\mathscr{D}} \sup_{0 \leqslant t \leqslant 1} \left(\sum_{1 \leqslant i \leqslant d} B_i^2(t) \right) \Big/ q^2(t), \tag{1.10}$$

where $B_1, B_2, ..., B_d$ are independent Brownian bridges.

(ii) *Also,*

$$\frac{1}{n^2} \sum_{1 \leqslant k < n-1} Z(k)\, k(n-k) \int_{k/n}^{(k+1)/n} \frac{1}{t(1-t)}\, dt \xrightarrow{\mathscr{D}} \sum_{1 \leqslant i \leqslant d} \int_0^1 \frac{B_i^2(t)}{t(1-t)}\, dt. \tag{1.11}$$

Scholz and Stephens (1987) contains selected values of the distribution of the limiting random variable in (1.11).

According to our discussion, the natural statistic for testing H_0 is $\max_{1 \leqslant k < n} (k(n-k)/n)\, \mathbf{\Delta}(k)\, D^{-1} \mathbf{\Delta}'(k)$. However, this statistic is not covered by Theorems 1.1 and 1.2. We now consider its asymptotic distribution. Let

$$a(x) = (2 \log x)^{1/2}$$

and

$$b_d(x) = 2 \log x + \frac{d}{2} \log \log x - \log \Gamma(d/2),$$

where $\Gamma(t)$ is the Gamma function.

THEOREM 1.3. *We assume that H_0 holds. If* (1.1)–(1.4) *and* (1.6) *are satisfied, then*

$$\lim_{n \to \infty} P\{a(\log n) \max_{1 \leqslant k < n} Z^{1/2}(k) \leqslant t + b_d(\log n)\} = \exp(-2e^{-t}) \tag{1.12}$$

for all t.

Hušková (1990), Steinebach (1994), and Steinebach and Eastwood (1996) argued that the increments of $k\hat{\boldsymbol{\mu}}(k)$ and $(n-k)\,\tilde{\boldsymbol{\mu}}(k)$ provide better tests for H_0. Let

$$M_n = \max_{1 \leqslant k \leqslant n-h} h^{-1/2} \left\{ \left(\sum_{k \leqslant i \leqslant k+h} \mathbf{X}_i - h\hat{\boldsymbol{\mu}}(n) \right) \mathbf{D}^{-1} \left(\sum_{k \leqslant i \leqslant k+h} \mathbf{X}_i - h\hat{\boldsymbol{\mu}}(n) \right)' \right\}^{1/2}.$$

THEOREM 1.4. *We assume that* H_0 *holds. If* (1.1)–(1.4) *and* (1.6) *are satisfied and* $h = h(n) \to \infty$, $n/h \to \infty$, *and*

$$\limsup_{n \to \infty} \frac{n^{1/2-\kappa}(\log(n/h))^{1/2}}{h^{1/2}} < \infty \qquad \textit{with some} \quad 0 < \kappa < \kappa_0, \tag{1.13}$$

where $\kappa_0 = (1/(4(2+d)))\min(1, v-2)/(5+\min(1, v-2))$. *Then we have*

$$\lim_{n \to \infty} P\{a(n/h)\, M_n \leqslant t + b_d(n/h)\} = \exp(-2e^{-t})$$

for all t.

In this section we assumed that $\mathbf{D}$ is known. This is, however, rarely the case in applications. We show in Section 3 that the results of Section 1 remain valid if $\mathbf{D}$ is replaced by any estimator $\mathbf{D}_n$ satisfying (3.1).

The proofs of Theorems 1.1–1.4 are postponed until Section 4. First we discuss the consistency of our tests against two possible alternatives. Section 3 contains the applications of our results to the average temperatures in Prague.

2. TESTS UNDER ALTERNATIVES

First we assume that a sudden change occurred. Namely,

$$H_A^{(1)}\text{: there is an integer } k^*,\ 1 \leqslant k^* < n \text{ such that}$$

$$\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2 = \cdots \boldsymbol{\mu}_{k^*} \neq \boldsymbol{\mu}_{k^*+1} = \cdots = \boldsymbol{\mu}_n.$$

We now show that the tests based on (1.8)–(1.12) are consistent against $H_A^{(1)}$. Let $\tau = (\boldsymbol{\mu} - \boldsymbol{\lambda})\,\mathbf{D}(\boldsymbol{\mu} - \boldsymbol{\lambda})'$ with $\boldsymbol{\mu} = \boldsymbol{\mu}_1$ and $\boldsymbol{\lambda} = \boldsymbol{\mu}_n$. We note that $\boldsymbol{\mu}$ as well as $\boldsymbol{\lambda}$ may depend on n.

THEOREM 2.1. *We assume that* $H_A^{(1)}$ *holds and* (1.1)–(1.4) *and* (1.6) *are satisfied.*

(i) *If*

$$\tau k^* \to \infty \tag{2.1}$$

and

$$\tau(n-k^*) \to \infty, \tag{2.2}$$

then we have

$$\frac{1}{\tau(k^*(n-k^*))^2} \sum_{1 \leqslant k < n} k(n-k)\, Z(k) \xrightarrow{P} 1/3, \tag{2.3}$$

$$\frac{n}{\tau} \int_0^1 \frac{T(nt)}{t(1-t)}\, dt \Big/ \left\{ (n-k^*)^2 \left(-\log\left(1-\frac{k^*}{n}\right) - \frac{k^*}{n} \right) + k^{*2} \left(-\log \frac{k^*}{n} - \left(1-\frac{k^*}{n}\right) \right) \right\} \xrightarrow{P} 1 \tag{2.4}$$

and

$$\frac{1}{\tau n} \max_{1 \leqslant k < n} \left(\frac{k}{n}\left(1-\frac{k}{n}\right)\right)^{1-\alpha} Z(k) \Big/ \left(\frac{k^*}{n}\left(1-\frac{k^*}{n}\right)\right)^{2-\alpha} \xrightarrow{P} 1 \tag{2.5}$$

for all $0 \leqslant \alpha < 1$.

(ii) *If*

$$\tau k^*/\log\log n \to \infty \tag{2.6}$$

and

$$\tau(n-k^*)/\log\log n \to \infty, \tag{2.7}$$

then we have

$$\frac{1}{\tau n} \max_{1 \leqslant k < n} Z(k) \Big/ \left(\frac{k^*}{n}\left(1-\frac{k^*}{n}\right)\right) \xrightarrow{P} 1.$$

We can also estimate k^* with

$$\hat{k}(\alpha) = \min\{k\colon \max_{1 \leqslant i < n} (i(n-i))^{1-\alpha} Z(i) = (k(n-k))^{1-\alpha} Z(k)\}, \tag{2.8}$$

$0 \leqslant \alpha < 1$.

THEOREM 2.2. *We assume that $H_A^{(1)}$ holds,* (1.1)–(1.4) *and* (1.6) *are satisfied and*

$$k^* = [n\theta] \qquad \text{with some} \quad 0<\theta<1. \tag{2.9}$$

(i) *If* $0<\alpha<1$ *and* $\tau n \to \infty$, *then we have*

$$\hat{k}(\alpha)/n \xrightarrow{P} \theta. \tag{2.10}$$

(ii) *If* $\tau n/\log\log n \to \infty$, *then we have*

$$\hat{k}(1)/n \xrightarrow{P} \theta \tag{2.11}$$

We note that under the conditions of Theorem 2.2 we have

$$\frac{1}{n^3}\sum_{1\leqslant k<n} k(n-k)\,Z(k) \xrightarrow{P} \infty, \tag{2.12}$$

$$\frac{1}{n^2}\sum_{1\leqslant k<n-1} k(n-k)\,Z(k)\int_{k/n}^{(k+1)/n}\frac{1}{t(1-t)}\,dt \xrightarrow{P} \infty \tag{2.13}$$

and

$$a(\log n)\max_{1\leqslant k<n} Z^{1/2}(k) - b_d(\log n) \xrightarrow{P} \infty, \tag{2.14}$$

which give the asymptotic consistency of these tests against $H_A^{(1)}$.

Next we consider the case when the mean gradually increases. We assume

$$\begin{aligned} H_A^{(2)}\colon &\text{ there are integers } k^* \text{ and } l^*,\ 1\leqslant k^*<l^*<n\\ &\text{such that } \boldsymbol{\mu}_1=\cdots=\boldsymbol{\mu}_{k^*}\neq\boldsymbol{\mu}_{l^*+1}=\cdots=\boldsymbol{\mu}_n,\\ &\boldsymbol{\mu}_{k^*+i}=\boldsymbol{\mu}_{k^*}-\frac{\boldsymbol{\mu}_{k^*}-\boldsymbol{\mu}_n}{l^*-k^*}\,i,\ 1\leqslant i\leqslant l^*-k^*. \end{aligned}$$

THEOREM 2.3. *We assume that $H_A^{(2)}$ holds,* (1.1)–(1.4) *and* (1.6) *are satisfied, and*

$$0<\liminf_{n\to\infty} k^*/n \leqslant \limsup_{n\to\infty} l^*/n<1.$$

(i) *If* $\tau n\to\infty$, *then* (2.12) *and* (2.13) *hold.*

(ii) *If* $\tau n/\log\log n\to\infty$, *then* (2.14) *holds.*

We can also estimate the time of change. We assume that the period of gradual change is smaller than the time period when the observations were collected. Namely,

$$l^* - k^* = o(n), \qquad \text{as} \quad n \to \infty. \tag{2.15}$$

THEOREM 2.4. *We assume that $H_A^{(2)}$ holds and* (1.1)–(1.4), (1.6), (2.9), *and* (2.15) *are satisfied.*

(i) *If $\tau n \to \infty$, then* (2.10) *holds.*

(ii) *If $\tau n/\log\log n \to \infty$, then* (2.11) *holds.*

According to Theorems 2.2 and 2.4, the estimator in (2.8) can be used to get information about the time of change. The results say only that the changes occurred in a neighbourhood of $\hat{k}$ with large probability.

The consistency of tests based on M_n can be discussed along the lines of Theorems 2.1 and 2.2. For example, if $H_A^{(1)}$ and (2.9) hold and

$$h\tau/\log(n/h) \to \infty, \tag{2.16}$$

then we have

$$a(n/h)\, M_n - b_d(n/h) \xrightarrow{P} \infty. \tag{2.17}$$

If $H_A^{(1)}$, (2.9) and (2.15) are satisfied, then (2.16) implies (2.17).

We pointed out at the end of Section 1 that $\mathbf{D}$ is rarely known and must be estimated from the sample. It is clear from the proofs that the results of Section 2 remain valid if $\mathbf{D}$ is replaced by an estimator $\mathbf{D}_n$ which is bounded in probability under the alternative.

3. SIMULATIONS AND A STUDY OF TEMPERATURE CHANGES IN PRAGUE

We describe in this section the results of a small simulation study which illustrates some aspects of the finite sample behaviour of the procedures developed in the previous sections as well as the analysis of average monthly temperatures in Prague.

In order to be able to apply the results of the previous sections we must replace the matrix $\mathbf{D}$ by an appropriate estimator $\mathbf{D}_n$ based on the sample $\{\mathbf{X}_i, 1 \leqslant i \leqslant n\}$. Suppose $\mathbf{D}_n$ is an estimator satisfying

$$\|\mathbf{D}_n - \mathbf{D}\| = o_P((\log\log n)^{-1/2}). \tag{3.1}$$

If relation (3.1) holds, then all results of Section 1 remain valid with $Z(k)$ and $T(k)$ replaced by

$$\hat{Z}(k) = \frac{k(n-k)}{n} \mathbf{\Delta}(k) \, \mathbf{D}_n^{-1} \, \mathbf{\Delta}(k)', \qquad 1 \leqslant k < n \tag{3.2}$$

and

$$\hat{T}(k) = \frac{(k(n-k))^2}{n^3} \mathbf{\Delta}(k) \, \mathbf{D}_n^{-1} \, \mathbf{\Delta}(k)', \qquad 1 \leqslant k < n. \tag{3.3.}$$

In particular, under H_0 and the assumptions of Section 1, the following limit theorems hold:

$$n^{-3} \sum_{1 \leqslant k < n} k(n-k) \, \hat{Z}(k) \xrightarrow{\mathscr{D}} \sum_{1 \leqslant i \leqslant d} \int_0^1 B_i^2(t) \, dt, \tag{3.4}$$

$$n^{-2} \sum_{1 \leqslant k < n-1} k(n-k) \, \hat{Z}(k) \int_{k/n}^{(k+1)/n} \frac{1}{t(1-t)} \xrightarrow{\mathscr{D}} \sum_{1 \leqslant i \leqslant d} \int_0^1 \frac{B_i^2(t)}{t(1-t)} \, dt \tag{3.5}$$

and

$$\lim_{n \to \infty} P\{a(\log n) \max_{1 \leqslant k < n} \hat{Z}^{1/2}(k) \leqslant t + b_d(\log n)\} = \exp(-2e^{-t}), \tag{3.6}$$

for all t.

If (1.4) holds, then we can estimate $\mathbf{D}$ with

$$\begin{aligned} \mathbf{D}_n = \mathbf{D}_n(m) = &\frac{1}{n} \sum_{1 \leqslant i \leqslant n} (\mathbf{X}_i - \hat{\boldsymbol{\mu}}(n))' \, (\mathbf{X}_i - \hat{\boldsymbol{\mu}}(n)) \\ &+ \frac{2}{n} \sum_{1 \leqslant j \leqslant m} \; \sum_{1 \leqslant i \leqslant n-j} (\mathbf{X}_i - \hat{\boldsymbol{\mu}}(n))' \, (\mathbf{X}_{i+j} - \hat{\boldsymbol{\mu}}(n)), \end{aligned}$$

if $m \geqslant 1$, and

$$\mathbf{D}_n(0) = \frac{1}{n} \sum_{1 \leqslant i \leqslant n} (\mathbf{X}_i - \hat{\boldsymbol{\mu}}(n))' \, (\mathbf{X}_i - \hat{\boldsymbol{\mu}}(n)).$$

Using the Marcinkiewicz–Zygmund law of large numbers, it can be verified that (3.1) holds under H_0.

Under the alternatives discussed in section 2, $\mathbf{D}_n \xrightarrow{P} \hat{\mathbf{D}}$, with some matrix $\hat{\mathbf{D}}$. Theorems 2.1 and 2.2 remain true if $Z(k)$, $T(k)$ and τ are replaced with $\hat{Z}(k)$, $\hat{T}(k)$ and $\hat{\tau} = (\boldsymbol{\mu} - \boldsymbol{\lambda}) \, \hat{\mathbf{D}}(\boldsymbol{\mu} - \boldsymbol{\lambda})'$. In particular, if $\hat{\tau} k^* \to \infty$, $\hat{\tau}(n-k^*) \to \infty$ and (1.1)–(1.6), (2.9) are satisfied, then under either $H_A^{(1)}$ or $H_A^{(2)}$ we have

$$n^{-3} \sum_{1 \leqslant k < n} k(n-k)\, \hat{Z}(k) \xrightarrow{P} \infty, \tag{3.7}$$

$$n^{-2} \sum_{1 \leqslant k < n-1} k(n-k)\, \hat{Z}(k) \int_{k/n}^{(k+1)/n} \frac{1}{t(1-t)}\, dt \xrightarrow{P} \infty, \tag{3.8}$$

$$a(\log n) \max_{1 \leqslant k < n} \hat{Z}^{1/2}(k) - b_d(\log n) \xrightarrow{P} \infty. \tag{3.9}$$

For the temperature data considered below, the second term in the definition of $\mathbf{D}_n(m)$ is not significantly different from zero for lags $m = 1$, 2, 3, 4, 5, 6, 7 (confidence intervals valid under the assumption of normal errors were used). Therefore, in the simulations and computations presented in the sequel we use $\mathbf{D}_n(0)$.

We focus on (3.4) and (3.6), as they represent two distinct results: relation (3.4) is a functional limit theorem involving a weighted sum of the $\hat{Z}(k)$ whereas relation (3.6) is a Darling–Erdős type limit theorem based on the maximum of the $\hat{Z}(k)$.

First we consider inference based on relation (3.4).

Table I gives selected values of the cumulative distribution function of the right hand side of (3.4) for $d = 12$. The values were obtained using formula (4.4) (with h replaced by 12, our d is Kiefer's h) of Kiefer (1959b). For $d = 2$ and $d = 4$, the only even values of d included in Table 3 of Kiefer (1959b), our values agreed with those of Kiefer (1959b) up to the accuracy of Kiefer's table (six digits after the decimal point).

It is worthwhile noting that there are two misprints in formula (4.4) of Kiefer (1959b), the non-obvious one being the omission of the factor $2^{(h-2)/4}$ in its right hand side.

The asymptotic values in Table I are fairly good approximations of the actual critical values. To illustrate this point, we constructed a sequence of 80 identical years as follows: we simulated 80 independent 12-dimensional normal random vectors with independent standard normal components. Next, we generated a sample of 250 such sequences of 80 years and for each of the 250 sequences we evaluated the left hand side of (3.4). The percentiles of these 250 values are given in Table II.

TABLE I

Asymptotic distribution function in (3.4), $d = 12$

2.1000	2.2000	2.3000	2.4000	2.5000	2.6000	2.7000
0.6226	0.6892	0.7477	0.7979	0.8401	0.8750	0.9032
2.8000	2.9000	3.0000	3.1000	3.2000	3.3000	3.4000
0.9258	0.9437	0.9576	0.9683	0.9765	0.9827	0.9874
3.5000	3.6000	3.7000	3.8000	3.9000	4.0000	4.1000
0.9908	0.9933	0.9952	0.9965	0.9975	0.9983	0.9988

TABLE II

Percentiles of the simulated values of (3.4), $n = 80$, $d = 12$

80.0%	90.0%	95.0%	97.5%	99.0%
2.38	2.59	2.89	3.07	3.21

To get some idea of how sensitive the tests based on (3.4) and (3.6) are, we constructed three types of sequences that violate H_0. All of them consist of 80 years with independent standard normal components, but the means of components of certain years are changed, the change being the same for all components of a given year. Thus these alternatives do not allow a different change in, say, the January and the February components.

We now describe the alternatives in detail

1. *Abrupt change*: the first 40 years have mean zero, the last 40 have all monthly means increased by the same value.

2. *Gradual change*: the first 40 years have mean zero, the last 30 years have all monthly means increased by the same value, in the intermediate years the temperature increases linearly.

3. *Creeping change*: the temperature increases linearly from the first to the 80th years.

The cumulative increases considered are 0.25, 0.50 and 0.75. Table III gives estimates of the power of the test based on (3.4) for the 9 alternatives. Each entry is based on 100 simulations. We used the 95% critical value from Table II.

Now we consider procedures based on (3.6).

The most striking difference between (3.4) and (3.6) is that the asymptotic critical values based on the right hand side of (3.6) are very poor approximations of actual critical values for moderate sample sizes. For example, for $n = 80$ and $d = 12$ the asymptotic 95% critical value is so small that it practically always leads to the rejection of H_0 for the sequence of iid years. The discrepancy between the asymptotic and simulated critical values for (3.6)

TABLE III

Power of the test based on (3.4)

Alternative	0.25	0.50	0.75
Abrupt change	0.43	0.99	1.00
Gradual change	0.43	0.99	1.00
Creeping change	0.17	0.82	1.00

TABLE IV

Critical values for the test based on (3.6)

d	2	4	6	8	10	12
Asymptotic 95% critical value	4.08	4.31	4.13	3.71	3.14	2.43
Simulated 95% critical value	3.42	3.97	4.40	5.04	4.95	5.34

is illustrated in Table IV. The simulated critical values are based on 100 simulations for each d.

Table V was obtained in the same way as Table III. The null hypothesis was rejected if $\max_{1 \leqslant k < n} \hat{Z}^{1/2}(k) > 5.34$.

Conclusions. For the tests based on (3.4), it is acceptable to use asymptotic critical values instead of the simulated ones. Moreover, for any observed value of the test statistic the asymptotic P-value can readily be evaluated using formula (4.4) of Kiefer (1959). For tests based on (3.6) simulated critical values must be used. Tests based on (3.4) appear to be more powerful than those based on (3.6) when the change does not occur either very early or very late. Detecting a slow change taking place over the whole period when observations were taken is more difficult than detecting a fairly rapid change.

In the simulation study our main concern was the effect of d on the applicability of the limit results to finite samples. We generated independent observations since the dependence between annual observation vectors considered below is very weak. In case of dependent observations the rate of convergence in limit theorems of Sections 1–3 is likely to be slower. Davis, Huang and Yao (1995) used statistics similar to (3.6) in case of autoregressive process and observed that the dependence between the observations reduced the accuracy of the approximation in (3.6). Antoch, Hušková and Prašková (1997) studied the behaviour of test statistics and estimators for a change in the mean of a univariate linear process with first moment summable coefficients. Kokoszka and Horváth (1997) considered long-range dependent errors whereas Kokoszka and Leipus (1998) studied also non-linear processes.

TABLE V

Power of the test based on (3.6)

Alternative	0.25	0.50	0.75
Abrupt change	0.33	0.98	1.00
Gradual change	0.30	0.99	1.00
Creeping change	0.14	0.54	0.95

We now describe the results of the analysis of monthly temperatures in Prague by means of the tools we have developed. One of the goals of our study is to determine whether the data provide evidence for the global warming theory.

Our data set consists of average monthly temperatures in Prague from 1775 to 1989. The data have been compiled by Daniela Jarušková.

The data have been analysed as follows: Using the asymptotic P-value based on (3.4), and referred to in the sequel as *the P-value*, we determine whether a shift in mean has taken place. If the P-value indicates no change, we confirm this result using simulated critical values based on (3.4) and also on (3.6). If the P-value indicates a change, we estimate it using the estimator $\hat{k}(1)$, which is the argmax of the $\hat{Z}(k)$. This divides the data into two subsets of years. The whole procedure is then repeated for each subsets until periods of constant mean are obtained.

Figure 1 shows the graphs of the average monthly temperatures (dotted) and $\hat{Z}(k)$ (continuous).

Using the above procedure we obtained the following segmentation:

- *Years* 1775–1989: P-value $= 0.00005$ indicates change. The estimated value of the time of change is 1835.

- *Years* 1775–1835: P-value $= 0.88$ indicates absense of change. The observed value of the LHS of (3.4) is 1.43 and lies below the simulated 25th percentile which is 1.66. We conclude that there was no significant change in mean from 1775 tot 1835.

- *Years* 1836–1989: P-value $= 1.8 \times 10^{-6}$ indicates change. The estimated time of change is 1893.

- *Years* 1836–1893: P-value $= 0.54$ indicates absense of change. The observed value of the left hand side of (3.4) is 1.89 and lies below the

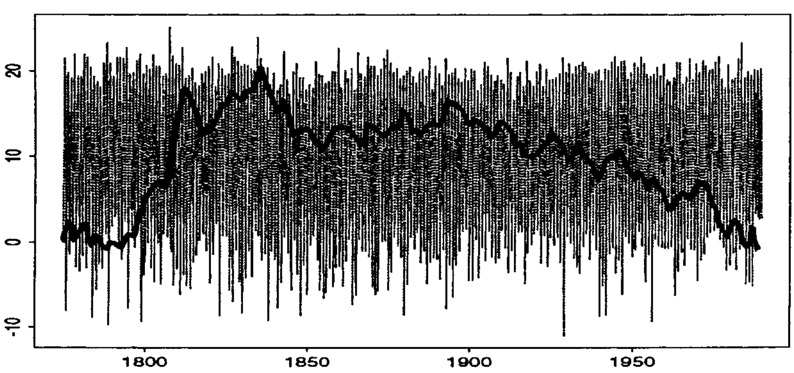

FIG. 1. The average monthly temperatures and $\hat{Z}(k)$.

simulated 50th percentile which is 2.06. We conclude that there was no significant change in mean form 1836 to 1893.

- *Years* 1894–1989: P-value = 0.006 indicates change. The estimated time of the change is 1927.
- *Years* 1894–1927: P-value = 0.18 indicates absense of change but is not very large. The same picture emerges when we use simulated critical values: The observed value of the left hand side of (3.4) is 2.43 and lies between the simulated 80th and 90th percentiles which are 2.42, 2.61, respectively. The observed value of max $\hat{Z}^{1/2}(k)$ is 4.72 and also lies between the simulated 80th and 90th percentiles which are 4.70 and 4.81, respectively. The above values indicate a possibility of a "creeping change", as studied in the previous subsection. This alternative is also reinforced by direct examination of Fig. 1.
- *Years* 1928–1989: P-value = 0.50 indicates absense of change. The observed value of he LHS of (3.4) is 1.94 and lies below the simulated 50th percentile which is 1.97. We conclude that there was no significant change in mean from 1928 to 1989.

The segmentation gave the following subsets: 1775–1835, 1835–1893, 1894–1927, 1928–1989. The average temperatures in these subsets are 9.79, 8.93, 9.26 and 9.80, respectively.

Our analysis strongly supports changes in the climate around 1835 and 1893. A warm period ended in 1835 and the climate became cooler between 1835 and 1893. The next warm period started in 1893 temperature was increasing slightly until 1927. There is no evidence of change in the data after 1927.

4. PROOFS OF THEOREMS 1.1–1.4

Throughout this section we assume that H_0 holds. We can and shall assume without loss of generality that $\boldsymbol{\mu}_i = \mathbf{0}$. Let

$$\mathbf{S}(k) = \sum_{1 \leqslant i \leqslant k} \mathbf{e}_i.$$

LEMMA 4.1. *We assume that* (1.1)–(1.4) *and* (1.6) *hold. We can find a Gaussian process* $\mathbf{G}(t)$ *with* $E\mathbf{G}(t) = \mathbf{0}$ *and* $E\mathbf{G}'(t)\,\mathbf{G}(t) = \mathbf{D}\min(t, s)$ *such that*

$$\|\mathbf{S}(k) - \mathbf{G}(k)\| = o(k^{1/2-\kappa}) \qquad \text{a.s.} \tag{4.1}$$

with any $0 < \kappa < \kappa_0$, *where* $\kappa_0 = (1/(4(2+d)))\min(1, \nu-2)/(5+\min(1, \nu-2))$.

Proof. We apply Theorem 1 of Eberlein (1986). Let $\mathscr{F}_k=\sigma(\mathbf{e}_1, ..., \mathbf{e}_k)$. By (1.1) and (1.4) we have

$$\left\| E\left(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \,\middle|\, \mathscr{F}_k \right)\right\| \leqslant \sum_{k+1\leqslant i\leqslant k+m} E\,\|\mathbf{e}_i\|, \tag{4.2}$$

and therefore (1.2) and (1.3) yield

$$\sup_{1\leqslant k,\, n<\infty} E\left\| E\left(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \,\middle|\, \mathscr{F}_k \right)\right\| < \infty. \tag{4.3}$$

Similarly to (4.3) one can show that for all $1\leqslant k,\ n<\infty$,

$$\begin{aligned} E\Bigg\| E\Bigg\{ \Bigg(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \Bigg)' \Bigg(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \Bigg) \Bigg| \mathscr{F}_k \\ - E\Bigg(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \Bigg)' \Bigg(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \Bigg)\Bigg\| < C_1 \end{aligned} \tag{4.4}$$

and

$$\left\| E\left(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \right)' \left(\sum_{k+1\leqslant i\leqslant n+k} \mathbf{e}_i \right) - n\mathbf{D} \right\| < C_2 \tag{4.5}$$

with some constants C_1 and C_2. Thus we established that all conditions of Theorem 1 of Eberlein (1986) are satisfied. Checking the proofs in Eberlein (1986), one can get the upper bound for the rate of approximation.

Let

$$\mathbf{U}_n(t)=n^{-1/2}(\mathbf{S}(nt)-t\mathbf{S}(n)), \qquad 0\leqslant t\leqslant 1. \tag{4.6}$$

LEMMA 4.2. *We assume that* (1.1)–(1.4) *and* (1.6) *hold. We can find a sequence of Gaussian processes* $\{\mathbf{V}_n(t), 0\leqslant t\leqslant 1\}$ *with* $E\mathbf{V}_n(t)=\mathbf{0}$ *and* $E\mathbf{V}_n'(t)\,\mathbf{V}_n(s)=\mathbf{D}\{\min(t, s)-ts\}$ *such that*

$$n^{\alpha} \sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} \|\mathbf{U}_n(t)-\mathbf{V}_n(t)\|/(t(1-t))^{1/2-\alpha}=O_P(1) \tag{4.7}$$

for all $0\leqslant\alpha<\kappa_0$

Proof. First we define

$$\mathbf{S}^*(k)=\begin{cases} \mathbf{S}(k), & 1\leqslant k<(n-m)/2 \\ \mathbf{S}((n-m)/2), & (n-m)/2\leqslant k\leqslant (n+m)/2 \\ \mathbf{S}((n-m)/2)+\mathbf{S}(k)-\mathbf{S}((n+m)/2), & (n+m)/2<k\leqslant n \end{cases}$$

and

$$\mathbf{U}_n^*(t) = n^{-1/2}(\mathbf{S}^*(nt) - t\mathbf{S}^*(n)), \qquad 0 \leqslant t \leqslant 1.$$

It is easy to see that

$$\max_{1 \leqslant k \leqslant n} \|\mathbf{S}(k) - \mathbf{S}^*(k)\| = O_P(1)$$

and therefore

$$n^{1/2} \max_{1/(n+1) \leqslant t \leqslant n/(n+1)} \|\mathbf{U}_n(t) - \mathbf{U}_n^*(t)\| / (t(1-t))^{1/2-\alpha} = O_P(1) \tag{4.8}$$

for all α. By (1.4) and Lemma 4.1 we can find $\{\mathbf{G}^{(1)}(x), 0 \leqslant x \leqslant n/2\}$ and $\{\mathbf{G}^{(2)}(x)\ 0 \leqslant x \leqslant n/2\}$, two independent copies of $\mathbf{G}$ such that

$$\max_{1 \leqslant k \leqslant n/2} \sup_{k-1/2 \leqslant x < k+1/2} \|\mathbf{S}^*(k) - \mathbf{G}^{(1)}(x)\| / x^{1/2-\kappa} = O_P(1) \tag{4.9}$$

and

$$\max_{n/2 \leqslant k < n} \sup_{k-1/2 \leqslant x < k+1/2} \|\mathbf{S}^*(n) - \mathbf{S}^*(k) - \mathbf{G}^{(2)}(n-x)\| / (n-x)^{1/2-\kappa} = O_P(1) \tag{4.10}$$

for any $0 < \kappa < \kappa_0$. Next we introduce

$$\mathbf{V}_n(t) = \begin{cases} n^{-1/2}(\mathbf{G}^{(1)}(nt) - t(\mathbf{G}^{(1)}(n/2) + \mathbf{G}^{(2)}(n/2)), & 0 \leqslant t \leqslant 1/2 \\ n^{-1/2}(-\mathbf{G}^{(2)}(n-nt) + (1-t)(\mathbf{G}^{(1)}(n/2) + \mathbf{G}^{(2)}(n/2)), & 1/2 \leqslant t \leqslant 1. \end{cases}$$

It follows from (4.9) and (4.10) that

$$n^{\alpha} \sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \|\mathbf{U}_n^*(t) - \mathbf{V}_n(t)\| / (t(1-t))^{1/2-\alpha} = O_P(1) \tag{4.11}$$

for any $0 \leqslant \alpha < \kappa_0$. Thus (4.7) follows from (4.8) and (4.11). The construction yields that $\{\mathbf{V}_n(t), 0 \leqslant t \leqslant 1\}$ is Gaussian and elementary calculations give the covariance structure.

LEMMA 4.3. *We assume that* (1.1)–(1.4) *and* (1.6) *hold. We can find independent Brownian bridges* $\{B_{n,1}(t), 0 \leqslant t \leqslant 1\}, ..., \{B_{n,d}(t), 0 \leqslant t \leqslant 1\}$ *such that*

$$n^{2\alpha} \sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \left| \mathbf{U}_n(t)\, \mathbf{D}^{-1} \mathbf{U}_n'(t) - \sum_{1 \leqslant i \leqslant d} B_{n,i}^2(t) \right| \Big/ (t(1-t))^{1-2\alpha}$$
$$= O_P(1)$$

for all $0 \leqslant \alpha < \kappa_0/2$.

Proof. Elementary arguments yield

$$\begin{aligned}\|\mathbf{U}_n(t)\,\mathbf{D}^{-1}\mathbf{U}_n'(t)-\mathbf{V}_n(t)\,\mathbf{D}^{-1}\mathbf{V}_n'(t)\|&\\ \leqslant c\{\|\mathbf{U}_n(t)-\mathbf{V}_n(t)\|^2+\|\mathbf{V}_n(t)\|\,\|\mathbf{U}_n(t)-\mathbf{V}_n(t)\|\}&\end{aligned}\tag{4.12}$$

with some constant c. Lemma 4.2 implies that

$$n^{2\alpha}\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\|\mathbf{U}_n(t)-\mathbf{V}_n(t)\|^2/(t(1-t))^{1-2\alpha}=O_P(1).\tag{4.13}$$

Since the coordinates of $\mathbf{V}_n(t)$ are Brownian bridges multiplied with constants, we have that

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\|\mathbf{V}_n(t)\|/(t(1-t))^{1/2+\varepsilon}=O_P(n^{\varepsilon})\tag{4.14}$$

for any $\varepsilon>0$ (cf. Csörgő and Horváth (1993), p. 259). Using again Lemma 4.2 with (4.14) we obtain

$$\begin{aligned}&n^{2\alpha}\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\|\mathbf{V}_n(t)\|\,\|\mathbf{U}_n(t)-\mathbf{V}_n(t)\|/(t(1-t))^{1-2\alpha}\\ &\quad\leqslant n^{2\alpha}\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\|\mathbf{V}_n(t)\|/(t(1-t))^{1/2+\varepsilon}\\ &\qquad\times\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\|\mathbf{U}_n(t)-\mathbf{V}_n(t)\|/(t(1-t))^{1/2-\varepsilon-2\alpha}\\ &\quad=n^{2\alpha}O_P(n^{\varepsilon})\,O_P(n^{-(\varepsilon+2\alpha)})\\ &\quad=O_P(1),\end{aligned}$$

where $0<\varepsilon<\kappa_0-2\alpha$. Observing that

$$\{\mathbf{V}_n(t)\,\mathbf{D}^{-1}\mathbf{V}_n'(t),\,0\leqslant t\leqslant 1\}\overset{\mathscr{D}}{=}\left\{\sum_{1\leqslant i\leqslant d}B_i^2(t),\,0\leqslant t\leqslant 1\right\},$$

where $B_1, B_2, ..., B_d$ are independent Brownian bridges, the proof of Lemma 4.3 is complete.

Proof of Theorem 1.1. Since H_0 holds, we have

$$T(nt)=\mathbf{U}_n(t)\,\mathbf{D}^{-1}\mathbf{U}_n'(t),\qquad 0\leqslant t\leqslant 1,\tag{4.15}$$

and therefore Lemma 4.3 yields

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\left|T(nt)-\sum_{1\leqslant i\leqslant d}B_{n,i}^2(t)\right|=o_P(1).\tag{4.16}$$

It is easy to see that

$$\sup_{0<t\leqslant 1/(n+1)} |T(nt)| = O_P(1/n) \tag{4.17}$$

and

$$\sup_{n/(n+1)\leqslant t\leqslant 1} |T(nt)| = O_P(1/n). \tag{4.18}$$

The scale transformation of the Wiener process yields

$$\sup_{0<t\leqslant 1/(n+1)} \left| \sum_{1\leqslant i\leqslant d} B^2_{n,i}(t) \right| = O_P(1/n) \tag{4.19}$$

and

$$\sup_{n/(n+1)\leqslant t\leqslant 1} \left| \sum_{1\leqslant i\leqslant d} B^2_{n,i}(t) \right| = O_P(1/n). \tag{4.20}$$

Putting together (4.15)–(4.20) we get Theorem 1.1.

Proof of Theorem 1.2. We use the construction of Lemma 4.3. It follows from (4.16) that

$$\sup_{\lambda\leqslant t\leqslant 1-\lambda} \left| T(nt) - \sum_{1\leqslant i\leqslant d} B^2_{n,i}(t) \right| \Big/ q^2(t) = o_P(1) \tag{4.21}$$

for all $0<\lambda<1/2$. If $I_{0,1}(q, c)<\infty$ for some $c>0$, then

$$\lim_{t\to 0} q(t)/t^{1/2} = \infty \tag{4.22}$$

and

$$\lim_{t\to 1} q(t)/(1-t)^{1/2} = \infty \tag{4.23}$$

(cf. Csörgő and Horváth (1993), p. 180). Lemma 4.3 with $\alpha=0$ yields

$$\begin{aligned}
&\sup_{1/(n+1)\leqslant t\leqslant \lambda} \left| T(nt) - \sum_{1\leqslant i\leqslant d} B^2_{n,i}(t) \right| \Big/ q^2(t) \\
&\quad\leqslant \sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} \left| T(nt) - \sum_{1\leqslant i\leqslant d} B^2_{n,i}(t) \right| \Big/ (t(1-t)) \sup_{0<t\leqslant\lambda} t(1-t)/q^2(t),
\end{aligned}$$

and therefore by (4.22) we have

$$\lim_{\lambda\to 0}\limsup_{n\to\infty} P\left\{\sup_{1/(n+1)\leqslant t\leqslant\lambda}\left|T(nt)-\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)\right|\Big/ q^2(t)>\varepsilon\right\}=0 \tag{4.24}$$

for all $\varepsilon>0$. Similarly,

$$\lim_{\lambda\to 1}\limsup_{n\to\infty} P\left\{\sup_{\lambda\leqslant t\leqslant n/(n+1)}\left|T(nt)-\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)\right|\Big/ q^2(t)>\varepsilon\right\}=0 \tag{4.25}$$

for all $\varepsilon>0$. From (4.21), (4.24), and (4.25) we get

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\left|T(nt)-\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)\right|\Big/ q^2(t)=o_P(1). \tag{4.26}$$

Next we note that

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)/q^2(t)\xrightarrow{\mathscr{D}}\sup_{0<t<1}\sum_{1\leqslant i\leqslant d} B_i^2(t)/q^2(t) \tag{4.27}$$

and the limit is finite with probability one (cf. Csörgő and Horváth (1993) p. 189). Now (1.10) follows from (4.26) and (4.27).

To prove (ii) of Theorem 1.2 it is enough to show that

$$\int_{1/(n+1)}^{n/(n+1)} T(nt)/(t(1-t))\,dt\xrightarrow{\mathscr{D}}\sum_{1\leqslant i\leqslant d}\int_0^1\frac{B_i^2(t)}{t(1-t)}\,dt.$$

Let $0<\kappa<\kappa_0/2$. Using Lemma 4.3 and (4.15) we obtain

$$\begin{aligned}
&\int_{1/(n+1)}^{n/(n+1)}\left|T(nt)-\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)\right|\Big/(t(1-t))\,dt\\
&\quad\leqslant\int_{1/(n+1)}^{n/(n+1)}\frac{|T(nt)-\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)|}{(t(1-t))^{1-\varepsilon}}(t(1-t))^{-\varepsilon}\,dt\\
&\quad\leqslant\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)}\left|T(nt)-\sum_{1\leqslant i\leqslant d} B_{n,i}^2(t)\right|\Big/(t(1-t))^{1-\varepsilon}\\
&\qquad\times\int_{1/(n+1)}^{n/(n+1)}(s(1-s))^{-\varepsilon}\,ds\\
&\quad=O_P(n^{-\varepsilon}).
\end{aligned}$$

Since

$$\int_{1/(n+1)}^{n/(n+1)} \sum_{1 \leqslant i \leqslant d} B_{n,i}^2(t)/(t(1-t))\, dt \xrightarrow{\mathscr{D}} \sum_{1 \leqslant i \leqslant d} \int_0^1 \frac{B_i^2(t)}{t(1-t)}\, dt,$$

the proof of (1.11) is complete.

Proof of Theorem 1.3. It is enough to show that

$$\lim_{n \to \infty} P\{a(\log n) \sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} T^{1/2}(nt)/(t(1-t))^{1/2} \leqslant x + b_d(\log n)\}$$
$$= \exp(-2e^{-x}) \tag{4.28}$$

for all x. Horváth (1993) showed

$$\sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \left(\sum_{1 \leqslant i \leqslant d} B_{n,i}^2(t)/(t(1-t)) \right)^{1/2} \Big/ (2 \log\log n)^{1/2} \xrightarrow{P} 1, \tag{4.29}$$

$$\sup_{1/(n+1) \leqslant t \leqslant (\log n)/n} \left(\sum_{1 \leqslant i \leqslant d} B_{n,i}^2(t)/(t(1-t)) \right)^{1/2} = O_P((\log\log\log n)^{1/2}), \tag{4.30}$$

and similarly

$$\sup_{1-(\log n)/n \leqslant t \leqslant n/(n+1)} \left(\sum_{1 \leqslant i \leqslant d} B_{n,i}^2(t)/(t(1-t)) \right)^{1/2} = O_P((\log\log\log n)^{1/2}). \tag{4.31}$$

If $\xi = \xi(n)$ satisfies

$$\sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} \sum_{1 \leqslant i \leqslant d} B_{n,i}^2(t)/(t(1-t)) = \sum_{1 \leqslant i \leqslant d} B_{n,i}^2(\xi)/(\xi(1-\xi)),$$

then by (4.29)–(4.31) we have

$$\lim_{n \to \infty} P\{(\log n)/n \leqslant \xi \leqslant 1-(\log n)/n\} = 1. \tag{4.32}$$

Combining Lemma 4.3 with (4.29)–(4.31) we conclude

$$\sup_{1/(n+1) \leqslant t \leqslant n/(n+1)} (T(nt)/(t(1-t)))^{1/2}/(2 \log\log n)^{1/2} \xrightarrow{P} 1,$$

$$\sup_{1/(n+1) \leqslant t \leqslant (\log n)/n} (T(nt)/(t(1-t)))^{1/2} = O_P((\log\log\log n)^{1/2}),$$

$$\sup_{1-(\log n)/n \leqslant t \leqslant n/(n+1)} (T(nt)/(t(1-t)))^{1/2} = O_P((\log\log\log n)^{1/2})$$

and therefore

$$\lim_{n\to\infty} P\{(\log n)/n \leqslant \eta \leqslant 1-(\log n)/n\} = 1, \tag{4.33}$$

where $\eta = \eta(n)$ satisfies

$$\sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} T(nt)/(t(1-t)) = T(n\eta)/(\eta(1-\eta)).$$

Lemma 4.3 yields

$$\begin{aligned}
&\sup_{(\log n)/n\leqslant t\leqslant 1-(\log n)/n} \left| T(nt) - \sum_{1\leqslant i\leqslant d} B^2_{n,k}(t) \right| \Big/ (t(1-t))\\
&\quad\leqslant \sup_{(\log n)/n\leqslant t\leqslant 1-(\log n)/n} \left| T(nt) - \sum_{1\leqslant i\leqslant d} B^2_{n,i}(t) \right| \Big/ (t(1-t))^{1-\varepsilon}\\
&\qquad\times \sup_{(\log n)/n\leqslant t\leqslant 1-(\log n)/n} (t(1-t))^{-\varepsilon}\\
&\quad= O_P(n^{-\varepsilon})\, O((n/\log n)^{\varepsilon})\\
&\quad= O_P(\log n)^{-\varepsilon})
\end{aligned}$$

for all $0<\varepsilon<\kappa_0$. Hence (4.32) and (4.33) imply

$$\begin{aligned}
&\Bigg| \sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} (T(nt)/(t(1-t)))^{1/2}\\
&\qquad - \sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} \left(\sum_{1\leqslant i\leqslant d} B^2_{n,i}(t)/(t(1-t)) \right)^{1/2} \Bigg| = O_P((\log n)^{-\varepsilon})
\end{aligned} \tag{4.34}$$

with some $\varepsilon>0$. Horváth (1993) proved that

$$\begin{aligned}
&\lim_{n\to\infty} P\left\{ a(\log n) \sup_{1/(n+1)\leqslant t\leqslant n/(n+1)} \left(\sum_{1\leqslant i\leqslant d} B^2_{n,i}(t)/(t(1-t)) \right)^{1/2} \leqslant x + b_d(\log n) \right\}\\
&\quad = \exp(-2e^{-x})
\end{aligned} \tag{4.35}$$

for all x, and therefore Theorem 1.3 follows from (4.34).

Proof of Theorem 1.4. First we note that

$$\left| M_n - \max_{1\leqslant k\leqslant n-h} h^{-1/2} \left\{ \left(\sum_{k\leqslant i\leqslant k+h} \mathbf{e}_i \right) \mathbf{D}^{-1} \left(\sum_{k\leqslant i\leqslant k+h} \mathbf{e}_1 \right)' \right\}^{1/2} \right| = O_P((h/n)^{1/2}).$$

Using Lemma 4.1 and condition (1.13) we obtain

$$\left| \max_{1 \leqslant k \leqslant n-h} h^{-1/2} \left\{ \left(\sum_{k \leqslant i \leqslant k+h} \mathbf{e}_i \right) \mathbf{D}^{-1} \left(\sum_{k \leqslant i \leqslant k+h} \mathbf{e}_i \right)' \right\}^{1/2} \right.$$
$$\left. - \max_{1 \leqslant k \leqslant n-h} h^{-1/2} \{ (\mathbf{G}(k+h) - \mathbf{G}(k))\, \mathbf{D}^{-1} (\mathbf{G}(k+h) - \mathbf{G}(k))' \}^{1/2} \right|$$
$$= o_P(1/a(n/h)).$$

We note that

$$\max_{1 \leqslant k \leqslant n-h} h^{-1/2} \{ (\mathbf{G}(k+h) - \mathbf{G}(k))\, \mathbf{D}^{-1} (\mathbf{G}(k+h) - \mathbf{G}(k))' \}^{1/2} \}$$
$$\overset{\mathscr{D}}{=} \max_{1 \leqslant k \leqslant n-h} h^{-1/2} \left(\sum_{1 \leqslant i \leqslant d} (W_i(k+h) - W_i(k))^2 \right)^{1/2}, \tag{4.36}$$

where W_1, W_2, ..., W_d are independent Wiener processes. According to (4.36), it is enough to consider the increments of independent Wiener processes, so Lemma 3.1 in Steinebach and Eastwood (1996) implies Theorem 1.4.

5. PROOFS OF THEOREMS 2.1–2.4

Let

$$\mathbf{u}(k) = \begin{cases} \dfrac{k(n-k^*)}{n} (\boldsymbol{\mu} - \boldsymbol{\lambda}) & \text{if} \quad 1 \leqslant k \leqslant k^* \\ \dfrac{(n-k)k^*}{n} (\boldsymbol{\mu} - \boldsymbol{\lambda}) & \text{if} \quad k^* \leqslant k < n. \end{cases} \tag{5.1}$$

LEMMA 5.1. *If the conditions of Theorem* 2.1 *are satisfied, then*

$$\frac{k(n-k)}{n} \boldsymbol{\Delta}(k) = \mathbf{u}(k) + \mathbf{R}(k) \tag{5.2}$$

and

$$\max_{1 \leqslant k < n} \frac{\|\mathbf{R}(k)\|}{(k(n-k))^{\alpha/2}} = \begin{cases} O_P(n^{1/2-\alpha}), & \textit{if} \quad 0 \leqslant \alpha < 1 \\ O_P((\log\log n/n)^{1/2}), & \textit{if} \quad \alpha = 1 \end{cases} \tag{5.3}$$

Proof. It is easy to see that (5.2) holds with

$$\mathbf{R}(k) = \mathbf{S}(k) - \frac{k}{n}\mathbf{S}(n),$$

and therefore (5.3) follows from (1.10) and Theorem 1.3.

Proof of Theorem 2.1. The proof follows immediately from Lemma 5.1.

Proof of Theorem 2.2. Observing that $\mathbf{u}(k)\,\mathbf{D}^{-1}\mathbf{u}'(k)/(k(n-k))^{\alpha}$ reaches its maximum at k^*, Lemma 5.1 implies Theorem 2.2.

Next we define

$$\mathbf{v}(k) = \begin{cases} \dfrac{k(\boldsymbol{\mu}-\boldsymbol{\lambda})(2n-k^*-l^*+1)}{2n}, & \text{if} \quad 1 \leqslant k \leqslant k^* \\ \dfrac{\boldsymbol{\mu}-\boldsymbol{\lambda}}{2n}\left\{2k(n-l^*)+k(l^*-k^*+1)-\dfrac{n}{l^*-k^*}(k-k^*)(k-k^*+1)\right\}, & \\ & \text{if} \quad k^* < k \leqslant l^* \\ \dfrac{(n-k)(\boldsymbol{\mu}-\boldsymbol{\lambda})(k^*+l^*-1)}{2n}, & \text{if} \quad l^* < k \leqslant n, \end{cases}$$

and similarly to Lemma 5.1 the next result can be proven.

LEMMA 5.2. *If the conditions of Theorem* 2.3 *are satisfied, then*

$$\frac{k(n-k)}{n}\boldsymbol{\Delta}(k) = \mathbf{v}(k) + \mathbf{R}(k)$$

and (5.3) *holds.*

Proofs of Theorems 2.3 *and* 2.4. The proofs follow immediately from Lemma 5.2.

ACKNOWLEDGMENTS

We thank the referee for a very careful reading of the manuscript and Daniella Jarušková for providing the Prague temperature data.

REFERENCES

J. Antoch, M. Hušková, and Z. Prašková, Effect of dependence on statistics for determination of change, *J. Statist. Plann. Inference* **60** (1997), 291–310.

M. Csörgő and L. Horváth, Nonparametric methods for changepoint problems, *in* "Handbook of Statistics," Vol. 7, pp. 403–425, North Holland, Amsterdam, 1988.

M. Csörgő and L. Horváth, "Weighted Approximations in Probability and Statistics," Wiley, Chichester, 1993.

R. Davis, D. Huang, and Y.-C. Yao, Testing for a change in the parameter values and order of an autoregressive model, *Ann. Statist.* **23** (1995), 282–304.

E. Eberlein, On strong invariance principles under dependence assumptions, *Ann. Probab.* **14** (1986), 260–270.

D. L. Hawkins, A $U-I$ approach to retrospective testing for shifting parameters in a linear model, *Comm. Statist. Theory Methods* **18** (1989), 3117–3134.

L. Horváth, The maximum likelihood method for testing changes in the parameters of normal observations, *Ann. Statist.* **21** (1993), 671–680.

M. Hušková, Asymptotics for robust MOSUM, *Comment. Math. Univ. Carolin.* **31** (1990), 345–356.

P. Kokoszka and L. Horváth, The effect of long-range dependence on change-point estimators, *J. Statist. Plann. Inference* **64** (1997), 57–81.

P. Kokoszka and R. Leipus, Change-point in the mean of dependent observations, *Statist. Probab. Lett.* (1998), in press.

J. Kiefer, A functional equation technique for obtaining Wiener process probabilities associated with theorems of Kolmogorov–Smirnov type, *Math. Proc. Cambridge Philos. Soc.* **55** (1959a), 328–332.

J. Kiefer, *K*-sample analogues of the Kolmogorov–Smirnov and Cramér–v. Mises tests, *Ann. Math. Statist.* **30** (1959b), 420–447.

F. W. Scholz and M. A. Stephens, *K*-sample Anderson–Darling tests, *J. Amer. Statist. Assoc.* **82** (1987), 918–924.

J. Steinebach, Change point and jump estimates in an AMOC renewal model, *in* "Asymptotic Statistics, Proc. Fifth Prague Symposium," pp. 447–457, Physica-Verlag, Heidelberg, 1994.

J. Steinebach and V. R. Eastwood, Extreme value asymptotics for multivariate renewal processes, *J. Multivariate Anal.* **56** (1996), 284–302.