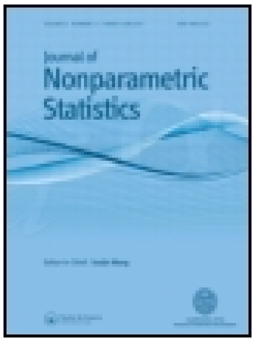

Taylor & Francis
Taylor & Francis Group

Journal of Nonparametric Statistics

ISSN: 1048-5252 (Print) 1029-0311 (Online) Journal homepage: http://www.tandfonline.com/loi/gnst20

Monotone Regression and Density Function Estimation at a Point of Discontinuity

D. Anevski & O. Hössjer

Nonparametric Statistics, 2002, Vol. 14(3), pp. 279–294

MONOTONE REGRESSION AND DENSITY FUNCTION ESTIMATION AT A POINT OF DISCONTINUITY

D. ANEVSKI[a]* and O. HÖSSJER[a]

[a]*Centre for Mathematical Sciences, Lund University, P.O. Box 118, Se22100 Lund, Sweden.*

(Received 20 April 2000; Revised 17 May 2001; In final form 23 October 2000)

Pointwise limit distribution results are given for the isotonic regression estimator at a point of discontinuity. The cases treated are independent data, ϕ- and α-mixing data and subordinated Gaussian long range dependent data. Pointwise limit results for the nonparametric maximum likelihood estimator of a monotone density are given at a point of discontinuity, for independent data. The limit distributions are non-standard and differ from the ones obtained for differentiable regression and density functions.

Keywords: Limit distributions; Density estimation; Regression function estimation; Discontinuity; Dependence; Monotone

AMS 1991 Subject Classification: Primary: 62E20, 62G07; Secondary: 60G18

1 INTRODUCTION

This paper deals with two nonparametric estimation problems: isotonic regression and nonparametric maximum likelihood estimation of a monotone density. We consider a regression model

$$y_i = m(t_i) + \epsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

with $m : (0, 1) \mapsto \mathbb{R}$ increasing, $t_i = i/n$ equidistant design points and $\{\epsilon_i\}$ a stationary sequence of error terms with $E(\epsilon_i) = 0$ and $\mathrm{Var}(\epsilon_i) = \sigma^2$. Define the partial sum process

$$x_n = \text{linear interpolation of} \left\{ \left(t_i + \frac{1}{2n}, \sum_{j=1}^{i} y_j \right)_{i=0}^{n} \right\}, \tag{2}$$

*Corresponding author.

ISSN 1048-5252 print; ISSN 1029-0311X online
DOI: 10.1080/10485250290000806

where $t_0 = 0$ is assumed. The isotonic regression estimator

$$\hat{m} = \text{argmin}\left\{\sum_{i=1}^{n}(y_i - v(t_i))^2\colon v \text{ increasing on } (0,1)\right\} \tag{3}$$

of Brunk [2] is not uniquely defined. However, if we require v to be piecewise constant on all intervals $(t_i - 1/(2n), t_i + 1/(2n))$, the solution is given by

$$\hat{m} = T_{[0,1]}(x_n)',$$

where T_J maps a function defined on an interval J to its greatest convex minorant,

$$T_J(x) = \sup\{v;\ v: J \mapsto \mathbb{R},\ v \text{ convex and } v \le x\}.$$

Note that we use the convention $T(x)'(t) = T(x)'(t+)$.

Asymptotic results for $\hat{m}(t_0)$ at a fixed interior point $t_o \in (0,1)$ were considered by Brunk in [3] when $m'(t_0) > 0$ and $\{\epsilon_i\}$ are independent. Then Wright [13] (cf. also Leurgans [9]) generalized Brunk's results to the case

$$m(t) = m(t_0) + a\,\text{sgn}(t - t_0)|t - t_0|^{p-1} + o\big(|t - t_0|^{p-1}\big) \tag{4}$$

as $t \to t_0$ and $1 < p < \infty$. A special case of Theorem 1 in Wright [13] is

$$Cn^{(p-1)/(2p-1)}(\hat{m}(t_0) - m(t_0)) \stackrel{\mathcal{L}}{\to} T(B(s) + |s|^p)'(0), \tag{5}$$

with $C = (p(\sigma^{2p-2}a)^{-1})^{1/(2p-1)}$, $T = T_{\mathbb{R}}$ and B a twosided standard Brownian motion. Thus the local behaviour of m around t_0 (*i.e.* the choice of p) influences both the convergence rate and the limit distribution. The case $p = 2$ corresponds to the classical result of Brunk [3]. Moreover, when $p = 2$, the right hand side of (5) can be replaced by

$$2\ \text{argmin}_{s\in\mathbb{R}}\big(s^2 + B(s)\big),$$

where we use the convention that $\text{argmin}_{s\in\mathbb{R}}(x(s))$ is the infimum of all points in $\mathbb{R}$ at which the minimum of x is attained. For $p = \infty$ the convergence rate is $n^{-1/2}$, cf. Parsons [10], Groeneboom and Pyke [8] and Dykstra and Carolan [4].

In Anevski and Hössjer [1], a general asymptotic scheme for isotonic functional estimation was treated. Theorem 3 in Anevski and Hössjer [1] admits generalization of (5) to weakly dependent mixing data as well as long range dependent subordinated Gaussian data. In fact, a somewhat weaker formulation of (4), which suffices for (5) to be deduced as a special case of Theorem 3 in Anevski and Hössjer [1], is

$$\delta^{-p}(M(t_0 + s\delta) - M(t_0) - m(t_0)s\delta) \to A|s|^p, \tag{6}$$

uniformly for s on compact sets, as $\delta \to 0$, with $A = a/p$ and $p > 1$, where $M(t) = \int_0^t m(u)\,du$ is the primitive function of m. One can show that the fraction of data that $\hat{m}$ uses is of the order

$$d_n = n^{-1/(2p-1)}, \tag{7}$$

in the sense that data in a shrinking neighbourhood of t_0 with length proportional to $n^{-1/(2p-1)}$ determines the limit distribution. If m grows faster locally around t_0, a smaller proportion of data is used in determining the large sample distributional properties of $\hat{m}(t_0)$; this can be interpreted as $d_n = d_n(t_0)$ being a spatially adaptive bandwidth and $\hat{m}$ smooth less at points of large increase of m.

In the present paper we extend the results in Anevski and Hössjer [1] to cover the case when m is not continuous at the point t_0, *i.e.* when $m(t_0-), m(t_0+)$ both exist and differ (m is monotone and thus right and left hand limits exist at each point); this is the case $p = 1$ with $2A = m(t_0+) - m(t_0-)$ in (6). Since it is of interest to see how $\hat{m}$ is influenced by the discontinuity at t_0 for points close to t_0, we will study the estimate on a local scale, *i.e.* we study $\hat{m}(t_0 + s_0 n^{-1})$ for a fixed $s_0 \in \mathbb{R}$ (note that formally $d_n = n^{-1}$ in (7) if $p = 1$).

Suppose now that $\{t_i\}_{i=1}^n$ is a stationary sequence of random variables with marginal distribution function F and density $f = F'$. We assume that f is supported on [0, 1] and increasing. In Grenander [6] it was first shown that for independent data, the nonparametric maximum likelihood estimator (NPMLE)

$$\hat{f} = \text{argmax}\left\{\prod_{i=1}^{n} v(t_i) : v \text{ increasing}, v \geq 0, \int_0^1 v(t)\,\mathrm{d}t = 1\right\},$$

is given by $\hat{f} = T_{[0,1]}(F_n)'$ and

$$F_n(t) = \frac{1}{n}\sum_{i=1}^{n} 1_{\{t_i \leq t\}} \tag{8}$$

the empirical distribution function formed by $\{t_i\}$. We are not aware of any limit distribution results for $T_{[0,1]}(F_n)'$ in the i.i.d. data case when the restriction on F is

$$\delta^{-p}(F(t_0 + s\delta) - F(t_0) - f(t_0)s\delta) \to A|s|^p, \tag{9}$$

as $\delta \to 0$ when $1 < p < \infty$, other then the classical case $p = 2$ treated in Prakasa Rao [12] and Groeneboom [7]. However, it can be deduced from Theorem 2.1 in Leurgans [9] and Section 3 and 5 in Anevski and Hössjer [1] that if $1 < p < \infty$,

$$Cn^{(p-1)/(2p-1)}\left(\hat{f}(t_0) - f(t_0)\right) \xrightarrow{\mathcal{L}} T_{\mathbb{R}}(|s|^p + B(s))'(0), \tag{10}$$

as $n \to \infty$, with $t_0 > 0$ fixed and $C = A^{-1/(2p-l)} f(t_0)^{-(p-1)/(2p-1)}$. In particular, when $p = 2$ and $A = f'(t_0)/2$, (10) reduces to the result obtained in Prakasa Rao [12].

In this paper, we consider the case when $f(t_0-)$ and $f(t_0+)$ both exist and are different *i.e.* $p = 1$ in (9). It is of interest to note that $T(F_n)'$ is the NPMLE of a monotone density when data are i.i.d. no matter what the restrictions on the unknown density f are, and thus also in the present case.

When $p = 1$, we obtain "convergence rate" $n^0 = 1$ in (5) and (10), *i.e.* $\hat{m}(t_0)$ and $\hat{f}(t_0)$ are not consistent. Further, the Brownian motion must be replaced by other stochastic processes. The reason is that the effective number nd_n of data points used for computing $\hat{m}(t_0)$ and $\hat{f}(t_0)$ does not increase with n.

The article is organized as follows: In Section 2 we state the main results of the paper *i.e.* the pointwise, at a point of discontinuity, limit distribution results for the isotonic regression function estimate of an increasing regression function (Theorem 1) and for the NPMLE of an increasing density (Theorem 2), respectively. In Section 3 we give a unified derivation of the limit distributions of $T(x_n)$ and $T(x_n)'$ for a large class of stochastic processes x_n, which equals F_n for density estimation and the partial sum process, defined in (2), for regression. These results are next applied to the proofs of Theorem 1 and 2. Finally we have collected some technical results in the appendix.

The results in this paper depend on a truncation result, Theorem 1 in Anevski and Hössjer [1]. When referring to assumptions in Anevski and Hössjer [1], we denote *e.g.* **A1** in Anevski and Hössjer [1] as **B1**.

2 MAIN RESULTS

In this section we give pointwise limit distribution results for the isotonic regression estimator and its primitive function at a point of discontinuity and under various dependence assumptions on the error terms. Furthermore we state the pointwise limit distribution for the NPMLE of a monotone density at a point of discontinuity in the independent data case.

For $I \subset \mathbb{R}$ an arbitrary interval, define $D(I)$ as the set of real valued functions on I that are right continuous with left hand limits. We equip $D(I)$ with the supnorm metric when I is compact, and $D(-\infty, \infty)$ with the supnorm metric on compact intervals. To avoid measurability problems for processes in $D(-\infty, \infty)$, we use the σ-algebra generated by the open balls in this metric, cf. Pollard [11].

We will give a somewhat unified approach to these estimation problems and thus we write either of the partial sum process defined in (2) and the empirical distribution function defined in (8) as

$$x_n(t) = x_{b,n}(t) + v_n(t), \qquad t \in J \tag{11}$$

where J is the domain of x_n, and $x_n, v_n \in D(J)$. Here v_n is a sequence of stochastic processes, and $x_{b,n}$ is a sequence of deterministic functions. Furthermore, given a sequence $d_n \downarrow 0$, we will rescale the stochastic part of x_n locally around an interior point t_0 of J, according to

$$\tilde{v}_n(s) = \tilde{v}_n(s; t_0, s_0) = d_n^{-1}(v_n(t_0 + (s + s_0)d_n) - v_n(t_0 + s_0 d_n)), \tag{12}$$

and the deterministic part of x_n according to

$$g_n(s) = d_n^{-1}\big(x_{b,n}(t_0 + (s + s_0)d_n) - x_{b,n}(t_0 + s_0 d_n) - s d_n \bar{x}'_{b,n}(t_0)\big), \tag{13}$$

for s_0 fixed and $s + s_0 \in J_{n,t_0} = d_n^{-1}(J - t_0)$, and with $\bar{x}'_{b,n}(t_0) = (x'_{b,n}(t_0-) + x'_{b,n}(t_0+))/2$.

Consider again the regression model (1). The partial sum process defined in (2) can also be written

$$x_n(t) = \int_{1/(2n)}^{t} \tilde{y}_n(u)\,\mathrm{d}u = x_{b,n}(t) + v_n(t),$$

with

$$\tilde{y}_n(t) = y_n, \quad t \in \left(t_i - \frac{1}{2n}, t_i + \frac{1}{2n}\right], \quad i = 1, \ldots, n,$$

$$x_{b,n}(t) = \int_{1/(2n)}^{t} \tilde{m}_n(u)\,\mathrm{d}u,$$

$$v_n(t) = \int_{1/(2n)}^{t} \tilde{\epsilon}_n(u)\,\mathrm{d}u,$$

and $\tilde{m}_n$, $\tilde{\epsilon}_n$ defined as $\tilde{y}_n$ with $\{y_i\}_{i=1}^n$ replaced by $\{m(t_i)\}_{i=1}^n$ and $\{\epsilon_i\}_{i=1}^n$ respectively. Notice that $x_{b,n}$ is convex, since m is increasing.

As noted in Section 1, $T_{[0,1]}(x_n)'$ is the solution (3) to the isotonic regression problem. Classically pointwise limit distribution results have been shown at points when m is continuous, such as in Eq. (4). Instead, we will assume:

A1 *The regression function m is right continuous and increasing on* [0,1], *and* t_0 *is a fixed point in* (0,1) *such that*

$$m(t_0+) - m(t_0-) = 2A > 0.$$

The behaviour of $\hat{m}$ close to t_0 depends critically on the positioning of the surrounding design points $t_i < t_0 \leq t_{i+1}$. Since $n(t_i - t_0)$ is not convergent, we will not center our local scale at $t_0 + s_0 n^{-1}$, but rather at $\tilde{t}_0 + s_0 n^{-1}$, where

$$\tilde{t}_0 = \tilde{t}_0(n) = \max\{t_i : t_i < t_0\} + \frac{1}{2n}.$$

Thus the rescaled deterministic part of x_n is

$$\begin{aligned} g_n(s) &= d_n^{-1}\big(x_{b,n}\big(\tilde{t}_0 + (s+s_0)d_n\big) - x_{b,n}\big(\tilde{t}_0 + s_0 d_n\big) - s d_n \bar{x}'_{b,n}\big(\tilde{t}_0\big)\big) \\ &= \int_0^s \big(\tilde{m}_n\big(\tilde{t}_0 + (u+s_0)d_n\big) - \bar{m}_n\big)\,\mathrm{d}u \end{aligned}$$

with $\bar{m}_n = (m(\tilde{t}_0 - 1/(2n)) + m(\tilde{t}_0 + 1/(2n)))/2$, and the rescaled stochastic part

$$\begin{aligned} \tilde{v}_n(s) &= d_n^{-1}\big(v_n\big(\tilde{t}_0 + (s+s_0)d_n\big) - v_n\big(\tilde{t}_0 + s_0 d_n\big)\big) \\ &= \int_0^s \tilde{\epsilon}_n\big(\tilde{t}_0 + (u+s_0)d_n\big)\,\mathrm{d}u. \end{aligned}$$

Assume we can extend the error terms to a doubly infinite stationary sequence $\{\epsilon_i\}_{i=-\infty}^{\infty}$. Then with

$$\tilde{\epsilon}_1(t) = \epsilon_i, \quad t \in \left(i - \frac{1}{2}, i + \frac{1}{2}\right], \quad i \in \mathbf{Z},$$

our discrete analogue of the Brownian motion becomes

$$Z(t) = \frac{1}{\sigma}\int_0^t \tilde{\epsilon}_1\left(u + \frac{1}{2}\right) \mathrm{d}u,$$

which is defined by linear interpolation between the points $(i, -\Sigma_{j=i}^{0}\epsilon_j/\sigma)_{i<0}$ and $(i, \Sigma_{j=1}^{i}\epsilon_j/\sigma)_{i\geq 0}$. Then if $d_n = n^{-1}$ it is straightforward to check that

$$\begin{aligned}\tilde{v}_n(s) &= \int_0^s \tilde{\epsilon}_1\big(n\tilde{t}_0 + u + s_0\big)\,\mathrm{d}u \stackrel{\mathcal{L}}{=} \int_0^s \tilde{\epsilon}_1\left(\frac{1}{2} + u + s_0\right)\mathrm{d}u \\ &= \sigma(Z(s+s_0) - Z(s_0)) =: \tilde{v}(s).\end{aligned} \tag{14}$$

Define $\bar{m} = (m(t_0-) + m(t_0+))/2$. In order to be able to state the limit results more compactly, define the function

$$\rho_{B,C}(s) = \begin{cases} Bs, & s > 0, \\ -Cs, & s < 0, \end{cases}$$

for $B, C > 0$, and denote $\rho_B = \rho_{B,B}$. The next theorem gives the limit distribution of the isotonic regression estimate and of it's primitive function, for an increasing regression function satisfying the discontinuity condition **A1**.

THEOREM 1 *Assume m satisfies* **A1**, $\{\epsilon_i\}$ *are independent and identically distributed, and* $E(\epsilon_i) = 0, \sigma^2 = Var(\epsilon_i) < \infty$. *Then the solution to the isotonic regression problem* $\hat{m} = T(x_n)'$ *satisfies*

$$\begin{aligned}&P\big(\hat{m}\big(\tilde{t}_0 + s_0 n^{-1}\big) - \tilde{m} < a\big) \\ &\quad\to\ P\Big(\mathrm{argmin}_{s\in\mathbb{R}}\Big[\rho_{(A-a)/\sigma,(A+a)/\sigma}(s+s_0) + Z(s+s_0)\Big] > 0\Big),\end{aligned}$$

as $n \to \infty$, *for each* $|a| < A$ *such that the limit is continuous (viewed as a function of a). Further*

$$\begin{aligned}&n\left(\int_0^{\tilde{t}_0 + s_0 n^{-1}} \hat{m}(u)\,\mathrm{d}u - x_n\big(\tilde{t}_0 + s_0 n^{-1}\big)\right) \\ &\quad\stackrel{\mathcal{L}}{\to}\ T[A((|s+s_0| - |s_0|)) + \sigma(Z(s+s_0) - Z(s_0))](0),\end{aligned}$$

as $n \to \infty$, *provided* $|a| < A$.

When $\{\epsilon_i\}$ are dependent, $\tilde{v}_n$ is still given by (14), although $Z(\cdot)$ has a different distribution then. It is possible to establish Theorem 3 for weakly dependent and subordinated Gaussian long range dependent data as well. The main technical difficulty is to verify Eq. (29) in the Appendix, which is done in Appendix II of Anevski and Hössjer [1].

Next we will treat monotone density estimation. Assume that $\{t_i\}_{i=-\infty}^{\infty}$ is an i.i.d. sequence of random variables with distribution function F. For the density function $f = F'$ we assume the following:

A2 *The density function f is increasing on* [0,1], *and t_0 is a fixed point in* (0,1) *such that*

$$f(t_0+) - f(t_0-) = 2A > 0.$$

As mentioned in Section 1, the NPMLE is $\hat{f} = T(F_n)'$, where F_n is the empirical distribution function defined in (8). Thus we put $x_n = F_n$. Further, assume $x_{b,n} = x_b = F$, so that

$$v_n(t) = F_n(t) - F(t) = \frac{1}{n}\sum_{i=1}^{n}\big(1_{\{t_i \le t\}} - F(t)\big),$$

is the centered empirical process.

Notice that with $d_n = n^{-1}$, the rescaled stochastic part is

$$\begin{aligned}\tilde{v}_n(s) = \sum_{i=1}^{n}\big(&1_{\{t_i \le t_0 + (s+s_0)n^{-1}\}} - 1_{\{t_i \le t_0 + s_0 n^{-1}\}}\\ &-F\big(t_0 + (s+s_0)n^{-1}\big) + F\big(t_0 + s_0 n^{-1}\big)\big).\end{aligned}$$

Let $s \mapsto N(s)$ be a twosided Poisson process with constant intensity 1, *i.e.* if $\{Y_i\}_{-\infty}^{\infty}$ are i.i.d. Exp(1) random variables we put $T_i = \Sigma_{j=1}^{i} Y_j$ if $i \ge 0$, $T_i = -\Sigma_{j=I}^{1} Y_j$ if $i < 0$ and

$$N(t) = \begin{cases} \Sigma_{i=1}^{\infty} 1_{\{T_i \le t\}}, & t \ge 0,\\ -\Sigma_{i=1}^{\infty} 1_{\{T_{-i} > t\}}, & t < 0.\end{cases}$$

Define the martingale

$$N_0(s) = N(s) - s$$

as the centered version of $N(\cdot)$. Then from *e.g.* Corollary 2.2 in Einmahl [5], it follows that

$$\tilde{v}_n(s) \xrightarrow{\mathcal{L}} \tilde{v}(s) := N_0\bigg(\int_0^s \lambda(u)\,\mathrm{d}u\bigg), \tag{15}$$

as $n \to \infty$, where

$$\lambda(u) = \begin{cases} f(t_0+), & u > -s_0,\\ f(t_0-), & u < -s_0.\end{cases}$$

Let $\bar{f} = (f(t_0-) + f(t_0+))/2$. The next result gives the limit distribution for the NPMLE, and it's primitive function, of an increasing density satisfying the discontinuity condition **A2**.

THEOREM 2 *Let $\{t_i\}_{i\ge1}$ be an i.i.d. sequence with a marginal density function f satisfying* **A2**. *Then if $F_n(t)$ is the empirical distribution function and $\hat{f}_n(t) = T(F_n)'(t)$,*

$$\begin{aligned}&P\{\hat{f}_n\big(t_0 + s_0 n^{-1}\big) - \bar{f} < a\}\\ &\quad\to P\{\mathrm{argmin}_{s\in\mathbb{R}}\big[\rho_{A-a,A+a}(s+s_0) + \tilde{v}(s)\big] > 0\},\end{aligned}$$

as $n \to \infty$, *for each* $|a| < A$ *such that the limit is continuous (viewed as a function of* a*). Further*

$$n\left(\int_0^{t_0+s_0n^{-1}} \hat{f}_n(u)\,\mathrm{d}u - F_n\big(t_0+s_0n^{-1}\big)\right) \xrightarrow{\mathcal{L}} T(A(|s+s_0|-|s_0|)+\tilde{v}(s))(0),$$

as $n \to \infty$, *for each* $s_0 \in \mathbb{R}$ *with* $\tilde{v}$ *defined in* (15).

Theorem 2 can also be established for weakly dependent data, since the local Poisson behaviour of F_n (on a scale n^{-1}) can be proved in this case as well.

We would here shortly like to discuss why the discontinuities of the target functions (m and f) give essentially more different limit results in Theorems 1 and 2 than in the regular cases. For instance for the regression problem, in the regular case (*i.e.* when the primitive function of the target function satisfies (6) with $p > 1$) the isotonic functional estimation $\hat{m}$ uses data in a shrinking neighbourhood of t_0 of length $d_n = n^{-1/(2p-1)}$, as noted in the introduction, and which follows from Lemma 2 in Section 3; this means that $\hat{m}$ is a local estimator analogously to a kernel regression estimator. Thus in the regular case, we have $d_n \gg n^{-1}$ and the number of data in this shrinking interval converges to ∞. Since $\hat{m}$, for a finite n, is a (nonlinear) functional of the partial sum process, it will follow that the limit functional in the regular case is a nonlinear functional of a Brownian motion (for the independent data case), cf. Eq. (5). As a contrast, in the discontinuous case, *i.e.* when $p = 1$ and thus $d_n = n^{-1}$, the amount of data in the shrinking interval does not increase; thus the functional central limit theorem is not in force to obtain a Brownian motion as a limit of the partial sum process and the limit random variable will not be a functional of a Brownian motion, cf. Theorem 4 of Section 3. The fact that $\hat{m}$ (and $\hat{f}$) is not consistent is also explained by the fact that the amount of data does not increase in the shrinking interval determining the limit distribution of the estimator. Furthermore the deterministic part $\rho_{(A-a)/\sigma,(A+a)/\sigma}$ (and $\rho_{A-a,A+a}$) in the limit random variable is simply the limit of the rescaled deterministic part g_n (on the same scale as the rescaled process part $\tilde{v}_n$), and the derivative of ρ is essentially the (rescaled) target function, which explains the form of the deterministic part ρ.

3 DERIVATION OF LIMIT DISTRIBUTIONS

In this section we prove limit distributions for $T(x_n)$ and $T(x_n)'$ for processes x_n satisfying (11) at a point t_0 of discontinuity of $x_{b,n}$. The main results (Theorems 3 and 4) are obtained under somewhat more general assumptions than strictly necessary for the proof of Theorems 1 and 2, to allow for possible future work on *e.g.* other forms of dependence. Next we apply the obtained results to the proofs of Theorems 1 and 2, *i.e.* to the cases x_n equal to the partial sum process and x_n equal to the empirical distribution function, respectively.

Denote $T_c = T_{[-c,c]}$ and $T = T_{\mathbb{R}}$. The next two assumptions are the main restrictions on the stochastic and deterministic part of x_n respectively, and are sufficient to imply local limit distribution results for $T(x_n)$ and $T(x_n)'$.

A3 [Weak convergence of rescaled stochastic term] *Assume there exists a stochastic process* $\tilde{v}(\cdot) = \tilde{v}(\cdot; t_0, s_0) \neq 0$ *such that*

$$\tilde{v}_n(s) \stackrel{\mathcal{L}}{\to} \tilde{v}(s),$$

on $D(-\infty, \infty)$ *as* $n \to \infty$.

The limit process $\tilde{v}$ corresponds *e.g.* to the increment in the interpolated random walk $\sigma(Z(s+s_0) - Z(s_0))$ in (14), for the regression problem, or to the centered nonhomogenous Poisson process $N_0\big(\int_0^s \lambda(u)\,\mathrm{d}u\big)$ in (15), for the density estimation problem.

A4 [Local uniform convergence of rescaled bias term] *Assume that* $\{x_{b,n}\}_{n\geq 1}$ *are convex functions. Assume that, with* g_n *defined in* (13), *for some* $A > 0$ *and each* $c > 0$,

$$\sup_{|s|\leq c} |g_n(s) - A(|s+s_0| - |s_0|)| \to 0, \tag{16}$$

as $n \to \infty$.

We remark that for most of our applications there exists a function x_b such that

$$x_b(t) = \lim_{t\to\infty} x_{b,n}(t), \qquad t \in J,$$

and

$$x_b'(t_0+) - x_b'(t_0-) = 2A > 0,$$

(cf. **A1** and **A2**). To see the connection with (16), assume for simplicity that $x_{b,n} = x_b$. Then

$$\begin{aligned} g_n(s) &= d_n^{-1} \int_{t_{0+s_0 d_n}}^{t_0+(s+s_0)d_n} \big(x_b'(u) - \bar{x}_b'(t_0)\big)\,\mathrm{d}u \\ &= d_n^{-1} \int_{t_0+s_0 d_n}^{t_0+(s+s_0)d_n} A\,\mathrm{sgn}(u-t_0)\,\mathrm{d}u + o(1) \\ &= A(|s+s_0| - |s_0|) + o(1), \end{aligned}$$

as $n \to \infty$.

The next lemma is the local limit distribution result for $T(x_n)$. Define the greatest convex minorant on a shrinking interval $T_{c,n} = T_{[t_0+(s_0-c)d_n, t_0+(s_0+c)d_n]}$.

LEMMA 1 *Let* $t_0 \in J$ *be fixed, and assume that* **A3** *and* **A4** *hold. Then for each* $c > 0$

$$\begin{aligned} &d_n^{-1}\big[T_{c,n}(x_n)(t_0+s_0 d_n) - x_n(t_0+s_0 d_n)\big] \\ &\qquad \stackrel{\mathcal{L}}{\to} T_c[A(|s+s_0| - |s_0|) + \tilde{v}(s)](0), \end{aligned}$$

as $n \to \infty$, *with* A *as in* **A4**.

Lemma 1 is proved in the Appendix. The next restriction on the tail behaviour of $\tilde{v}_n$ is assumption **B6** in Anevski and Hössjer [1]: it is used in Lemma 2 below to show that the local limit distribution result implies the global limit distribution, *i.e.* that truncation does not matter.

A5 [**Tail behaviour of rescaled process**] *The function*

$$y_n(s) = g_n(s) + \tilde{v}_n(s), \tag{17}$$

satisfies: given $\delta > 0$ *there are finite* $\tau = \tau(\delta) > 0$ *and* $\kappa = \kappa(\delta) > 0$ *such that*

$$\liminf_{n\to\infty} P\left(\inf_{|s|>\tau}(y_n(s) - \kappa|s|) > 0\right) > 1 - \delta,$$

and given $\epsilon, \delta, \tilde{\tau} > 0$

$$\limsup_{n\to\infty} P\left(\inf_{\tilde{\tau}\le s\le c}\frac{y_n(s)}{s} - \inf_{\tilde{\tau}\le s}\frac{y_n(s)}{s} > \epsilon\right) < \delta,$$
$$\limsup_{n\to\infty} P\left(\inf_{-c\le s\le -\tilde{\tau}}\frac{y_n(s)}{s} - \inf_{s\le -\tilde{\tau}}\frac{y_n(s)}{s} < -\epsilon\right) < \delta,$$

for all large enough $c > 0$.

The following result uses Theorem 1 in Anevski and Hössjer [1] to show that the difference between the local map $T_{c,n}$ and global map T_J diminishes as first n and then c grows.

LEMMA 2 *Assume that* **A3**, **A4** *and* **A5** *hold. Define* $A_{n,\Delta} = [t_0 - \Delta d_n, t_0 + \Delta d_n]$. *Then for each finite* $\Delta > 0$ *and* $\epsilon > 0$

$$\lim_{c\to\infty}\lim_{n\to\infty}\mathbf{P}\left[\sup_{A_{n,\Delta}} d_n^{-1}|T_{c,n}(x_n)(\cdot) - T_J(x_n)(\cdot)| \le \epsilon\right] = 1,$$
$$\lim_{c\to\infty}\lim_{n\to\infty}\mathbf{P}\left[\sup_{A_{n,\Delta}} |T_{c,n}(x_n)'(\cdot) - T_J(x_n)'(\cdot)| \le \epsilon\right] = 1.$$

Lemma 2 is proved in the Appendix. The next assumption is the analog of assumption **A5** for the limit process, (cf. Proposition 1 in the Appendix).

A6 [**Tail behaviour of limit process**] *For each* $\epsilon, \delta > 0$ *there is a* $\tau = \tau(\epsilon, \delta) > 0$, *so that*

$$P\left(\sup_{|s|\ge\tau}\frac{|\tilde{v}(s)|}{|s|} > \epsilon\right) \le \delta.$$

Finally, we arrive at the following global limit distribution result for $T(x_n)$.

THEOREM 3 *Let t_0 and s_0 be fixed and suppose* **A3–A6** *hold. Then*

$$d_n^{-1}[T_J(x_n)(t_0 + s_0 d_n) - x_n(t_0 + s_0 d_n)] \xrightarrow{\mathcal{L}} T[A(|s + s_0| - |s_0|) + \tilde{v}(s)](0), \tag{18}$$

with $A > 0$ as in **A4**, *as $n \to \infty$.*

Proof Clearly

$$\begin{aligned} &d_n^{-1}(T_J(x_n)(t_0 + s_0 d_n) - x_n(t_0 + s_0 d_n)) \\ &= d_n^{-1}\big[T_J(x_n)(t_0 + s_0 d_n) - T_{c,n}(x_n)(t_0 + s_0 d_n)\big] \\ &\quad + d_n^{-1}\big[T_{c,n}(x_n)(t_0 + s_0 d_n) - x_n(t_0 + s_0 d_n)\big]. \end{aligned}$$

Lemma 2 implies

$$d_n^{-1}\big(T_{c,n}(x_n)(t_0 + s_0 d_n) - T_J(x_n)(t_0 + s_0 d_n)\big) \xrightarrow{P} 0,$$

if we let $n \to \infty$. Lemma 1 implies that

$$d_n^{-1}\big(T_{c,n}(x_n)(t_0 + s_0 d_n) - x_n(t_0 + s_0 d_n)\big) \xrightarrow{\mathcal{L}} T_c[A(|s + s_0| - |s_0|) + \tilde{v}(s)](0)$$

as $n \to \infty$. Then use Theorem 1 in Anevski and Hössjer [1] with $y_n(s) = A(|s + s_0| - |s_0|) + \tilde{v}(s)$, to deduce

$$T_c(A(|s + s_0| - |s_0|) + \tilde{v}(s))(0) - T(A(|s + s_0| - |s_0|) + \tilde{v}(s))(0) \xrightarrow{P} 0$$

as $c \to \infty$. Notice that **B1** and **B2** of Theorem 1 in Anevski and Hössjer [1] follow from **A6** and the convexity of $A(|s + s_0| - |s_0|)$ as in the proof of Proposition 1 in Anevski and Hössjer [1] and **B3** follows from **A3** and (16). The proof is completed by applying Slutsky's theorem, first letting $n \to \infty$ and then $c \to \infty$. □

The next condition replaces an assumption on the continuity of the functional $x \mapsto T_c(x)'(0)$. This functional is not continuous on all of $D(-\infty, \infty)$, cf. Proposition 2 and Notes 2 and 3 in Anevski and Hössjer [1], and thus we cannot use the continuous mapping theorem together with **A3** and **A4** to obtain limit results for $T(x_n)'$.

A7 *Define y_n as in* (17). *Then*

$$T_c(y_n)'(0) \xrightarrow{\mathcal{L}} T_c(A|s + s_0| + \tilde{v}(s))'(0)$$

as $n \to \infty$ for each $c > 0$.

The next result is the global limit distribution result for $T(x_n)'$.

THEOREM 4 *Assume* **A3–A7** *hold. Then*

$$T(x_n)'(t_0 + s_0 d_n) - \bar{x}'_{b,n}(t_0) \xrightarrow{\mathcal{L}} T[A|s + s_0| + \tilde{v}(s)]'(0), \tag{19}$$

as $n \to \infty$ *with* A *and* $\bar{x}'_{b,n}(t_0)$ *as in* **A2**. *If further* $|a| < A$ *and*

$$P\big(T[A|s + s_0| + \tilde{v}(s)]'(0) = a\big) = 0,$$

then

$$\lim_{n\to\infty} P\{T(x_n)'(t_0 + s_0 d_n) - \bar{x}'_{b,n}(t_0) < a\}$$
$$= P\big(\text{argmin}_{s\in\mathbb{R}}\big[\rho_{A-a,A+a}(s + s_0) + \tilde{v}(s)\big] > 0\big).$$

Proof From (27) and (26) in the Appendix we obtain

$$T_{c,n}(x_n)'(t_0 + s_0 d_n) - \bar{x}'_{b,n}(t_0) = T_c(y_n)'(0),$$

with y_n as defined in (17). Now **A3** and **A4** imply

$$y_n(s) \xrightarrow{\mathcal{L}} A(|s + s_0| - |s_0|) + \tilde{v}(s) =: y(s), \tag{20}$$

on $D(I_c)$, where $I_c = [-c, c]$. Then **A7** implies

$$T_{c,n}(x_n)'(t_0 + s_0 d_n) - \bar{x}'_{b,n}(t_0) \xrightarrow{\mathcal{L}} T_c(y)'(0)$$

for each $c > 0$. This result can be extended to (19) in the same way as in Theorem 3 in Anevski and Hössjer [1], using Lemma 2, Theorem 1 of Anevski and Hössjer [1] and Slutsky's Theorem.

The second part of the theorem follows from the first and the fact that

$$\{T(y)'(0) < a\} = \{\text{argmin}_{s\in\mathbb{R}}(y(s) - as) > 0\}$$
$$= \{\text{argmin}_{s\in\mathbb{R}}\big[\rho_{A-a,A+a}(s + s_0) + \tilde{v}(s)\big] > 0\}.$$

□

Notice that

$$P(|T(x_n)'(t_0 + s_0 d_n) - \bar{x}'_{b,n}(t_0)| \leq a) \to C$$

as $n \to \infty$, for any $a > 0$, with $C = 1$ if $a \geq A$ but $C < 1$ if $0 < a < A$. Thus the variance of $T(x_n)'(\cdot)$ does not tend to zero as $n \to \infty$ at neighbourhoods of t_0 of size d_n.

Next we prove the main results, Theorems 1 and 2.

Proof (Theorem 1) We will show that the assumptions of Theorem 1 imply assumptions **A3–A7** of Theorems 3 and 4.

Since $\tilde{v}_n = \tilde{v}$ on J_{n,t_0}, **A3** holds. Lemma 4 in the Appendix shows that **A1** implies **A4**. Applying Lemma 6 in Anevski and Hössjer [1] with $m = 1$ (we can put $m_0 = 1$ in that lemma for independent data) we deduce that for each $\epsilon, \delta > 0$ there is a finite $\tau = \tau(\epsilon, \delta)$ such that

$$\limsup_{n\to\infty} P\left(\sup_{|s|\geq\tau}\left|\frac{\tilde{v}_n(s)}{g_n(s)}\right| > \epsilon\right) < \delta. \tag{21}$$

By Proposition 1 in the Appendix, this implies **A5**. Since $\tilde{v}_n \stackrel{\mathcal{L}}{=} \tilde{v}$, **A6** follows from **A5**. In order to check **A7**, notice that $T(y_n) \stackrel{\mathcal{L}}{=} T(\tilde{y}_n)$, where $\tilde{y}_n = g_n + \tilde{v}_n$. Further

$$\tilde{y}_n(s) - y(s) = g_n(s) - A(|s + s_0| - |s_0|) =: \psi_n(s), \tag{22}$$

with y_n defined in (20). From the proof of Lemma 4 in the Appendix, it follows that for each $\epsilon > 0$ we can find $\delta > 0$ such that

$$|s| \leq \delta \Rightarrow |\psi_n(s)| \leq \epsilon|s|$$

whenever $n \geq n_0(\epsilon, \delta)$. Defining $X_n = T_c(\tilde{y}_n)'(0)$ and $X = T_c(y)'(0)$ we obtain

$$\begin{aligned} P(X_n \leq y) &= \int P(X_n \leq y|X = x)\,\mathrm{d}F_X(x) \\ &\to \int 1_{\{x<y\}}\,\mathrm{d}F_X(x) = P(X \leq y), \end{aligned}$$

for every continuity point y of F_X, and thus **A7** holds. Thus Theorems 3 and 4 are in force and this ends the proof. □

Proof (Theorem 2) We will show that the assumptions of Theorem 2 imply assumptions **A3–A7**.

Since $x_{b,n} = x_b$, **A4** follows from **A2** and the remark after **A4**. Further **A3** follows from (15) and **A6** from properties of the Poisson process. In order to check **A7**, we introduce the functional $h(z) = T(z)'(0)$, for which the set of discontinuities is

$$D_h = \{z : T(z)'(0) > T(z)'(0-)\}.$$

By **A3** and **A4**, $y_n \stackrel{\mathcal{L}}{\to} y$ on $D(-\infty, \infty)$ and

$$y(s) = N\left(\int_0^s \lambda(u)\,\mathrm{d}u\right) - \bar{f}s.$$

With probability one, N is differentiable at the origin, and so y is almost surely differentiable at the origin, implying by Note 2 in Anevski and Hössjer [1] that $P(y \in D_h) = 0$. However, even though $y_n \stackrel{\mathcal{L}}{\to} y$, we cannot use the continuous mapping theorem directly to conclude $h(y_n) \stackrel{\mathcal{L}}{\to} h(y)$, since the limit process y is not supported on a separable subset of

$D(-\infty, \infty)$, with probability one. Instead, we will use approximation arguments to deduce **A7**. Given $\epsilon > 0$, there exist $\delta > 0$ such that $P(A_n) \geq 1 - \epsilon$ and $P(A) \geq 1 - \epsilon$, where

$$A_n = \{t_i \notin [t_0 - \delta d_n, t_0 + \delta d_n], i = 1, \ldots, n\},$$
$$A = \{N(\delta) - N(\delta-) = 0\}.$$

Since **A7** only deals with distributional convergence we introduce a new probability space $\Omega = \Omega_\epsilon$ as

$$\Omega = A \times A_1 \times A_2 \times \cdots$$

and new independent random variables $\bar{y}, \bar{y}_n : \Omega \mapsto D(-\infty, \infty)$ defined "coordinatewise" from $\omega \in \Omega$, according to

$$\bar{y}_n \stackrel{\mathcal{L}}{=} y_n | A_n,$$
$$\bar{y} \stackrel{\mathcal{L}}{=} y | A,$$

with $|$ denoting restriction. Since $\bar{y}_n'(0)$ and $\bar{y}'(0)$ exist for all $\omega \in \Omega$, we have $\bar{y}_n(\Omega) \cap D_h = \bar{y}(\Omega) \cap D_h = \emptyset$. Therefore, h is continuous on $\bar{y}(\Omega) \cup (\cup_{n=1}^{\infty} \bar{y}_n(\Omega))$. It may be seen that $\bar{y}_n \stackrel{\mathcal{L}}{\longrightarrow} \bar{y}$ and thus by the continuous mapping theorem

$$h(\bar{y}_n) \stackrel{\mathcal{L}}{\to} h(\bar{y}). \tag{23}$$

But

$$P(h(y) \leq x) = P(A)P(h(y) \leq x | A) + P(A^c)P(h(y) \leq x | A^c),$$
$$P(h(\bar{y}) \leq x) = P(h(y) \leq x | A),$$

so that

$$|P(h(\bar{y}) \leq x) - P(h(y) \leq x)| \leq 2\epsilon \tag{24}$$

and similarly

$$|P(h(\bar{y}_n) \leq x) - P(h(y_n) \leq x)| \leq 2\epsilon. \tag{25}$$

Thus **A7** follows from (23)–(25), since $\epsilon > 0$ is arbitrary.

Finally, **A5** follows from Lemma 10 in Anevski and Hössjer [1] with $\delta_n = n^{-1}$, via Proposition 1. (A careful inspection of the proof of that lemma shows that the assumption $n\delta_n \to \infty$ can be dropped for independent data.) □

4 APPENDIX

Proof (Lemma 1) A t varying in $[t_0 + (s_0 - c)d_n, t_0 + (s_0 + c)d_n]$ can be written as $t = t_0 + (s_0 + s)d_n$ with $s_0 + s \in [-c, c]$. Using (12) and (13), we obtain

$$\begin{aligned} &x_n(t_0 + (s_0 + s)d_n) \\ &\quad = x_n(t_0 + s_0 d_n) + d_n(g_n(s) + \tilde{v}_n(s)) + s d_n \bar{x}'_{b,n}(t_0), \end{aligned} \tag{26}$$

with g_n defined in (13). From **A3** and **A4** we have

$$g_n(s) + \tilde{v}_n(s) \xrightarrow{\mathcal{L}} A(|s + s_0| - |s_0|) + \tilde{v}(s)$$

on $D(-c, c)$, as $n \to \infty$. Thus, by (27), (28) and the continuous mapping theorem

$$\begin{aligned} &d_n^{-1}\left[T_{c,n}(x_n)(t_0 + s_0 d_n) - x_n(t_0 + s_0 d_n)\right] \\ &\quad = T_c[g_n(s) + \tilde{v}_n(s)](0) \xrightarrow{\mathcal{L}} T_c[A(|s_0 + s| - |s_0|) + \tilde{v}(s)](0), \end{aligned}$$

as $n \to \infty$. □

Proof (Lemma 2) Define the rescaled process y_n as in (17) for $s \in J_{n,t_0}$, as a random element of $D(J_{n,t_0})$. From (27) and (26) it follows that

$$\begin{aligned} \sup_{A_{n,\Delta}} d_n^{-1}|T_{c,n}(x_n)(\cdot) - T_J(x_n)(\cdot)| &= \sup_{[-\Delta,\Delta]} |T_c(y_n)(\cdot) - T_{J_{n,t_0}}(y_n)(\cdot)|, \\ \sup_{A_{n,\Delta}} |T_{c,n}(x_n)'(\cdot) - T_J(x_n)'(\cdot)| &= \sup_{[-\Delta,\Delta]} |T_c(y_n)'(\cdot) - T_{J_{n,t_0}}(y_n)'(\cdot)|. \end{aligned}$$

Note first that if $J \neq \mathbb{R}$ then $J_{n,t_0} \neq \mathbb{R}$, so that g_n is not defined on all of $\mathbb{R}$ in that case. If $J = \mathbb{R}$ we use Theorem 1 in Anevski and Hössjer [1] with $I = [-\Delta, \Delta]$, and if $J \neq \mathbb{R}$ we use it with $O_n = J_{n,t_0}$ (cf. Note 1 in Anevski and Hössjer [1]).

It remains to verify conditions **B1–B3** of Theorem 1 in Anevski and Hössjer [1]. **B1–B2** coincide with **A3** and **B3** follows from **A3** and **A4**. Thus all the regularity conditions of Theorem 1 are satisfied, and the first part of the lemma follows. The proof of the second part is analogous. □

The following lemma is proved in *e.g.* Lemma 1 in Anevski and Hössjer [1].

LEMMA 3 *For any interval $I \subset \mathbb{R}$, and functions l, h on I, such that l is linear and any constant $a \in \mathbb{R}$, we have*

$$T_I(h + l) = T_I(h) + l, \quad T_I(ah) = aT_I(h), \tag{27}$$

Further, the map

$$D(-\infty, \infty) \ni y \mapsto T(y) \in \mathcal{C}(-\infty, \infty) \tag{28}$$

is continuous, with both spaces equipped with the supnorm metric on compact intervals.

LEMMA 4 *Suppose that m satisfies* **A1** *and $d_n = n^{-1}$. Then $x_{b,n}$ satisfies* **A4**.

Proof Clearly $x_{b,n}$ is convex, so it suffices to establish (16). Write $m(t) = m_0(t) + 2A1_{\{t \geq t_0\}}$. Then

$$\begin{aligned} g_n(s) &= \int_{s_0}^{s_0+s} \left(m\left(\tilde{t}_0 + \lfloor un \rfloor n^{-1} + \frac{1}{2n}\right) - \frac{1}{2} m\left(\tilde{t}_0 - \frac{1}{2n}\right) \right. \\ &\quad \left. - \frac{1}{2} m\left(\tilde{t}_0 + \frac{1}{2n}\right)\right) \mathrm{d}u \\ &= A(|s_0 + s| - |s_0|) + \int_{s_0}^{s+s_0} \left(m_0\left(\tilde{t}_0 + \lfloor un \rfloor n^{-1} + \frac{1}{2n}\right) \right. \\ &\quad \left. - \frac{1}{2} m_0\left(\tilde{t}_0 + \frac{1}{2n}\right) - \frac{1}{2} m_0\left(\tilde{t}_0 - \frac{1}{2n}\right)\right) \mathrm{d}u \\ &\to A(|s + s_0| - |s_0|), \end{aligned}$$

as $n \to \infty$, uniformly on compact sets, since m_0 is continuous at t_0. □

The next result corresponds to Proposition 1 in Anevski and Hössjer [1], the proof of which is given in Anevski and Hössjer [1].

PROPOSITION 1 *Suppose* **A4** *holds and that for each* $\epsilon, \delta > 0$ *there is a finite* $\tau = \tau(\epsilon, \delta)$ *such that*

$$\limsup_{n \to \infty} P\left(\sup_{|s| \geq \tau} \left| \frac{\tilde{v}_n(s)}{g_n(s)} \right| > \epsilon \right) < \delta. \tag{29}$$

Then **A5** *holds.*

Acknowledgement

We wish to thank the referee and associate editor for their comments, that improved the readability of the paper.

References

[1] Anevski, D. and Hössjer, O. (2000) "A general asymptotic scheme for inference under order restrictions", Lund Technical Report 1, Centre for Mathematical Sciences.
[2] Brunk, H. D. (1958) "On the estimation of parameters restricted by inequalities", *The Annals of Mathematical Statistics* **29**, 437–454.
[3] Brunk, H. D. (1970) "Estimation of isotonic regression", In: Pun, M. L. (Ed.), *Nonparametric Techniques in Statistical Inference* (Cambridge University Press, London), pp. 177–197.
[4] Carolan, C. and Dykstra, R. (1999) "Asymptotic behaviour of the Grenander estimator at density fiat regions", *The Canadian Journal of Statistics* **27**(3), 557–566.
[5] Einmahl, J. H. J. (1997) "Poisson and Gaussian approximation of weighted local empirical processes", *Stochastic Processes and their Applications* **70**, 31–58.
[6] Grenander, U. (1956) "On the theory of mortality measurements Part II", *Skand.Akt.* 39.
[7] Groeneboom, P. (1985) "Estimating a monotone density", In: Le cam, L.M. and Olshen, R. A. (Eds.), *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (Wadswordth, Belmont, calif).
[8] Groeneboom, P. and Pyke, R. (1983) "Asymptotic normality of statistics based on the convex minorants of empirical distribution functions", *The Annals of Probability* **11**, 328–345.
[9] Leurgans, S. (1982) "Asymptotic distributions of slope-of-greatest-convex-minorant estimators", *The Annals of Statistics* **10**, 287–296.
[10] Parsons, V. L. (1975) Distribution theory of the isotonic estimators, *PhD thesis*, University of Iowa.
[11] Pollard, D. (1984) *Convergence of Stochastic Processes* (Springer-Verlag, New York).
[12] Prakasa Rao, B. L. S. (1969) "Estimation of a unimodal density", *Sankhya Ser. A* **31**, 23–36.
[13] Wright, F. T. (1981) "The asymptotic behaviour of monotone regression estimates", *The Annals of Statistics* **9**, 443–448.