Statistics, 2002, Vol. 36(1), pp. 9–31

CHANGE-POINT DETECTION WITH NON-PARAMETRIC REGRESSION

LAJOS HORVÁTH[a,*,†] and PIOTR KOKOSZKA[b,‡]

[a]*Department of Mathematics, University of Utah, Salt Lake City, UT 84112, USA;* [b]*Department of Mathematical Sciences, University of Liverpool, Liverpool, L69 3BX, UK*

(Received 9 November 1998; In final form 4 April 2001)

We consider the regression model $y_i = f(x_i) + \varepsilon$ in which the function f or its pth derivative $f^{(p)}$ may have a discontinuity at some unknown point τ. By fitting local polynomials from the left and right, we test the null that $f^{(p)}$ is continuous against the alternative that $f^{(p)}(\tau-) \neq f^{(p)}(\tau+)$. We obtain Darling-Erdős type limit theorems for the test statistics under the null hypothesis of no change, as well as their limits in probability under the alternative. Consistency of the related change-point estimators is also established.

Keywords: Nonparametric change-point tests; Polynomial smoothing

AMS 1990 Subject classification: 62G07, 62G10, 62G20

1 INTRODUCTION AND RESULTS

We consider the regression model

$$y_i = f(x_i) + \varepsilon_i, \quad 1 \le i \le n,$$

where $x_i = i/n$, $1 \le i \le n$, and $f(t)$ is an unknown function. Let $0 \le p < \infty$ be a given integer. We wish to test the null-hypothesis

$$H_0 : f^{(p)} \text{exists and is continuous on } [0, 1]$$

against the alternative

$$H_a : \text{there is } \tau \in (0, 1) \text{ such that } f^{(p)}(\tau-) \neq f^{(p)}(\tau+).$$

The testing procedure proposed in this paper consists in fitting local polynomials from left and right and checking if the difference between appropriate coefficients of the two polyno-

*Corresponding author.
† E-mail: horvath@math.utah.edu
‡E-mail: P.S.Kokoszka@liverpool.ac.uk

ISSN 0233-1888 print; ISSN 1029-4910 online © 2002 Taylor & Francis Ltd
DOI: 10.1080/02331880290013965

mials is significantly different from zero at at least one point. This paper develops the asymptotic theory which is presented in Sections 1 and 2. Proofs are given in Sections 3 and 4.

Testing for the smoothness of f covers some very important examples in change-point analysis.

Example 1.1 Assume that $f(t)$ is constant under H_0. Under H_a there is $\tau \in (0,1)$ such that $f(t) = f(0)$, if $0 \le t \le \tau$, $f(t) = f(1)$, if $\tau < t \le 1$ and $f(0) \ne f(1)$. In this special case, we test the null hypothesis of constant mean against the alternative that the mean changed at an unknown time. For surveys on detection of changes in the mean we refer to Brodsky and Darkhovsky (1993) and Csörgő and Horváth (1997).

Example 1.2 We assume again that f is constant under H_0. Under the alternative the mean remains constant until $y_{[n\tau]}$ and after that it starts to increase linearly. This means that there are constants c_1 and $c_2 \ne 0$ such that $f(t) = c_1 + c_2(t-\tau)I\{\tau \le t \le 1\}$. For testing and estimating the time of the increase in the mean we refer to Jarušková (1998) and Hušková (1999).

Example 1.3 An epidemic (square wave) alternative to a constant f is defined by $f(t) = c_1 + c_2 I\{\tau_1 \le t \le \tau_2\}$, $c_2 \ne 0$, $0 < \tau_1 < \tau_2 < 1$. For a discussion of epidemic alternatives we refer to Yao (1993a) and Yao (1993b).

Our tests are based on a method proposed by Loader (1996). First we fit a local polynomial to (y_m, x_m) from the left using only (y_i, x_i), $1 \le i \le m-1$. The coefficients of the fitted polynomial minimize the weighted sums of squares

$$\sum_{0<i\le nhC} \left(y_{m-i} - \sum_{0\le j\le p} \alpha_j (x_{m-i} - x_m)^j \right)^2 K((x_{m-i} - x_m)/h), \tag{1.1}$$

where K is a weight function satisfying the following conditions:

$$K(u) = 0 \quad \text{if } |u| \ge C \tag{1.2}$$

$$K(u) = K(-u) \quad \text{for all } -\infty < u < \infty \tag{1.3}$$

and

$$K^{(2)} \text{ exists and is continuous on } (-C, C). \tag{1.4}$$

Throughout this paper we also assume

$$h > 0 \quad \text{and} \quad h = h(n) \to 0 \text{ as } n \to \infty. \tag{1.5}$$

The weighted least-squares estimators $\hat{\boldsymbol{\alpha}}(m) = (\hat{\alpha}_0(m), \dots, \hat{\alpha}_p(m))^T$ of (1.1) are given by

$$\hat{\boldsymbol{\alpha}}(m) = (X_-^T(m) Q_-(m) X_-(m))^{-1} X_-^T(m) Q_-(m) Y_-(m) \tag{1.6}$$

(cf. Fan and Gijbels (1996)), where

$$X_-(m) = \{(x_{m-i} - x_m)^j\}_{1\le i\le nhC, 0\le j\le p},$$

$Y_-(m) = (y_{m-1}, y_{m-2}, \ldots, y_{m-nhC})^T$ and $Q_-(m) = \{q_-(i, j)\}$ is a diagonal matrix with $q_-(i, i) = K((x_{m-i} - x_m)/h)$, $1 \le i \le nhC$.

Next we fit a local polynomial to (y_m, x_m) from the right using (y_i, x_i), $m + 1 \le i \le n$. The coefficients of the polynomial minimize

$$\sum_{0<i\le nhC} \left(y_{m+i} - \sum_{0\le j\le p} \beta_j (x_{m+i} - x_m)^j \right)^2 K((x_{m+i} - x_m)/h). \tag{1.7}$$

Similarly to (1.6), the coefficients $\hat{\boldsymbol{\beta}}(m) = (\hat{\beta}_0(m), \ldots, \hat{\beta}_p(m))^T$ satisfy

$$\hat{\boldsymbol{\beta}}(m) = (X_+^T(m) Q_+(m) X_+(m))^{-1} X_+^T(m) Q_+(m) Y_+(m), \tag{1.8}$$

where $X_+(m) = \{(x_{m+i} - x_m)^j\}_{1\le i\le nhC, 0\le j\le p}$, $Y_+(m) = (y_{m+1}, y_{m+2}, \ldots, y_{m+nhC})^T$ and $Q_+(m) = \{q_+(i, j)\}$ is a diagonal matrix with $q_+(i, i) = K((x_{m+i} - x_m)/h)$.

We compare the coefficient vectors $\hat{\boldsymbol{\alpha}}(m)$ and $\hat{\boldsymbol{\beta}}(m)$ and if they are statistically different for at least one m, $nhC \le m \le n - nhC$, then we reject H_0 in favour of H_a. For this we consider the asymptotic distribution of

$$Z_i(n) = \max_{nhC\le m\le n-nhC} |\hat{\alpha}_i(m) - \hat{\beta}_i(m)|, \quad 0 \le i \le p.$$

We wish to point out that if $p = 0$, then $\hat{\alpha}(m)$ and $\hat{\beta}(m)$ are the usual (one-sided) kernel estimators of Priestley and Chao (1972) for regression functions. For further results on kernel and related nonparametric estimators we refer to Clark (1977), Gasser and Müller (1979), Cheng and Lin (1981) and Stadtmüller (1986). In general, kernel estimators can be used to estimate one-sided (left and right) higher order derivatives of regression functions and thus, using analogues of $Z_i(n)$, to test for discontinuity in the derivatives of regression functions. This idea was developed by Müller (1992). However, Hastie and Loader (1993) pointed out the superiority of local polynomial smoothing to kernel regression estimators.

We assume that $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n$ are independent, identically distributed random variables with

$$E\varepsilon_i = 0, \quad 0 < \sigma^2 = E\varepsilon_i^2 < \infty \quad \text{and} \quad E|\varepsilon_i|^v < \infty \quad \text{for some } v > 2. \tag{1.9}$$

To formulate our main result we need some further notation. Let

$$A_- = \left\{ (-1)^{i+j} \int_0^\infty x^{i+j} K(x)\, \mathrm{d}x \right\}_{0\le i,j\le p}$$

and

$$A_+ = \left\{ \int_0^\infty x^{i+j} K(x)\, \mathrm{d}x \right\}_{0\le i,j\le p}.$$

We also assume that

$$A_-^{-1} = B_- = \{b_-(i, j)\} \quad \text{and} \quad A_+^{-1} = B_+ = \{b_+(i, j)\} \text{ exist.} \tag{1.10}$$

Next we define

$$\gamma_0(i,j) = (-1)^{i+j}\int_0^\infty x^{i+j}K^2(x)\,\mathrm{d}x, \quad \text{if } 0 \le i,j \le p$$

$$\gamma_1(i,j) = \begin{cases} -\frac{1}{2}K^2(0), & \text{if } i=j=0 \\ \dfrac{i-j}{2}(-1)^{i+j-1}\displaystyle\int_0^\infty x^{i+j-1}K^2(x)\,\mathrm{d}x & \text{otherwise} \end{cases}$$

$$\gamma_2(i,j) = \begin{cases} -\frac{1}{2}K(0)K'(0) - \frac{1}{2}\displaystyle\int_0^\infty (K'(x))^2\,\mathrm{d}x, \quad \text{if } i=j=0 \\ -\frac{1}{2}(-1)^{i+j}\displaystyle\int_0^\infty x^{i+j}(K'(x))^2\,\mathrm{d}x, \\ \qquad\qquad \text{if } 0 \le i,j \le 1 \text{ and } \max(i,j) > 0 \\ -\frac{1}{2}(-1)^{i+j}\left\{\int_0^\infty x^{i+j}(K'(x))^2\,\mathrm{d}x - \dfrac{i^2-i+j^2-j}{2}\displaystyle\int_0^\infty x^{i+j-2}K^2(x)\mathrm{d}\right\} \\ \qquad\qquad \text{otherwise} \end{cases}$$

$$\delta_0(i,j) = \int_0^\infty x^{i+j}K^2(x)\,\mathrm{d}x, \quad 0 \le i,j < \infty$$

$$\delta_1(i,j) = \begin{cases} -\frac{1}{2}K^2(0), & \text{if } i=j=0 \\ \dfrac{i-j}{2}\displaystyle\int_0^\infty x^{i+j-1}K^2(x)\,\mathrm{d}x & \text{otherwise} \end{cases}$$

$$\delta_2(i,j) = \begin{cases} -\frac{1}{2}\displaystyle\int_0^\infty (K'(x))^2\,\mathrm{d}x, \quad \text{if } i=j=0 \\ -\frac{1}{2}\displaystyle\int_0^\infty x^{i+j}(K'(x))^2\,\mathrm{d}x + \frac{1}{4}K^2(0), \quad \text{if } i=0, j=1 \\ -\frac{1}{2}\displaystyle\int_0^\infty x^{i+j}(K'(x))^2\,\mathrm{d}x - \frac{1}{4}K^2(0), \quad \text{if } i=1, j=0 \\ -\frac{1}{2}\left\{\displaystyle\int_0^\infty x^{i+j}(K'(x))^2\,\mathrm{d}x - \dfrac{i^2-i+j^2-j}{2}\displaystyle\int_0^\infty x^{i+j-2}K^2(x)\,\mathrm{d}x\right\} \\ \qquad\qquad \text{otherwise} \end{cases}$$

$$\tau_0(i,j) = 0, \quad \text{if } 0 \le i,j < \infty$$

$$\tau_1(i,j) = \begin{cases} K^2(0), & \text{if } i=j=0 \\ 0 & \text{otherwise} \end{cases}$$

$$\tau_2(i,j) = \begin{cases} -\frac{1}{2}K^2(0), & \text{if } i=0, j=1 \\ -\frac{1}{2}K^2(0), & \text{if } i=1, j=0 \\ 0 & \text{otherwise} \end{cases}$$

For any $0 \le i \le p$ and $\ell = 0, 1, 2$ we define

$$A_\ell(i) = \sum_{0\le j,k\le p} \{b_-(i,j)b_-(i,k)\gamma_\ell(j,k) + b_+(i,j)b_+(i,k)\delta_\ell(j,k) - b_+(i,j)b_-(i,k)\tau_\ell(j,k)\}.$$

THEOREM 1.1 *Let $0 \le i \le p$. We assume that (1.2)–(1.5), (1.9), (1.10) hold, $A_0(i) > 0$,*

$$f^{(p)} \text{ exists and is Lipschitz } 1 \text{ on } [0,1], \tag{1.11}$$

$$\limsup_{n\to\infty} (\log(1/h))^{1/2} n^{(2-\nu)/(2\nu)} h^{-1/2} < \infty. \tag{1.12}$$

and

$$\lim_{n\to\infty} (\log 1/h)^{1/2} n^{1/2} h^{p+3/2} = 0. \tag{1.13}$$

(i) *If $-\infty < A_1(i) < 0$, then*

$$\lim_{n\to\infty} P\Bigg\{(2\log(1/h))^{1/2} n^{1/2} h^{i+1/2} Z_i(n)/\sigma \le t + 2\log(1/h). + \frac{1}{2}\log\log(1/h) + \log\left(\frac{1}{\pi^{1/2}}\left(-\frac{A_1(i)}{A_0(i)}\right)\right)\Bigg\} = \exp(-2e^{-t}) \tag{1.14}$$

for all $-\infty < t < \infty$.

(ii) *If $A_1(i) = 0$ and $-\infty < A_2(i) < 0$, then*

$$\lim_{n\to\infty} P\Bigg\{(2\log(1/h))^{1/2} n^{1/2} h^{i+1/2} Z_i(n)/\sigma \le t + 2\log(1/h). + \log\left(\frac{1}{2^{1/2}\pi}\left(-\frac{A_2(i)}{A_0(i)}\right)^{1/2}\right)\Bigg\} = \exp(-2e^{-t}) \tag{1.15}$$

for all $-\infty < t < \infty$.

It is known that for every fixed $0 < t < 1$, $\hat{\alpha}_p([nt])$ and $\hat{\beta}_p([nt])$ are asymptotically normal (c.f. Fan and Gijbels (1996, p. 116)). However, less is known about the global properties of

local polynomial smoothing. For example, the asymptotic distributions of the Kolmogorov–Smirnov and related sup statistics are unknown. The proof of Theorem 1.1 indicates that these statistics must have double exponential (extreme value) limit distributions. Local polynomial smoothing is related to kernel-type estimators, c.f. Fan and Gijbels (1996, p. 63), and therefore it is not surprising that Theorem 1.1 resembles analogous limit theorems for the supremum of kernel-type estimators for densities and regression functions. As in the case of all kernel-type estimators, the rate of convergence can be slow, especially for large p.

Theorem 1.1 will be proven in Section 3. Here we consider two special cases. First we assume that $p = 0$, that is we fit constants to the observations.

COROLLARY 1.1 *Let $p = 0$. We assume that (1.2)–(1.5), (1.9)–(1.13) hold and $\int_0^\infty K(x)\,\mathrm{d}x \neq 0$.*

(i) *If $K^2(0) > 0$, then*

$$\begin{aligned}\lim_{n\to\infty} P\Bigg\{&(2\log(1/h))^{1/2}(nh)^{1/2}Z_0(n)/\sigma \le t + 2\log(1/h).\\ &+\frac{1}{2}\log\log(1/h) - \frac{1}{2}\log\pi + \log\left(K^2(0)\Big/\int_0^\infty K^2(x)\,\mathrm{d}x\right)\Bigg\}\\ &= \exp(-2e^{-t})\end{aligned} \tag{1.16}$$

for all $-\infty < t < \infty$.

(ii) *If $K(0) = 0$, then*

$$\begin{aligned}\lim_{n\to\infty} P\Bigg\{&(2\log(1/h))^{1/2}(nh)^{1/2}Z_0(n)/\sigma \le t + 2\log(1/h)\\ &-\log(2^{1/2}\pi) + \frac{1}{2}\log\left(\int_0^\infty (K'(x))^2\,\mathrm{d}x\Big/\int_{-\infty}^\infty K^2(x)\,\mathrm{d}x\right)\Bigg\}\\ &= \exp(-2e^{-t})\end{aligned} \tag{1.17}$$

for all $-\infty < t < \infty$.

If $p = 0$, then $\hat{\alpha}_0(m)$ and $\hat{\beta}_0(m)$ are kernel estimators of the regression function at m/n. In this case an analogue of Corollary 1.1 was adapted by Wu and Chu (1993) from Stadtmüller (1986). Wu and Chu (1993) considered the limit distribution of $\max_{n\delta\le m\le(1-\delta)n} |\hat{\alpha}_0(m) - \hat{\beta}_0(m)|$ with some $0 < \delta < 1/2$, and thus essentially assumed that no change can occur on the intervals $[0, \delta]$ and $[1-\delta, 1]$, so their test does not have power against changes on these intervals.

Now we assume that $p = 1$, that is we fit lines to the observations.

COROLLARY 1.2 *Let $p = 1$. We assume that (1.2)–(1.5), (1.9)–(1.13) hold,*

$$\int_0^\infty K(x)\,\mathrm{d}x \int_0^\infty x^2K(x)\,\mathrm{d}x \neq \left(\int_0^\infty xK(x)\,\mathrm{d}x\right)^2 \tag{1.18}$$

and

$$\omega_0 = \left(\int_0^\infty x^2 K(x)\,\mathrm{d}x\right)^2 \int_0^\infty K^2(x)\,\mathrm{d}x + \left(\int_0^\infty xK(x)\,\mathrm{d}x\right)^2 \times \int_0^\infty x^2 K^2(x)\,\mathrm{d}x - 2\int_0^\infty x^2 K(x)\,\mathrm{d}x \int_0^\infty xK(x)\,\mathrm{d}x \int_0^\infty xK^2(x)\,\mathrm{d}x > 0. \tag{1.19}$$

(i) *If*

$$\omega_1 = K^2(0)\left(\int_0^\infty x^2 K(x)\,\mathrm{d}x\right)^2 > 0, \tag{1.20}$$

then

$$\lim_{n\to\infty} P\Big\{(2\log(1/h))^{1/2}(nh)^{1/2} Z_0(n)/\sigma \le t + 2\log(1/h) + \tfrac{1}{2}\log\log(1/h) - \tfrac{1}{2}\log\pi + \log(\omega_1/\omega_0)\Big\} = \exp(-2e^{-t}) \tag{1.21}$$

for all $-\infty < t < \infty$.

(ii) *If* $K(0) = 0$ *and*

$$\omega_2 = \frac{1}{2}\left(\int_0^\infty x^2 K(x)\,\mathrm{d}x\right)^2 \int_0^\infty (K'(x))^2\,\mathrm{d}x + \frac{1}{2}\left(\int_0^\infty xK(x)\,\mathrm{d}x\right)^2 \int_0^\infty x^2 (K'(x))^2\,\mathrm{d}x - \int_0^\infty xK(x)\,\mathrm{d}x \int_0^\infty x^2 K(x)\,\mathrm{d}x \int_0^\infty x(K'(x))^2\,\mathrm{d}x > 0, \tag{1.22}$$

then

$$\lim_{n\to\infty} P\Big\{(2\log(1/h))^{1/2}(nh)^{1/2} Z_0(n)/\sigma \le t + 2\log(1/h) - \log(2^{1/2}\pi) + \tfrac{1}{2}\log(\omega_2/\omega_0)\Big\} = \exp(-2e^{-t}) \tag{1.23}$$

for all $-\infty < t < \infty$.

(iii) *If*

$$\omega_3 = \frac{1}{2}\left(\int_0^\infty K(x)\,\mathrm{d}x\right)^2 \int_0^\infty x^2 (K'(x))^2\,\mathrm{d}x - \frac{1}{2}K(0)K'(0)\int_0^\infty xK(x)\,\mathrm{d}x - \int_0^\infty xK(x)\,\mathrm{d}x \int_0^\infty K(x)\,\mathrm{d}x \int_0^\infty x(K'(x))^2\,\mathrm{d}x > 0, \tag{1.24}$$

then

$$\lim_{n\to\infty} P\Big\{(2\log(1/h))^{1/2} n^{1/2} h^{3/2} Z_1(n)/\sigma \le t + 2\log(1/h) - \log(2^{1/2}\pi) + \tfrac{1}{2}\log(\omega_3/\omega_0)\Big\} = \exp(-2e^{-t}) \tag{1.25}$$

for all $-\infty < t < \infty$.

Remark 1.1 If $h(n) = an^{-b}$, then (1.12) and (1.13) are satisfied for any $1/(2p+3) < b < 1 - 2/v$.

2 TESTS AND ESTIMATORS UNDER THE ALTERNATIVE

We discuss briefly the consistency of the tests described in Section 1 and point out that local polynomial smoothing can also be used to estimate the time of change. Let $0 < \tau < 1$ and define

$$f(t) = \begin{cases} f_1(t), & \text{if } 0 \le t \le \tau \\ f_2(t), & \text{if } \tau < t \le 1. \end{cases}$$

We assume that $f_1(t)$ and $f_2(t)$ are smooth functions. Namely,

$$f_1^{(p)} \text{ exists and is Lipschitz 1 on } [0, \tau] \tag{2.1}$$

and

$$f_2^{(p)} \text{ exists and is Lipschitz 1 on } [\tau, 1]. \tag{2.2}$$

The two parts of $f(t)$ connect smoothly up to the order $p-1$ and

$$f^{(p-1)} \text{ exists and is bounded on } [0, 1]. \tag{2.3}$$

However,

$$f^{(p)}(\tau-) \ne f^{(p)}(\tau+). \tag{2.4}$$

THEOREM 2.1 *We assume that (1.2)–(1.5), (1.9), (1.10), (1.12), (2.1)–(2.4) are satisfied.*

(i) *If*

$$\lim_{n\to\infty} n^{1/2} h^{p+1/2} (\log 1/h)^{1/2} = 0, \tag{2.5}$$

then the results of Theorem 1.1 hold for $0 \le i \le p-1$.

(ii) *Moreover*,

$$\hat{\alpha}_p([n\tau]) - \hat{\beta}_p([n\tau]) \xrightarrow{P} f^{(p)}(\tau-) - f^{(p)}(\tau+). \tag{2.6}$$

Next we consider an estimator for τ proposed by Loader (1996). Let

$$\hat{\tau} = \min\{t : h \le t \le 1 - h \text{ and } |\hat{\alpha}_p([n\tau]) - \hat{\beta}_p([n\tau])| = Z_p(n)\}.$$

Loader (1996) obtained the asymptotic distribution of $n(\hat{\tau} - \tau)$ assuming that the errors are normally distributed. However, his model, and consequently the assumptions on f, are different from ours. Here we show that the estimator $\hat{\tau}$ is weakly consistent.

THEOREM 2.2 *We assume that (1.2)–(1.5), (1.9), (1.10), (1.12) and (2.1)–(2.4) hold. Then*

$$|\hat{\tau} - \tau| = O_P(h). \tag{2.7}$$

In the case of $p = 0$, which corresponds to kernel regression estimators, Müller (1992) and Wu and Chu (1993) obtained several limit results for

$$\begin{aligned}\tau^* &= \min\Big\{t : \delta \le t \le 1 - \delta \text{ and } |\hat{\alpha}_0([nt]) - \hat{\beta}_0([nt])| \\ &= \max_{\delta \le s \le 1-\delta} |\hat{\alpha}_0([ns]) - \hat{\beta}_0([ns])|\Big\},\end{aligned}$$

where $0 < \delta < 1/2$. By choosing δ in the definition of τ^*, it is implicitely assumed that we know that no change occured on $[0, \delta]$ and $[1 - \delta, 1]$. Müller (1992) also considered kernel-type estimators for the location of the discontinuity of higher order derivatives of the regression function. Hastie and Loader (1993) argued that local polynomial smoothing is preferable to kernel estimators in the context of change-point (or discontinuity) detection.

Theorems 2.1 and 2.2 are proven in Section 4.

3 PROOFS OF THEOREM 1.1 AND COROLLARIES 1.1 AND 1.2

Let $H = \{h(i, j)\}_{0 \le i, j \le p}$ be a diagonal matrix with $h(i, i) = h^{-i-1/2}$, $0 \le i \le p$. Let $E_-(m) = (\varepsilon_{m-1}, \ldots, \varepsilon_{m-nhC})^T$, $Y_-^*(m) = (f(x_{m-1}), \ldots, f(x_{m-nhC}))^T$ and

$$\boldsymbol{\alpha}(m) = \left(f(x_m),\ f^{(1)}(x_m),\ \frac{1}{2!} f^{(2)}(x_m), \ldots, \frac{1}{p!} f^{(p)}(x_m)\right).$$

Elementary algebra shows that

$$\begin{aligned}&H^{-1}(\hat{\boldsymbol{\alpha}}(m) - \boldsymbol{\alpha}(m)) \\ &= A_-^{-1}(m) H X_-^T(m) Q_-(m) E_-(m) + A_-^{-1}(m) H X_-^T(m) Q_-(m) \\ &\quad \times \{Y_-^*(m) - X_-(m)\alpha(m)\},\end{aligned} \tag{3.1}$$

where

$$A_-(m) = HX_-^T(m)Q_-(m)X_-(m)H.$$

In (3.1) we decomposed $\hat{\boldsymbol{\alpha}}(m) - \boldsymbol{\alpha}(m)$ into stochastic and numerical terms. Putting together Lemmas 3.1 and 3.2 we get the exact order of the numerical term. In Lemmas 3.3–3.5 we consider the stochastic term.

Throughout this paper $|\cdot|$ denotes the maximum norm of vectors and matrices.

LEMMA 3.1 *If (1.2)–(1.5) and (1.10) hold, then*

$$\max_{nhC \le m \le n-nhC} \left| \frac{1}{n} A_-(m) - A_- \right| = O\left(\frac{1}{nh}\right) \tag{3.2}$$

and

$$\max_{nhC \le m \le n-nhC} \left| A_-^{-1}(m) - \frac{1}{n} B_- \right| = O\left(\frac{1}{n^2 h}\right). \tag{3.3}$$

Proof It is easy to see that $A_-(m) = \{u_{i,j}\}_{0 \le i,j \le p}$, where

$$u_{i,j} = h^{-(i+j+1)} \sum_{0<k\le nhC} (x_{m-k} - x_m)^{i+j} K((x_{m-k} - x_m)/h).$$

Since $x_i = i/n$ we get

$$u_{i,j} = (-1)^{i+j} n \frac{1}{nh} \sum_{0<k\le nhC} \left(\frac{k}{nh}\right)^{i+j} K\left(\frac{k}{nh}\right).$$

By (1.4) we have that

$$\left| \frac{1}{nh} \sum_{0<k\le nhC} \left(\frac{k}{nh}\right)^{i+j} K\left(\frac{k}{nh}\right) - \int_0^\infty x^{i+j} K(x)\,\mathrm{d}x \right| = O\left(\frac{1}{nh}\right), \tag{3.4}$$

which completes the proof of (3.2). The result in (3.3) follows immediately from (3.2).

LEMMA 3.2 *If (1.2)–(1.5) and (1.11) hold, then*

$$\max_{nhC \le m \le n-nhC} \left| HX_-^T(m)Q_-(m)\{Y_-^*(m) - X_-(m)\boldsymbol{\alpha}(m)\} \right| = O(nh^{p+3/2}). \tag{3.5}$$

Proof First we note that

$$
\begin{aligned}
&HX_-^T(m)Q_-(m)\{Y_-^*(m) - X_-(m)\boldsymbol{\alpha}(m)\} \\
&\quad = \Bigg(h^{-1/2} \sum_{0<k\le nhC} \Bigg[f(x_{m-k}) - \sum_{0\le i\le p} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \Bigg] \\
&\qquad \times K((x_{m-k} - x_m)/h), \ldots, \\
&\qquad h^{-p-1/2} \sum_{0<k\le nhC} \Bigg[f(x_{m-k}) - \sum_{0\le i\le p} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \Bigg] \\
&\qquad \times (x_{m-k} - x_m)^p K((x_{m-k} - x_m)/h) \Bigg)^T . \qquad (3.6)
\end{aligned}
$$

Taylor expansion yields

$$
\max_{0<k\le nhC} \left| f(x_{m-k}) - \sum_{0\le i\le p} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \right| = O(h^{p+1}). \qquad (3.7)
$$

By (3.4) we have

$$
\begin{aligned}
\max_{nhC\le m\le n-nhC} \Bigg| &\frac{1}{nh^{j+1}} \sum_{0<k\le nhC} (x_{m-k} - x_m)^j K((x_{m-k} - x_m)/h) \\
&- (-1)^j \int_0^C x^j K(x)\,\mathrm{d}x \Bigg| = O\left(\frac{1}{nh}\right) \qquad (3.8)
\end{aligned}
$$

and therefore the proof of Lemma 3.2 is complete.

We introduce the notation

$$
T_-^{(i)}(m) = h^{-i-1/2} \sum_{0<k\le nhC} \varepsilon_{m-k}(x_{m-k} - x_m)^i K((x_{m-k} - x_m)/h), \qquad (3.9)
$$

$0 \le i \le p$, and therefore we have

$$
HX_-^T(m)Q_-(m)E_-(m) = (T_-^{(0)}(m), \ldots, T_-^{(p)}(m))^T . \qquad (3.10)
$$

It is easy to see that

$$
\begin{aligned}
T_-^{(i)}(m) &= h^{-i-1/2} \sum_{0<k\le nhC} \varepsilon_{m-k} \left(-\frac{k}{n}\right)^i K\left(-\frac{k}{nh}\right) \\
&= h^{-i-1/2} \sum_{m-nhC\le j<m} \varepsilon_j \left(\frac{j-m}{n}\right)^i K\left(\frac{j-m}{nh}\right).
\end{aligned}
$$

Introducing $S(x) = \sum_{1 \le i \le x} \varepsilon_i$ we can write

$$T_-^{(i)}(m) = h^{-1/2} \int_{-\infty}^{\infty} \left(\frac{x}{nh} - \frac{m}{nh}\right)^i K\left(\frac{x}{nh} - \frac{m}{nh}\right) I\left\{\frac{m}{nh} - C \le \frac{x}{nh} < \frac{m}{nh}\right\} \mathrm{d}S(x).$$

In Komlós *et al.* (1975, 1976) a Wiener process $\{W(x), -\infty < x < \infty\}$ was constructed such that

$$|S(x) - \sigma W(x)| = o(x^{1/\nu}) \quad \text{a.s.} \tag{3.11}$$

as $x \to \infty$, assuming that (1.9) holds. The continuous Gaussian counterpart of $T_-^{(i)}$ is defined as

$$\Gamma_-^{(i)}(t) = h^{-1/2} \int_{-\infty}^{\infty} \left(\frac{x}{nh} - t\right)^i K\left(\frac{x}{nh} - t\right) I\left\{t - C \le \frac{x}{nh} < t\right\} \mathrm{d}W(x).$$

LEMMA 3.3 *If (1.2)–(1.5)* and (1.9) hold, then

$$\max_{nhC \le m \le n - nhC} \left| T_-^{(i)}(m) - \sigma \Gamma_-^{(i)}\left(\frac{m}{nh}\right) \right| = o_P(h^{-1/2} n^{1/\nu}), \quad 0 \le i \le p.$$

Proof Integration by parts gives

$$\begin{aligned}
&\max_{nhC \le m \le n - nhC} \int_{-\infty}^{\infty} \left(\frac{x}{nh} - \frac{m}{nh}\right)^i K\left(\frac{x}{nh} - \frac{m}{nh}\right) \\
&\quad \times I\left\{\frac{m}{nh} - C \le \frac{x}{nh} \le \frac{m}{nh}\right\} \mathrm{d}(S(x) - \sigma W(x)) \\
&\le 2C^i \sup_{0 \le x \le n} |S(x) - \sigma W(x)| \sup_{-C \le u \le 0} |K(u)| \\
&\quad + \max_{nhC \le m \le n - nhC} \left| \int_{m - nhC}^{m} (S(x) - \sigma W(x)) \left\{ \frac{i}{nh} \left(\frac{x}{nh} - \frac{m}{nh}\right)^{i-1} \right.\right. \\
&\quad \left.\left. \times K\left(\frac{x}{nh} - \frac{m}{nh}\right) + \frac{1}{nh} \left(\frac{x}{nh} - \frac{m}{nh}\right)^i K'\left(\frac{x}{nh} - \frac{m}{nh}\right) \right\} \mathrm{d}x \right| \\
&= a_{n,1} + a_{n,2}.
\end{aligned}$$

It follows immediately from (3.11) that

$$a_{n,1} = o_P(n^{1/\nu}).$$

Using (1.3), (1.4) and (3.11) we obtain that

$$a_{n,2} = o_P(n^{1/\nu}).$$

LEMMA 3.4 *If (1.2)–(1.5) and (1.13) hold, then*

$$\max_{nhC \le m \le n-nhC} \sup_{|s| \le 1/(nh)} \left| \Gamma_-^{(i)}\left(\frac{m}{nh}\right) - \Gamma_-^{i}\left(\frac{m}{nh}+s\right) \right| = O_P(h^{-1/2}(\log 1/h)^{1/2}), \quad 0 \le i \le p.$$

Proof By the scale transformation of the Wiener process we have that for each n and h

$$\{\Gamma_-^{(i)}(t), -\infty < t < \infty\} \stackrel{\mathcal{D}}{=} \{n^{1/2}\hat{\Gamma}_-^{(i)}(t), -\infty < t < \infty\}, \tag{3.12}$$

where

$$\hat{\Gamma}_-^{(i)}(t) = \int_{-\infty}^{\infty} \hat{K}^{(i)}(x-t)\mathrm{d}W(x)$$

with

$$\hat{K}^{(i)}(u) = u^i K(u) I\{-C \le u \le 0\}.$$

Using (1.4) we can find a constant c_1 such that

$$(E(\hat{\Gamma}_-^{(i)}(t) - \hat{\Gamma}_-^{(i)}(s))^2)^{1/2} \le c_1|t-s|^{1/2}. \tag{3.13}$$

By (3.13), we can apply Fernique's (1975) inequality and obtain

$$P\left\{ \sup_{|s| \le 1/(nh)} \left| \hat{\Gamma}_-^{(i)}\left(\frac{m}{nh}\right) - \hat{\Gamma}_-^{(i)}\left(\frac{m}{nh}+s\right) \right| > c_2(nh)^{-1/2}(\log n)^{1/2} \right\} \le \frac{c_3}{n^2}, \tag{3.14}$$

with some constants c_2 and c_3. Now (3.14) yields

$$P\left\{ \max_{nhC \le m \le n-nhC} \sup_{|s| \le 1/(nh)} \left| \hat{\Gamma}_-^{(i)}\left(\frac{m}{nh}\right) - \hat{\Gamma}_-^{(i)}\left(\frac{m}{nh}+s\right) \right| > c_2(nh)^{-1/2}(\log n)^{1/2} \right\} \le \frac{c_3}{n}. \tag{3.15}$$

Lemma 3.4 follows immediately from (3.12), (3.15) and condition (1.13).

LEMMA 3.5 *If (1.2)–(1.5) and (1.9)–(1.13) hold, then*

$$\sup_{C \le t \le 1/h-C} |T_-^{(i)}([nht]) - \sigma\Gamma_-^{(i)}(t)| = o_P(n^{1/\nu}h^{-1/2}) + O_P(h^{-1/2}(\log 1/h)^{1/2}).$$

Proof It follows immediately from Lemmas 3.2–3.4.

Similarly to (3.1) we have

$$\begin{aligned} &H^{-1}(\hat{\boldsymbol{\beta}}(m) - \boldsymbol{\alpha}(m)) \\ &\quad = A_+^{-1}(m)HX_+^T(m)Q_+(m)E_+(m) + A_+^{-1}(m)HX_+^T(m)Q_+(m) \\ &\qquad \times \{Y_+^*(m) - X_+(m)\boldsymbol{\alpha}(m)\}, \end{aligned} \tag{3.16}$$

where

$$A_+(m) = HX_+^T(m)Q_+(m)X_+(m)H,$$

$E_+(m) = (\varepsilon_{m+1}, \dots, \varepsilon_{m+nhC})^T$ and $Y_+^*(m) = (f(x_{m+1}), \dots, f(x_{m+nhC}))^T$. The following five lemmas can be proven with minor modifications of Lemmas 3.1–3.5.

LEMMA 3.6 *If (1.2)–(1.5) and (1.10) hold, then*

$$\max_{nhC \le m \le n-nhC} \left| \frac{1}{n} A_+(m) - A_+ \right| = O\left(\frac{1}{nh}\right)$$

and

$$\max_{nhC \le m \le n-nhC} \left| A_+^{-1}(m) - \frac{1}{n} B_+ \right| = O\left(\frac{1}{n^2 h}\right).$$

LEMMA 3.7 *If (1.2)–(1.5) and (1.11) hold, then*

$$\max_{nhC \le m \le n-nhC} \left| HX_+^T(m)Q_+(m)\{Y_+^*(m) - X_+(m)\boldsymbol{\alpha}(m)\} \right| = O(nh^{p+3/2}).$$

Next we introduce

$$T_+^{(i)}(m) = h^{-i-1/2} \sum_{0<k\le nhC} \varepsilon_{m+k}(x_{m+k} - x_m)^i K((x_{m+k} - x_m)/h)$$

and

$$\Gamma_+^{(i)}(t) = h^{-1/2} \int_{-\infty}^{\infty} \left(\frac{x}{nh} - t\right)^i K\left(\frac{x}{nh} - t\right) I\left\{t < \frac{x}{nh} \le t + C\right\} \mathrm{d}W(x),$$

where $\{W(x), -\infty < x < \infty\}$ is the Wiener process of (3.11). Similarly to (3.10) we have

$$HX_+^T(m)Q_+(m)E_+(m) = (T_+^{(0)}(m), \dots, T_+^{(p)}(m))^T. \tag{3.17}$$

LEMMA 3.8 *If (1.2)–(1.5) and (1.9) hold, then*

$$\max_{nhC \le m \le n-nhC} \left| T_+^{(i)}(m) - \sigma \Gamma_+^{(i)}\left(\frac{m}{nh}\right) \right| = o_P(h^{-1/2} n^{1/\nu}), \quad 0 \le i \le p.$$

LEMMA 3.9 *If (1.2)–(1.5) and (1.13) hold, then*

$$\max_{nhC \le m \le n-nhC} \sup_{|s| \le 1/(nh)} \left| \Gamma_+^{(i)}\left(\frac{m}{nh} + s\right) - \Gamma_+^{(i)}\left(\frac{m}{nh}\right) \right| = O_P(h^{-1/2}(\log 1/h)^{1/2}), \quad 0 \le i \le p.$$

LEMMA 3.10 *If (1.2)–(1.5) and (1.9)–(1.13) hold, then*

$$\sup_{C \le t \le 1/h - C} |T_+^{(i)}([nht]) - \sigma \Gamma_+^{(i)}(t)| = o_P(n^{1/\nu} h^{-1/2}) + O_P(h^{-1/2}(\log 1/h)^{1/2}).$$

Proof of Theorem 1.1. First we note that $\Gamma_-^{(i)}, \Gamma_+^{(i)} (0 \le i \le p)$ are stationary Gaussian processes and therefore using Pickands (1969) we get

$$\sup_{C \le t \le 1/h} |\Gamma_-^{(i)}(t)| = O_P((nh)^{1/2}(\log 1/h)^{1/2}), \quad 0 \le i \le p \tag{3.18}$$

and

$$\sup_{C \le t \le 1/h} |\Gamma_+^{(i)}(t)| = O_P((nh)^{1/2}(\log 1/h)^{1/2}), \quad 0 \le i \le p. \tag{3.19}$$

Putting together (3.1), Lemmas 3.1, 3.2, 3.5–3.7, 3.10, (3.16), (3.18) and (3.19) we get

$$\begin{aligned} & n^{1/2} h^{i+1/2} \max_{nhC \le m \le n-nhC} |\hat{\alpha}_i(m) - \hat{\beta}_i(m)| \\ & = n^{-1/2} \sigma \sup_{C \le t \le 1/h - C} \left| \sum_{0 \le j \le p} \{ b_-(i,j) \Gamma_-^{(i)}(t) - b_+(i,j) \Gamma_+^{(i)}(t) \} \right| \\ & \quad + O_P(h^{-1/2} n^{(2-\nu)/(2\nu)}) + O_P(n^{1/2} h^{p+3/2}) \\ & \quad + O_P((nh)^{-1/2}(\log 1/h)^{1/2}). \end{aligned} \tag{3.20}$$

Let

$$\Delta_-^{(i)}(t) = \int_{-\infty}^{\infty} (x-t)^i K(x-t) I\{-C \le x - t < 0\} \, \mathrm{d}W(x), \quad 0 \le i \le p$$

and

$$\Delta_+^{(i)}(t) = \int_{-\infty}^{\infty} (x-t)^i K(x-t) I\{0 < x-t \le C\}\, \mathrm{d}W(x), \quad 0 \le i \le p.$$

By the scale transformation of the Wiener process we have

$$\begin{aligned} &\{n^{-1/2}\Gamma_-^{(i)}(t), n^{-1/2}\Gamma_+^{(i)}(t), \quad 0 \le i \le p, -\infty < t < \infty\} \\ &\stackrel{\mathcal{D}}{=} \{\Delta_-^{(i)}(t), \Delta_+^{(i)}(t), \quad 0 \le i \le p, -\infty < t < \infty\}. \end{aligned} \tag{3.21}$$

It is easy to see that

$$\Delta^{(i)}(t) = \sum_{0 \le j \le p} \{b_-(i,j)\Delta_-^{(i)}(t) - b_+(i,j)\Delta_+^{(i)}(t)\}, \quad 0 \le i \le p,$$

is a stationary Gaussian process with $E\Delta^{(i)}(t) = 0 (0 \le i \le p)$. Next we show that

$$E\Delta_-^{(i)}(t)\Delta_-^{(j)}(t+h) = \gamma_0(i,j) + \gamma_1(i,j)h + \gamma_2(i,j)h^2 + o(h^2), \tag{3.22}$$

$$E\Delta_+^{(i)}(t)\Delta_+^{(j)}(t+h) = \delta_0(i,j) + \delta_1(i,j)h + \delta_2(i,j)h^2 + o(h^2), \tag{3.23}$$

$$E\Delta_+^{(i)}(t)\Delta_-^{(j)}(t+h) = \tau_0(i,j) + \tau_1(i,j)h + \tau_2(i,j)h^2 + o(h^2) \tag{3.24}$$

as $h \downarrow 0$ and

$$E\Delta_-^{(i)}(t)\Delta_+^{(j)}(t+h) = 0 \quad \text{for all } h > 0, \tag{3.25}$$

$0 \le i, j \le p$. First we note that (3.25) follows from the fact that W has independent increments. Since the proofs of (3.22)–(3.24) are similar, we prove only (3.22) when $i = j = 0$. It is easy to see that

$$\begin{aligned} &E\Delta_-^{(0)}(t)\Delta_-^{(0)}(t+h) \\ &\quad = \int_{-\infty}^{\infty} K(x-t)K(x-t-h) I\{-C \le x-t-h < -h\}\, \mathrm{d}x \\ &\quad = \int_{-C}^{-h} K(u)K(u+h)\mathrm{d}u \\ &\quad = \int_{-C}^{0} K(u)K(u+h)\, \mathrm{d}u - \int_{-h}^{0} K(u)K(u+h)\, \mathrm{d}u. \end{aligned}$$

A two-term Taylor expansion gives

$$\begin{aligned} \int_{-C}^{0} K(u)K(u+h)\, \mathrm{d}u = &\int_{-C}^{0} K^2(u)\, \mathrm{d}u + h\int_{-C}^{0} K(u)K'(u)\, \mathrm{d}u \\ &+ \frac{1}{2}h^2 \int_{-C}^{0} K(u)K''(u)\, \mathrm{d}u + o(h^2). \end{aligned}$$

Note that

$$\int_{-C}^{0} K(u)K'(u)\,\mathrm{d}u = \frac{1}{2}(K^2(0) - K^2(-C)) = \frac{1}{2}K^2(0),$$

and, using integration by parts,

$$\int_{-C}^{0} K(u)K''(u)\,\mathrm{d}u = \int_{-C}^{0} K(u)\,\mathrm{d}K'(u) = K(0)K'(0) - \int_{-C}^{0} (K'(u))^2\,\mathrm{d}u.$$

Using again Taylor expansion we obtain that

$$\begin{aligned}\int_{-h}^{0} K(u)K(u+h)\,\mathrm{d}u &= \int_{-h}^{0} (K(0) + K'(0)u + O(h^2)) \\ &\quad \times (K(0) + K'(0)(u+h) + O(h^2))\,\mathrm{d}u \\ &= \int_{-h}^{0} (K^2(0) + 2K(0)K'(0)u + K(0)K'(0)h)\,\mathrm{d}u \\ &\quad + O(h^3) = K^2(0)h + O(h^3).\end{aligned}$$

and therefore we have

$$\begin{aligned}E\Delta_-^{(0)}(t)\Delta_-^{(0)}(t+h) &= \int_{-C}^{0} K^2(u)\,\mathrm{d}u - \frac{1}{2}K^2(0)h \\ &\quad + \frac{1}{2}\left(K(0)K'(0) - \int_{-C}^{0} (K'(u))^2\,\mathrm{d}u\right)h^2 + O(h^3).\end{aligned}$$

Putting together (3.21)–(3.25) we have immediately that

$$E\Delta^{(i)}(t)\Delta^{(i)}(t+h) = A_0(i) + A_1(i)|h| + A_2(i)h^2 + o(h^2), \quad 0 \le i \le p.$$

Also, $E\Delta^{(i)}(t)\Delta^{(i)}(t+h) = 0$, if $t > C, 0 \le i \le p$. By the stationarity of $\Delta^{(i)}(t)$ we have

$$\sup_{C \le t \le 1/h - C} |\Delta^{(i)}(t)| \stackrel{\mathcal{D}}{=} \sup_{0 \le t \le 1/h - 2C} |\Delta^{(i)}(t)|. \tag{3.26}$$

Thus we use Pickands (1969) (cf also Chapter 12 of Leadbetter *et al.* (1983)) and obtain the following limit theorems:

(i) If $-\infty < A_1(i)/A_0(i) < 0$, then

$$\begin{aligned}\lim_{n\to\infty} P\Bigg\{&(2\log(1/h))^{1/2} \sup_{C \le t \le 1/h - C} |\Delta^{(i)}(t)| \le x + 2\log(1/h) \\ &+ \frac{1}{2}\log\log(1/h) + \log\left(\frac{1}{\pi^{1/2}}\left(-\frac{A_1(i)}{A_0(i)}\right)\right)\Bigg\} = \exp(-2e^{-x}).\end{aligned} \tag{3.27}$$

(ii) If $A_1(i) = 0$ and $-\infty < A_2(i)/A_0(i) < 0$, then

$$\lim_{n\to\infty} P\left\{(2\log(1/h))^{1/2} \sup_{C\le t\le 1/h-C} |\Delta^{(i)}(t)| \le x + 2\log(1/h) + \log\left(\frac{1}{2^{1/2}\pi}\left(-\frac{A_2(i)}{A_0(i)}\right)^{1/2}\right)\right\} = \exp(-2e^{-x}), \tag{3.28}$$

THEOREM 1.1 now follows from (3.20), (3.21) and (3.26)–(3.28).

Proof of Corollary 1.1 First we note that

$$b_-(0,0) = b_+(0,0) = 1\Big/\int_0^\infty K(x)\,\mathrm{d}x$$

and therefore

$$A_0(0) = \int_{-\infty}^\infty K^2(x)\,\mathrm{d}x\Big/\left(\int_0^\infty K(x)\mathrm{d}X\right)^2,$$

$$A_1(0) = -2K^2(0)\Big/\left(\int_0^\infty K(x)\,\mathrm{d}x\right)^2$$

and if $K(0) = 0$, then

$$A_2(0) = -\int_0^\infty (K'(x))^2\,\mathrm{d}x\Big/\left(\int_0^\infty K(x)\,\mathrm{d}x\right)^2.$$

Now Theorem 1.1 implies immediately Corollary 1.1.

Proof of Corollary 1.2 First we observe that

$$b_-(0,0) = b_+(0,0) = \frac{1}{D}\int_0^\infty x^2K(x)\,\mathrm{d}x$$

$$b_-(1,1) = b_+(1,1) = \frac{1}{D}\int_0^\infty K(x)\,\mathrm{d}x$$

$$b_-(0,1) = b_-(1,0) = \frac{1}{D}\int_0^\infty xK(x)\,\mathrm{d}x$$

and

$$b_+(0,1) = b_+(1,0) = -\frac{1}{D}\int_0^\infty xK(x)\,\mathrm{d}x,$$

where

$$D = \left(\int_0^\infty K(x)\,\mathrm{d}x\right)\left(\int_0^\infty x^2K(x)\,\mathrm{d}x\right) - \left(\int_0^\infty xK(x)\,\mathrm{d}x\right)^2.$$

Hence

$$\begin{aligned} A_0(0) = &\frac{2}{D^2}\left\{\left(\int_0^\infty x^2K(x)\,\mathrm{d}x\right)^2\int_0^\infty K^2(x)\,\mathrm{d}x\right. \\ &+\left(\int_0^\infty xK(x)\,\mathrm{d}x\right)^2\int_0^\infty x^2K^2(x)\,\mathrm{d}x \\ &\left.-2\int_0^\infty x^2K(x)\,\mathrm{d}x\int_0^\infty xK(x)\,\mathrm{d}x\int_0^\infty xK^2(x)\,\mathrm{d}x\right\}. \end{aligned}$$

and

$$A_1(0) = -\frac{2K^2(0)}{D^2}\left(\int_0^\infty x^2K(x)\,\mathrm{d}x\right)^2.$$

If $K(0) = 0$, then

$$\begin{aligned} A_2(0) = -\frac{1}{D^2}&\left\{\left(\int_0^\infty x^2K(x)\,\mathrm{d}x\right)^2\int_0^\infty (K'(x))^2\,\mathrm{d}x\right. \\ &+\left(\int_0^\infty xK(x)\,\mathrm{d}x\right)^2\int_0^\infty x^2(K'(x))^2\,\mathrm{d}x \\ &\left.-2\int_0^\infty x^2K(x)\,\mathrm{d}x\int_0^\infty xK(x)\,\mathrm{d}x\int_0^\infty x(K'(x))^2\,\mathrm{d}x\right\}. \end{aligned}$$

Similarly,

$$\begin{aligned} A_0(1) = &\frac{2}{D^2}\left\{\left(\int_0^\infty xK(x)\mathrm{d}x\right)^2\int_0^\infty K^2(x)\mathrm{d}x\right. \\ &+\left(\int_0^\infty K(x)\mathrm{d}x\right)^2\int_0^\infty x^2K^2(x)\mathrm{d}x \\ &\left.-2\int_0^\infty K(x)\mathrm{d}x\int_0^\infty xK(x)\mathrm{d}x\int_0^\infty x(K'(x))^2\mathrm{d}x\right\}, \\ A_1(1) = &\ 0 \end{aligned}$$

and

$$\begin{aligned} A_2(1) = -\frac{1}{D^2}&\left\{\left(\int_0^\infty K(x)\mathrm{d}x\right)^2\int_0^\infty x^2(K'(x))^2\mathrm{d}x\right. \\ &-K(0)K'(0)\int_0^\infty xK(x)\mathrm{d}x \\ &\left.-2\int_0^\infty xK(x)\mathrm{d}x\int_0^\infty K(x)\mathrm{d}x\int_0^\infty x(K'(x))^2\mathrm{d}x\right\}. \end{aligned}$$

Using these formulas we can compute the constants in Theorem 1.1 and the proof of Corollary 1.2 is complete.

4 PROOFS OF THEOREMS 2.1–2

We assume for the sake of simplicity that $m^* = n\tau$ is an integer. Let $\Delta = f^{(p)}(\tau-) - f^{(p)}(\tau+)$,

$$\mathbf{U}_p^+(t) = (U_{0,p}^+(t), \ldots, U_{p,p}^+(t))^T, \quad 0 \le t \le C,$$

with

$$U_{i,p}^+(t) = \frac{1}{p!}\int_t^C s^{p+i}K(s)\mathrm{d}s$$

and

$$\mathbf{U}_p^-(t) = (U_{0,p}^-(t), \ldots, U_{p,p}^-(t))^T, \quad 0 \le t \le C,$$

with

$$U_{i,p}^-(t) = \frac{(-1)^{p+i}}{p!}\int_t^C s^{p+i}K(s)\mathrm{d}s.$$

LEMMA 4.1 *If the conditions of Theorem 2.1 are satisfied, then*

$$\max_{m\notin[m^*-nhC,m^*+nhC]} |H^{-1}E(\hat{\boldsymbol{\alpha}}(m) - \hat{\boldsymbol{\beta}}(m))| = O(h^{p+3/2}), \tag{4.1}$$

$$\begin{aligned}\max_{m^*\le m\le m^*+nhC} &\left|H^{-1}E(\hat{\boldsymbol{\alpha}}(m) - \hat{\boldsymbol{\beta}}(m)) - h^{p+1/2}\Delta B_-\boldsymbol{U}_p^-\left(\frac{m-m^*}{nh}\right)\right| \\ &= o(h^{p+1/2})\end{aligned} \tag{4.2}$$

and

$$\max_{m^*-nhC\le m\le m^*} \left|H^{-1}E(\hat{\alpha}(m) - \hat{\beta}(m)) - h^{p+1/2}\Delta B_+\boldsymbol{U}_p^+\left(\frac{m^*-m}{nh}\right)\right| = o(h^{p+1/2}). \tag{4.3}$$

Proof Since the change-point does not have any effect on (4.1), the result in (4.1) follows immediately from Lemmas 3.2 and 3.7. Let

$$\boldsymbol{\beta}^*(m) = \left(f(x_m), f^{(1)}(x_m), \ldots, \frac{1}{(p-1)!}f^{(p-1)}(x_m), \frac{1}{p!}f^{(p)}(x_m)\right)^T.$$

Lemma 3.7 yields

$$\max_{m^* \le m \le m^* + nhC} |H^{-1}\{E\hat{\boldsymbol{\alpha}}(m) - \boldsymbol{\beta}^*(m)| = O(h^{p+3/2}). \tag{4.4}$$

Following the proof of Lemma 3.2 we obtain that

$$\begin{aligned}
&HX_-^T(m)Q_-(m)[Y_-^*(m) - X_-(m)E\hat{\boldsymbol{\alpha}}(m)] \\
&= \Bigg(h^{-1/2} \sum_{m-m^*<k\le nhC} \Bigg[f_1(x_{m-k}) - \sum_{0\le i\le p-1} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \\
&\quad - \frac{1}{p!} f_2^{(p)}(x_m)(x_{m-k} - x_m)^p \Bigg] K((x_{m-k} - x_m)/h) \\
&\quad + h^{-1/2} \sum_{0<k\le m-m^*} \Bigg[f_2(x_{m-k}) - \sum_{0\le i\le p-1} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \\
&\quad - \frac{1}{p!} f_2^{(p)}(x_m)(x_{m-k} - x_m)^p \Bigg] K((x_{m-k} - x_m)/h), \ldots, h^{-p-1/2} \\
&\quad \times \sum_{m-m^*<k\le nhC} \Bigg[f_1(x_{m-k}) - \sum_{0\le i\le p-1} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \\
&\quad - \frac{1}{p!} f_2^{(p)}(x_m)(x_{m-k} - x_m)^p \Bigg] (x_{m-k} - x_m)^p K((x_{m-k} - x_m)/h) \\
&\quad + h^{-p-1/2} \sum_{0<k\le m-m^*} \Bigg[f_2(x_{m-k}) - \sum_{0\le i\le p-1} \frac{1}{i!} f^{(i)}(x_m)(x_{m-k} - x_m)^i \\
&\quad - \frac{1}{p!} f_2^{(p)}(x_m)(x_{m-k} - x_m)^p \Bigg] (x_{m-k} - x_m)^p K((x_{m-k} - x_m)/h) \Bigg)^T.
\end{aligned}$$

Taylor expansion gives

$$\begin{aligned}
&\max_{m^* \le m \le m^* + nhC} |HX_-^T(m)Q_-(m)\{Y_-^*(m) - X_-(m)E\hat{\alpha}(m)\} - \boldsymbol{L}(m)| \\
&\quad = o(nh^{p+1/2}),
\end{aligned} \tag{4.5}$$

where $\mathbf{L}(m) = (L_0(m), \ldots, L_p(m))^T$ with

$$\begin{aligned}
L_i(m) = h^{-i-1/2} \frac{1}{p!} \{ f_1^{(p)}(x_m) - f_2^{(p)}(x_m) \} \\
\times \sum_{m-m^*<k\le nhC} (x_{m-k} - x_m)^{i+p} K((x_{m-k} - x_m)/h).
\end{aligned}$$

Elementary verifications show that

$$\max_{m^* \le m \le m^* + nhC} |L_i(m) - nh^{p+1/2}\{f_1^{(p)}(x_m) - f_2^{(p)}(x_m)\}\boldsymbol{U}_p^-((m - m^*)(nh))|$$
$$= o(nh^{p+1/2}), \tag{4.6}$$

and therefore (4.2) follows from Lemma 3.1.

The proof of (4.3) is similar to that of (4.2) and is therefore omitted.

Proof of Theorem 2.1 (i) Using Lemma 4.1 and (2.5) we obtain that the same numerical term is negligible in $n^{1/2}h^{i+1/2}Z_i(n)$, if $0 \le i \le p-1$. Hence Lemmas 3.3–3.5 and 3.8–3.10 imply the first part of Theorem 2.1.

(ii) By Lemmas 3.3–3.5, 3.8–3.10 and Lemma 4.1 we get that the random term is smaller in $\hat{\boldsymbol{\alpha}}(n\tau) - \hat{\boldsymbol{\beta}}(n\tau)$ than the numerical term. The last element of $B_-\boldsymbol{U}_p^-(0)$ (and $B_+\boldsymbol{U}_p^+(0)$) is 1, so (2.7) follows from (4.2).

Proof of Theorem 2.2 It follows from the proof of Theorem 2.1 that

$$\sup_{nhC \le m \le m^* - nhC} |\hat{\alpha}_p(m) - \hat{\beta}_p(m)| = o_P(1)$$

and

$$\sup_{m^* + nhC \le m \le n - nhC} |\hat{\alpha}_p(m) - \hat{\beta}_p(m)| = o_P(1).$$

Hence using (2.6) we get

$$\lim_{n \to \infty} P\{m^* - nhC \le n\hat{\tau} \le m^* + nhC\} = 1,$$

which gives (2.7).

Acknowledgements

We are grateful to the four referees for the careful reading of our paper and pointing out several inacuracies. One of the referees informed us that Dr Zouhir Hamrouni (Université Joseph Fourier de Grenoble, Grenoble, France) obtained similar results in an unpublished doctoral thesis. He considered a regression model allowing for random design and heteroskedastic errors. We were not aware of this research and our results were independently obtained.

Research partially supported by a NATO Scientific and Environmental Division Grant CRG 960503.

References

Brodsky, B.E. and Darkhovsky, B.S. (1993) *Nonparametric Methods in Change-Point Problems*. Kluwer, Dodrecht.

Cheng, K.F. and Lin, P.E. (1981) Nonparametric Estimation of a Regression Function, *Z. Wahrschein. verw. Gebiete*, **57**, 223–233.

Clark, R.M. (1977) Nonparametric Estimation of a Regression Function, *J. Royal Statist. Soc. Ser. B*, **39**, 107–113.

Csörgö, M. and Horváth, L. (1997) *Limit Theorems in Change-Point Analysis*. Wiley, Chichester.

Fan, J. and Gijbels, I. (1996) *Local Polynomial Modelling and its Applications*. Chapman and Hall, London.

Fernique, X. (1975) Regularité des trajectoires des fonctions aléatoires gaussiénnes. In *Ecole d'Ete de Probabilites de Saint–Flour IV–1974*, pp. 1–96. Springer-Verlag, Berlin.

Gasser, T. and Müller, H.G. (1979) *Kernel Estimators of Regression Functions*. Springer-Verlag, Berlin.

Hastie, T. and Loader, C. (1993) Local Regression: Automatic Kernel Carpentry (with Discussion), *Statistical Science*, **8**, 120–143.

Hušková, M. (1999) Estimators in Location Model with Gradual Change, *J. Stat. Plan. Inf.*, **76**, 109–125.

Jarušková, D. (1998) Testing Apperance of Linear Trend, *J. Stat. Plan. Inf.*, **70**, 263–276.

Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes*. Springer, New York.

Loader, C.R. (1996) Change Point Estimation using Nonparametric Regression, *Ann. Stat.*, **25**, 1667–1678.

Müller, H.G. (1992) Change-points in Nonparametric Regression Analysis. *Ann. Stat.*, **20**, 737–761.

Pickands, J. (1969) Upcrossing Probabilities for Stationary Gaussian Processes, *Trans. Amer. Math. Soc.*, **145**, 51–74.

Priestley, M.B. and Chao, M.T. (1972) Nonparametric Function Fitting, *J. Royal. Statist. Soc. Ser. B*, **34**, 385–396.

Stadtmüller, U. (1986) Asymptotic Properties of Nonparametric Curve Estimates, *Periodica Mathematica Hungarica*, **17**, 83–108.

Wu, J.S. and Chu, C.K. (1993) Kernel-type Estimators of Jump Points and Values of a Regression Function, *Ann. Stat.*, **21**, 1545–1566.

Yao, Q. (1993) Boundary-crossing Probabilities of Some Random Fields Related to Likelihood Ratio Tests for Epidemic Alternatives, *J. Appl. Prob.*, **30**, 52–65.

Yao, Q. (1993) Tests for Change-points with Epidemic Alternatives, *Biometrika*, **80**, 179–191.

Copyright © 2003 EBSCO Publishing