JOURNAL OF MULTIVARIATE ANALYSIS 53, 159–179 (1995)

The Law of the Iterated Logarithm for the Multivariate Nearest Neighbor Density Estimators

STEFAN S. RALESCU*

Queens College, City University of New York

We consider estimation of a multivariate probability density function $f(x)$ by kernel type nearest neighbor (nn) estimators $g_n(x)$. The development of nn density estimation theory has had a rich history since Loftsgaarden and Quesenberry proposed the idea in 1965. In particular, there is a vast literature on convergence properties of $g_n(x)$ to $f(x)$. For statistical purposes, however, it is of importance to study also the speed of almost sure convergence. For pointwise estimation, this problem appears to have received no attention in the literature. The aim of the present paper is to obtain sharp pointwise rates of strong consistency by establishing a law of the iterated logarithm for this class of estimators. We also study the local estimation of a density function based on censored data by the kernel smoothing method using a nearest neighbor approach and derive a law of the iterated logarithm.

1. INTRODUCTION AND BACKGROUND

In many areas of statistics there has been a long standing need for a multidimensional density estimator. Density estimators are particularly important in data exploration and the presentation of results since they allow identification of interesting features and help one to draw conclusions about the data. Take $X_1, ..., X_n$ to be n points in the p-dimensional Euclidean space selected independently from a distribution with (unknown) density $f(x)$. Among the numerous schemes proposed for density estimation, one of the most popular is the Rosenblatt–Parzen kernel method which proposes the estimators

$$f_n(x) = \frac{1}{nh_n^p} \sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right), \tag{1.1}$$

Received June 2, 1993; revised June 1994.

AMS 1980 subject classifications: 62G05, 60F15.

Key words and phrases: density estimators, nearest neighbor, kernel estimators, law of the iterated logarithm, order statistics.

* Research supported in part by PSC-CUNY Grant 665358.

0047-259X/95 $6.00

where K is a positive real-valued function on $\mathbf{R}^p$ that integrates to 1 and $h_n > 0$ is a smoothing parameter that tends to 0 as $n \to \infty$. Of particular interest within this class are the moving window estimators $f_n^*(x)$ (Devroye, 1987, p. 19) obtained from f_n when K is the uniform probability density over the unit ball in $\mathbf{R}^p$. Note that $f_n^*(x)$ takes the form

$$f_n^*(x) = N_n(h_n)/nV_p(h_n) = N_n(h_n)/nc_0 h_n^p, \tag{1.2}$$

where $N_n(h_n)$ is the number of observations $X_1, ..., X_n$ that lie in the ball of radius h_n, $V_p(h_n)$ is the volume of the p-dimensional ball of radius h_n, and $c_0 = \pi^{p/2}/\Gamma(p/2+1)$ is the volume of the unit p-dimensional ball. Such estimators are known to have good pointwise properties. See for example Kreiger and Pickands (1981). A major drawback of f_n is that it cannot respond appropriately to variations in the magnitude of $f(x)$, i.e., the peakedness of the kernel is not data-responsive and the estimator is unable to adjust to denseness and sparseness of data near x. To obtain some assurance of the local control of the esimation process, Loftsgaarden and Quesenberry (1965) proposed the k nearest neighbor (nn) density estimator as a modification of $f_n^*(x)$ according to

$$g_n^*(x) = \frac{k_n}{nc_0 R_{n,k}^p(x)}, \tag{1.3}$$

where $R_{n,k}(x)$ is the distance from x to its kth nearest neighbor among $X_1, ..., X_n$ and $k = k_n$ is a sequence of positive integers satisfing $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$. In what follows, we suppress the dependence of $R_{n,k}(x)$ and related quantities on x and k unless confusion is likely. The nearest neighbor estimator was designed for *multivariate* density estimation to achieve two purposes:

1. to act like an "automatic" density estimator, thus overcoming the well known difficult obstacle of selecting the smoothing parameter involved in the construction of the estimator, and

2. to respond to local smoothing and provide an attractive estimator with a natural determination of high and low density regions.

More generally, the kernel-type nn estimators take the form

$$g_n(x) = \frac{1}{nR_n^p} \sum_{i=1}^{n} K\left(\frac{x - X_i}{R_n}\right). \tag{1.4}$$

Clearly, $g_n^*(x)$ could be viewed as the uniform kernel case of $g_n(x)$.

Among the many applications of nn estimators we cite discrimination and pattern recognition problems. It should be noted that Mack and

Rosenblatt (1979) signalled that caution in the use of nn estimators should be employed since, due to the fast growth of the nearest neighbor distance in the tails of the density, $g_n(x)$ may perform more poorly than $f_n(x)$ for tails values x where the bias of $g_n(x)$ is likely to be larger. For fixups and aspects of dealing with such difficulties, the reader is referred to Hall (1983) and Burman and Nolan (1992).

Many researchers have been concerned with the issue of consistency (weak, strong, pointwise, uniform) of nn estimators and an extensive amount of research has been done and reported in the literature on this subject. The problem has been treated in increasing generality by a number of authors: Loftsgaarden and Quesenberry (1965), Moore and Henrichon (1969), Wagner (1973), Moore and Yackel (1977a, 1977b), Devroye and Wagner (1977), and Devroye (1985), among others. (For more work and related problems concerning both nn estimators and other types of density estimators the reader is referred to the monograph by Prakasa Rao, 1983.) The equally important studies of the speed of asymptotic convergence related to consistency questions have received substantially less attention, however. A notable exception is a paper by Mack (1983) where for $p = 1$, the exact rate of uniform consistency of g_n^* was established (see also Deheuvels and Mason, 1992, for another proof of this result). Concerning laws of the iterated logarithm for the Rosenblatt–Parzen estimators (1.1), in the case $p = 1$ a fundamental contribution was made by Hall (1981) whose main device was the strong uniform approximation of the empirical process by a Kiefer process via the Hungarian embedding. These issues were further taken up by Härdle (1984), who obtained parallel results on the convergence of nonparametric kernel estimators of regression functions. Closely related laws of the iterated logarithm for sequentially calculated density estimators have been given by Wegman and Davis (1979).

When fixed bandwidth is replaced by nearest neighbor bandwidth (or more generally by a function of nearest neighbor distances), the problem becomes more complex primarily due to additional computational requirements. With few exceptions, this may be the main reason why the asymptotic properties dealing with rates of a.s. convergence for nearest neighbor based estimators have "eluded most researchers" as reported by Devroye and Györfi (1985).

In view of the well-documented consistency results, it seems highly relevant to ask what the corresponding rate for the almost sure pointwise convergence of $g_n(x)$ to $f(x)$ is. This is imperative in order to arrive at a much better understanding that is now available of the full meaning of the strong consistency result. The aim of the present paper is to answer this question by deriving the exact rate of strong pointwise consistency and we regard solving this problem as a significant component in the ambitious goal of establishing the theoretical asymptotic properties of $g_n(x)$. As concluded by

Terrell and Scott (1992) in their recent study, the nn estimator is of less interest in low dimensions ($p = 1, 2$) but "it is well worth considering in higher dimensions" ($p \geq 3$), thus "reassuring workers in classification and other areas who use these estimators." This puts more weight on the significance of our result which is valid in a multivariate setting.

In Section 2 we show a law of the iterated logarithm (LIL) for the nn estimator $g_n^*(x)$ properly centered at $\bar{g}_n(x)$ under the *weakest possible conditions* on $\{k_n\}$. This result gives the exact order of convergence for $g_n^*(x) - \bar{g}_n(x)$. For statistical interpretations, it is desirable to have exact pointwise strong convergence rates for $g_n^*(x) - f(x)$ but since $\bar{g}_n(x) - f(x)$ is deterministic and purely analytically handled, the corresponding LIL is presented without difficulty as a corollary. In Section 3, under slightly stronger conditions on $\{k_n\}$, we extend the LIL for the general class of spherically symmetric kernel-type nn estimators (1.4). Finally, a point of interest is to obtain similar results for other types of nonparametric curve estimators. While we do not attempt to take up this question in its full generality, we present in Section 3 a LIL for nn kernel estimators of a density function in the case of *censored data*. This solves an open problem (see Lo *et al.*, 1989, p. 468) and hints that analogues of our results may be possible in other cases of multivariate density estimation (e.g., nn regression estimators).

2. The LIL, Uniform Kernel Case

In the following we shall fix some point x where the density is to be estimated and assume $f(x) > 0$.

When f is continuous at x and $\{k_n\}$ satisfies $k_n \to \infty$, $k_n/n \to 0$, and $k_n/\log\log n \to \infty$ as $n \to \infty$, up to now only strong (pointwise) consistency is known from Moore and Yackel (1977b), i.e.,

$$g_n^*(x) \to f(x) \qquad \text{a.s.} \tag{2.1}$$

(see also Wagner (1973) who obtained this result under a condition equivalent to $k_n/\log n \to \infty$). The important problem of establishing the rate of convergence in (2.1) has remained open. We demonstrate that

$$\limsup_{n \to \infty} \pm \sqrt{\frac{k_n}{2 \log\log n}}\,[g_n^*(x) - f(x)] = f(x) \qquad \text{a.s.}$$

In looking for a practical technique applicable to nn estimators, we proceed in the spirit of Moore and Yackel (1977b, Theorem 1) and take as our starting point the observation, essentially contained in the original work of Loftsgaarden and Quesenberry (1965), that the variables $R_{n,k}(x)$ are functions of the kth order statistic of a uniform (0, 1) sample. The reasoning is simple and presented here to introduce the relevant notation: let

$$H(r) = P\{\|X_1 - x\| \leqslant r\} = \int_{\|u-x\| \leqslant r} f(u)\,du$$

and $Y_i = \|X_i - x\|$, $1 \leqslant i \leqslant n$. By the probability integral transform, $U_i = H(Y_i)$ are i.i.d. uniform (0, 1) r.v.'s, $1 \leqslant i \leqslant n$, and if $Y_{n:1} \leqslant \cdots \leqslant Y_{n:n}$ denote the order statistics of the Y sample, it follows that $R_{n,k}(x) = Y_{n:k}$. Hence, since H is nondecreasing, if $U_{n:1} \leqslant \cdots \leqslant U_{n:n}$ are the oder statistics of the U sample we can write $H(R_{n,k}(x)) = U_{n:k}$. This representation is attractive both theoretically and from the computational viewpoint.

To obtain the LIL we need to assume slightly stronger regularity conditions on $\{k_n\}$, namely that

$$k_n \uparrow \infty, \qquad k_n/n \sim \delta_n \downarrow 0$$

such that

$$\frac{n\,\delta_n}{\log\log n} \uparrow \infty. \tag{2.2}$$

We note that the introduction of an auxiliary sequence $\{\delta_n\}$ to ensure regularity on the rate of increase of k_n is necessary due to a technical requirement in view of the fact that there exists no nonultimately constant, nondecreasing sequence of positive integers $k_n \to \infty$ such that $k_n/n \downarrow 0$.

Define $b_n(x)$ by $H(b_n(x)) = k_n/n$ and set $\bar{g}_n(x) = k_n/nc_0(b_n(x))^p$.

THEOREM 1. *If f is continuous at x and $\{k_n\}$ satisfies condition* (2.2), *then*

$$\limsup_{n\to\infty} \pm \sqrt{\frac{k_n}{2\log\log n}}\,[g_n^*(x) - \bar{g}_n(x)] = f(x) \qquad a.s. \tag{2.3}$$

Proof. To demonstrate our result, we first note that the mean value theorem entails

$$U_{n:k} - k/n = (R_n - b_n)\,H'(\xi)$$

for some random quantity $\xi = \xi(n, x)$ such that

$$\min(b_n, R_n) < \xi < \max(b_n, R_n). \tag{2.4}$$

This in turn implies that $\sqrt{k_n/2\log\log n}\,[g_n^*(x)-g_n(x)]$ may be expressed as $(1/c_0)\,\mathrm{I}_n\,\mathrm{II}_n\,\mathrm{III}_n\,\mathrm{IV}_n$ where

$$\mathrm{I}_n=\frac{-nU_{n:k}+k_n}{\sqrt{2k_n\log\log n}},$$

$$\mathrm{II}_n=\left(\frac{k_n}{nb_n^p}\right)^2$$

$$\mathrm{III}_n=(k_n/n)^{(p-1)/p}/H'(\xi)$$

and

$$\mathrm{IV}_n=(R_n/b_n)^{-p}\sum_{i=0}^{p-1}[b_n(k_n/n)^{-1/p}]^{p-1-i}[R_n(k_n/n)^{-1/p}]^i.$$

Under condition (2.2) of the theorem, on account of Theorem 6 of Kiefer (1972) we have

$$\limsup_{n\to\infty}\pm\mathrm{I}_n=1 \qquad \text{a.s.}$$

Next, observe that by the continuity of f at x,

$$k_n/n=H(b_n)=\int_{\|u-x\|\le b_n}f(u)\,du=c_0f(x)\,b_n^p+o(b_n^p) \qquad \text{as} \quad n\to\infty \tag{2.5}$$

which implies

$$\lim_{n\to\infty}\mathrm{II}_n=[c_0f(x)]^2.$$

To handle III_n note that

$$H'(r)=\lim_{\varepsilon\to 0}\varepsilon^{-1}\left[\int_{\|u-x\|\le r+\varepsilon}f(u)\,du-\int_{\|u-x\|\le r}f(u)\,du\right]$$

$$=\int_{\|u-x\|=r}f(u)\,d\sigma_r(u),$$

where $d\sigma_r(u)$ is an element of surface area (sphere measure) of radius r and center x. By passing to spherical coordinates we can write

$$H'(\xi)=\xi^{p-1}\int_0^\pi\cdots\int_0^\pi\int_0^{2\pi}f(x_1+\xi\cos\theta_1,\ldots,x_p+\xi\sin\theta_1\cdots\sin\theta_{p-1})$$
$$\times\sin^{p-2}\theta_1\cdots\sin\theta_{p-2}\,d\theta_1\cdots d\theta_{p-1}.$$

Now, since $b_n \to 0$ and $R_n \to 0$ a.s. as $n \to \infty$, it follows that $\xi \to 0$ a.s., and by the continuity of f at x we deduce that

$$H'(\xi)/\xi^{p-1} \to pc_0 f(x) \qquad \text{a.s.}$$

Observe also that (2.5) and the a.s. convergence of $g_n^*(x)$ to $f(x)$ entail that $k_n/n\xi^p \to c_0 f(x)$ a.s. This and the previous statement clearly imply that

$$\mathrm{III}_n = \left(\frac{k_n}{n\xi^p}\right)^{(p-1)/p} \left(\frac{\xi^{p-1}}{H'(\xi)}\right) \to [c_0 f(x)]^{-1/p}/p \qquad \text{a.s.}$$

To complete the proof of the theorem it is only necessary to observe that almost surely $(R_n/b_n)^p \to 1$ as $n \to \infty$ and

$$\sum_{i=0}^{p-1} [b_n(k_n/n)^{-1/p}]^{p-1-i} [R_n(k_n/n)^{-1/p}]^i \to p/[c_0 f(x)]^{(p-1)/p} \qquad \text{a.s.} \quad \blacksquare$$

In practice it is appropriate to replace $\bar{g}_n(x)$ by $f(x)$ in the statement (2.3). This leads to the following result.

COROLLARY. *If f is Lipschitz at x and $\{k_n\}$ satisfies (2.2) as well as*

$$\frac{k_n^{p+2}}{n^2(\log\log n)^p} \to 0 \qquad as \quad n \to \infty, \tag{2.6}$$

then

$$\limsup_{n \to \infty} \pm \sqrt{\frac{k_n}{2\log\log n}}\, [g_n^*(x) - f(x)] = f(x) \qquad a.s. \tag{2.7}$$

Proof. It suffices to note that

$$|k_n/n - c_0 f(x)\, b_n^p| = \left| \int_{\|u-x\| \leqslant b_n} [f(u) - f(x)]\, du \right| \leqslant (\text{const})\, b_n^{p+1}.$$

This entails $|\bar{g}_n(x) - f(x)| = 0((k_n/n)^{1/p})$ which together with (2.6) and (2.3) demonstrates the result. $\blacksquare$

Remarks. Observe that

1. Theorem 1 is true under minimal assumptions on f and $\{k_n\}$.

2. Conclusion (2.3) of Theorem 1 yields the exact rate of almost sure convergence for $g_n^*(x) - \bar{g}_n(x)$. The above estimation of $\bar{g}_n(x) - f(x)$ under the conditions of the corollary ensures the validity of (2.7) and the requirement (2.6) determining additional growth on the variation of k_n may be

interpreted as the price one has to pay to replace the centering term $\bar{g}_n(x)$ by $f(x)$ in the LIL.

3. The exact rate of a.s. pointwise convergence is $\sqrt{(\log\log n)/k_n}$, while the exact rate of a.s. uniform convergence (which is available in the literature only in $p=1$ dimension) is slower, namely $\sqrt{(\log n)/k_n}$ (Mack, 1983).

4. An associated LIL could be derived for the conceptually similar (but substantially simpler) estimator $f_n^*(x)$. Specifically, with $m_n(x) = H(h_n)/(c_0 h_n^p)$, we have

$$\limsup_{n\to\infty} \pm \sqrt{\frac{nh_n^p}{2\log\log n}}\,[f_n^*(x) - m_n(x)] = \sqrt{f(x)/c_0} \qquad \text{a.s.,} \tag{2.8}$$

provided f is continuous at x, $f(x) > 0$, and

$$h_n \downarrow 0, \qquad nh_n \uparrow \infty, \qquad \text{and} \qquad \frac{\log\log n}{nh_n^p} \to 0 \qquad \text{as} \quad n\to\infty. \tag{2.9}$$

To demonstrate (2.8), observe that $f_n^*(x)$ can be expressed as $H_n(h_n)/(c_0 h_n^p)$ where H_n is the empirical distribution function (e.d.f) of $Y_1, ..., Y_n$. Then, by using the probability integral transformation $H_n(h_n) = \bar{F}_n(H(h_n))$ where $\bar{F}_n$ denotes the e.d.f. of an i.i.d. uniform (0, 1) sample of size n, and from assumption (2.9) and in view of the fact that $H(h_n)/h_n^p \to c_0 f(x)$ as $n\to\infty$, we deduce that

$$\limsup_{n\to\infty} \pm \frac{\sqrt{n}\,[H_n(h_n) - H(h_n)]}{\sqrt{2h_n^p \log\log n}} = \sqrt{c_0 f(x)} \qquad \text{a.s.}$$

This follows from Kiefer (1972) and is equivalent to (2.8). Furthermore, if in addition f is Lipschitz at x and $\{h_n\}$ also satisfies

$$\frac{nh_n^{p+2}}{\log\log n} \to 0 \qquad \text{as} \quad n\to\infty, \tag{2.10}$$

then the difference $m_n(x) - f(x)$ vanishes of higher order and we obtain

$$\limsup_{n\to\infty} \pm \sqrt{\frac{nh_n^p}{2\log\log n}}\,[f_n^*(x) - f(x)] = \sqrt{f(x)/c_0} \qquad \text{a.s.}$$

Conditions (2.9) and (2.10) will be verified in practice if one takes $h_n = an^{-b}$ for positive a and $1/(p+2) \leqslant b < 1/p$.

5. The result (2.3) is preserved by the general class of spherically symmetric kernel-type nn estimators. See Section 3.

3. Kernel-Type NN Estimators

In this section we extend our analysis to a larger class of nn estimators, namely that of the spherically symmetric (ss) kernel type. For multivariate F ($p > 1$), the shape of the p-kernel used for density estimation becomes a more important consideration than in the univariate case. As noted by Terrel and Scott (1992), possibilities include two obvious approaches: the first one is based on hyper-rectangles and constructs a "coordinate-wise product" kernel. The second "allows for the full range of linear scaling, by making the kernel spherically symmetric." For estimating a density at a fixed point, Elkins (1968) found some differences between cubical and spherical kernels and form Sacks and Ylvisaker (1981) it is known that when minimizing the maximum mean square error in dimensions $p > 1$, the asymptotically optimal kernel is *spherically symmetric* rather than a coordinatewise product kernel. Guided by these facts we consider ss kernels, i.e., $K(x) = K_0(\|x\|)$. We now make the following assumptions:

$$k_n \uparrow \infty, \qquad \frac{k_n}{n} \sim \delta_n \downarrow 0, \qquad \frac{(\log(n/\delta_n))^4}{n\,\delta_n(\log\log n)} \to 0, \tag{3.1}$$

and

$$\frac{\delta_m}{\delta_n} \to 1 \qquad \text{as} \quad m, n \to \infty \qquad \text{with} \quad \frac{m}{n} \to 1.$$

Our next result yields an extension of (2.3) for the general class of ss kernel-type nn estimators.

THEOREM 2. *Let g_n be defined by* (1.4) *with $K(x) = K_0(\|x\|)$ where K_0 has bounded support and is continuous a.e. Assume that f satisfies a Lipschitz condition in a neighborhood of x and that $\{k_n\}$ satisfies* (3.1). *Then*

$$\limsup_{n \to \infty} \pm \sqrt{\frac{k_n}{2\log\log n}}\,[g_n(x) - \bar{g}_n(x)]$$
$$= c_0 f(x) \sqrt{\int_0^\infty p z^{p-1} K_0^2(z)\, dz} \tag{3.2}$$

where $\bar{g}_n(x) = (1/b_n^p) \int_0^\infty K_0(t/b_n)\, dH(t)$.

Proof. For the sake of readability, we begin by summarizing the steps to prove Theorem 2. First, we show that $g_n(x)$ is close to

$$\tilde{g}_n(x) = \frac{1}{b_n^p} \int_0^\infty K_0\left(\frac{t}{b_n}\right) dH_n(t)$$

by decomposing $g_n(x) - \tilde{g}_n(x)$ into a sum of three terms. Here $\tilde{g}_n(x)$ is a fixed-width estimator, but with an h_n unknown to the statistician. This again reflects the property of nn estimators of being self-adjusting to the level of f. Next, error bounds are obtained on the components in the decomposition by utilizing sharp bounds for the oscillation modulus of the empirical process and the precise almost sure behavior of the uniform empirical process in the lower tail (i.e., near the origin). Finally, we make use of a result which extends Theorem 2 of Hall (1981) and infer (3.2).

Note that $g_n(x)$ takes the form

$$g_n(x) = \frac{1}{R_n^p} \int_0^\infty K_0\left(\frac{t}{R_n}\right) dH_n(t) = -\frac{1}{R_n^p} \int_0^\infty H_n(R_n z)\, dK_0(z) \qquad (3.3)$$

where the last equality follows by integration by parts. Similarly, $\tilde{g}_n(x)$ may be expressed as in (3.3) with R_n replaced by b_n. Now observe that for $w > 0$,

$$\frac{1}{w^p} \int_0^\infty H(wz)\, dK_0(z) = -\int_0^\infty z^{p-1} A(wz)\, K_0(z)\, dz \qquad (3.4)$$

where

$$A(u) = \int_0^\pi \cdots \int_0^\pi \int_0^{2\pi} f(x_1 + u\cos\theta_1, \ldots, x_p$$

$$+ u \sin\theta_1 \cdots \sin\theta_{p-1}) \sin^{p-2}\theta_1 \cdots \sin\theta_{p-2}\, d\theta_1 \cdots d\theta_{p-1}.$$

We will apply (3.4) with $w = R_n$ and $w = b_n$. Therefore, in view of (3.3) $g_n(x) - \tilde{g}_n(x)$ can be written as $\mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3$ where

$$\mathrm{I}_1 = \left(\frac{1}{R_n^p} - \frac{1}{b_n^p}\right) \int_0^\infty [H(b_n z) - H_n(b_n z)]\, dK_0(z)$$

$$\mathrm{I}_2 = \frac{1}{R_n^p} \int_0^\infty \{H(R_n z) - H_n(R_n z) - [H(b_n z) - H_n(b_n z)]\, dK_0(z)$$

and

$$\mathrm{I}_3 = \int_0^\infty [A(R_n z) - A(b_n z)]\, z^{p-1} K_0(z)\, dz.$$

Throughout the proof, the symbols c, C, and M denote positive generic constants, possibly different at different appearances. To simplify the presentation we shall note that the stochastic estimates below hold a.s. for n large enough. In view of the conditions on f and K_0, if $|z| \leqslant c$,

$$|A(R_n z) - A(b_n z)| \leqslant C\, |R_n - b_n|\, |z|$$

which implies

$$|I_3| \leqslant C\,|R_n - b_n| \int_0^{\infty} z^{p-1} K_0(z)\,dz.$$

Set $\varepsilon_n = k_n/n$ and observe that on account of the proof of Theorem 1,

$$|R_n - b_n| = O\left(\varepsilon_n^{(1-p)/p} \sqrt{\frac{\varepsilon_n \log\log n}{n}}\right) \qquad \text{a.s.} \tag{3.5}$$

Hence

$$\mathrm{I}_3 = O\left(\varepsilon_n^{(1-p)/p} \sqrt{\frac{\varepsilon_n \log\log n}{n}}\right) \qquad \text{a.s.} \tag{3.6}$$

To handle I_1, note that $H_n(t) - H(t) = \bar{F}_n(H(t)) - H(t)$ where $\bar{F}_n$ is the e.d.f. of a uniform (0, 1) sample. Since $b_n z$ in the integrand of I_1 varies in the lower tail of the uniform empirical process, the Lemma below provides a means to estimate I_1. This is a well-known result (cf. Einmahl and Mason, 1988, p. 1635) and is stated here only for the sake of reference.

LEMMA 1. *Let $0 < \varepsilon_n < 1$ be such that*

$$\varepsilon_n \downarrow 0, \qquad n\varepsilon_n \uparrow \qquad \textit{and} \qquad \frac{n\varepsilon_n}{\log\log n} \to \infty.$$

Then, for all $\lambda > 1$,

$$\limsup_{n\to\infty} \sup_{0 \leqslant s \leqslant \lambda\varepsilon_n} \sqrt{\frac{n}{\varepsilon_n \log\log n}}\, |\bar{F}_n(s) - s| = \sqrt{2} \qquad \textit{a.s.}$$

Now, clearly

$$|\mathrm{I}_1| \leqslant \left|\frac{1}{R_n^p} - \frac{1}{b_n^p}\right| \sup_{0 < u < Mb_n} |H_n(u) - H(u)| \tag{3.7}$$

and if $0 < u < Mb_n < c\,\delta_n^{1/p}$, then $0 < H(u) < c\,\delta_n$. From (2.3),

$$\limsup_{n\to\infty} \pm\varepsilon_n^{3/2} \sqrt{\frac{n}{2\log\log n}} \left[\frac{1}{R_n^p} - \frac{1}{b_n^p}\right] = c_0 f(x) > 0 \qquad \text{a.s.,}$$

which implies

$$\left|\frac{1}{R_n^p} - \frac{1}{b_n^p}\right| = O\left(\frac{1}{\delta_n}\sqrt{\frac{\log\log n}{\delta_n}}\right) \qquad \text{a.s.}$$

This together with (3.7) and an application of Lemma 1 entails

$$\mathrm{I}_1 = O\left(\frac{\log\log n}{n\,\delta_n}\right) \qquad \text{a.s.} \tag{3.8}$$

It remains to consider I_2. From Section 2 it is clear that there exists $c > 0$ such that for almost all ω there exists $n_0 = n_0(\omega)$ such that $\max\{b_n, R_n\} \leqslant c\,\delta_n^{1/p}$ holds for all $n \geqslant n_0$. Now, with $u = H(b_n z)$ and $v = H(R_n z)$, an application of the mean value theorem gives for $|z| \leqslant C$

$$|u - v| \leqslant M\,|R_n - b_n|\,\delta_n^{(1/p-1)} \tag{3.9}$$

(see the expression for $H'(\xi)$ in Section 2 with ξ in (2.4) replaced by ξ/z and satisfying $\xi < c\,\delta_n^{1/p}$). Writing

$$\mathrm{I}_2 = \frac{1}{R_n^p} \int_0^C \{\bar{F}_n(u) - u - [\bar{F}_n(v) - v]\}\, dK_0(z),$$

we have in view of (3.9) and (3.5) that

$$|\mathrm{I}_2| \leqslant \frac{1}{R_n^p\sqrt{n}}\,\omega_n(a_n)$$

where $\omega_n(a) = \sup_{|t-s| \leqslant a} \sqrt{n}\,|\bar{F}_n(t) - t - \bar{F}_n(s) + s|$ denotes the oscillation modulus of the uniform empirical process and $a_n = c\sqrt{(\delta_n \log\log n)/n}$. Now Theorem 0.2 of Stute (1982) comes into play. For this, we need to note that the assumption (3.1) entails

$$a_n \downarrow 0, \qquad na_n \uparrow \infty, \qquad \frac{\log(1/a_n)}{\log\log n} \to \infty, \qquad \text{and} \qquad \frac{\log(1/a_n)}{na_n} \to 0.$$

Hence, from Stute's result and the fact that $R_n^{1/p} = O(\delta_n^{-1})$ as $n \to \infty$,

$$\mathrm{I}_2 = O\left(\frac{1}{\delta_n\sqrt{n}}\sqrt{a_n \log\frac{1}{a_n}}\right) = O\left(\frac{(\log\log n)^{1/4}}{(n\,\delta_n)^{3/4}}\sqrt{\log\frac{n}{\delta_n}}\right) \qquad \text{a.s.} \tag{3.10}$$

The final step of the proof consists in an application of the following result which extends Hall's (1981) Theorem 2.

Lemma 2. *Let K_0 be a function of bounded variation on $[0, \infty)$ satisfying*

$$z^p K_0(z) \to 0 \qquad \textit{as} \quad z \to \infty \qquad \textit{and} \qquad \int_0^\infty z^{p-1} K_0^2(z)\, dz < \infty.$$

Let $X_1, ..., X_n$ be i.i.d. p-dimensional random vectors with unknown d.f. F and density $f(x) > 0$. (Here x is fixed.) Assume that F satisfies a Lipschitz condition in a neighborhood of x. Let $f_n(x)$ be given by (1.1) with $K(u) = K_0(\|u\|)$ and assume that

$$\frac{h_m}{h_n} \to 1 \qquad as \quad m, n \to \infty \qquad with \quad \frac{m}{n} \to 1 \tag{3.11}$$

and

$$\frac{(\log n)^4}{nh_n^p(\log \log n)} \to 0 \qquad as \quad n \to \infty.$$

Then

$$\limsup_{n \to \infty} \pm \sqrt{\frac{nh_n^p}{2 \log \log n}} \, [f_n(x) - Ef_n(x)]$$

$$= \sqrt{c_0 f(x) \int_0^{\infty} pz^{p-1} K_0^2(z)\, dz} \qquad a.s. \tag{3.12}$$

The proof of Lemma 2 is based on Theorem 1 of Hall (1981) and is sketched in the Appendix.

Applying Lemma 2 with $f_n(x) = \tilde{g}_n(x)$ and $h_n = b_n$, it remains to show that $c_n(\mathrm{I}_1 + \mathrm{I}_2 + \mathrm{I}_3) = o(1)$ a.s. as $n \to \infty$, where $c_n = \sqrt{nb_n^p/(\log \log n)}$. From (3.6), $c_n \mathrm{I}_3 \to 0$ a.s. reduces to $\varepsilon_n^{1/p} \to 0$. Also, on account of (3.8), $c_n \mathrm{I}_1 \to 0$ a.s. is equivalent to $(\log \log n)/n\,\delta_n \to 0$. Finally, from (3.10) $c_n \mathrm{I}_2 \to 0$ a.s. amounts to $(\log(n/\delta_n))^2/n\,\delta_n(\log \log n) \to 0$ which follows from (3.1).

The proof is complete.

Remarks. We observe that

6. The condition (3.1) is relatively mild and needed for technical reasons. It is however stronger than (2.2) which entails the LIL for the naive nn estimator $g_n^*(x)$. Informally, the third condition in (3.1) prevents k_n from being too large. All the requirements (reguarlity and growth) contained in (3.1) would be satisfied in practice if we take $\delta_n \sim an^{-b}$ for positive a and b with $b > 1$.

7. It is important to recognize that the order of magnitude $O(\sqrt{(\log \log n)/k_n})$ of the almost sure convergence to zero for $g_n(x) - \bar{g}(x)$ is maintained under the milder conditions (2.2). The proof of this situation is given in the Appendix.

8. In the context of other types of nonparametric curve estimation, our method may be adapted to yield appropriate laws of the iterated logarithm. To illustrate this aspect, we shall focus here on the LIL for the nn kernel estimator of a density function in the case of censored data. (A similar conclusion could be obtained for the nn kernel estimator of a hazard rate function.)

Consider a random censorship model with a sequence $X_1, ..., X_n$ of i.i.d. random variables with density f on the real line. In survival analysis the X's are often lifetimes or known monotone transformations of them and, due to causes of failure, the X's are at risk of being censored from the right: that is, along with the sequences $X_1, ..., X_n$, there exist i.i.d. random variables $Y_1, ..., Y_n$ (independent of the X-sequence) such that only

$$Z_i = \min(X_i, Y_i) \qquad \text{and} \qquad \gamma_i = 1_{\{X_i \leqslant Y_i\}}, \qquad 1 \leqslant i \leqslant n,$$

are observed (here γ_i indicates whether X_i has been observed or not). The objective is to estimate the true density f from the sample (Z_i, γ_i), $1 \leqslant i \leqslant n$. For this, consider the well-known Kaplan–Meier (1958) estimator $1 - \Gamma_n(u)$ of $1 - F(u) = P\{X_1 \geqslant u\}$ and the nn kernel density estimator

$$g_n(x) = \frac{1}{\rho_n} \int_{-\infty}^{\infty} K\left(\frac{x-y}{\rho_n}\right) d\Gamma_n(y), \tag{3.13}$$

where $\rho_n = \rho_{n,k}(x)$ is the distance from x to its kth nearest neighbor among $Z_1, ..., Z_n$ and $k = k_n$ is as before. In the following it will be assumed that Y_1 has a density $g = G'$, that $0 < F(x)\, G(x) < 1$, and that both f and g are continuous at x. Since censored data traditionally occur in lifetime analysis, we also assume that the X's and Y's are nonnegative. The actually observed Z's have a d.f. H^* satisfying $1 - H^*(x) = (1 - F(x))(1 - G(x))$ for $x \geqslant 0$. In the case of fixed bandwidth kernel estimators of f based on Γ_n, laws of iterated logarithm were obtained by Diehl and Stute (1988) and Lo *et al.* (1989). However, it was signaled that "a nearest neighbor approach may be preferable to the fixed bandwidth approach from an extensive simulation experiment" as reported by Lo, *et al.* (1989, Section 5, p. 468, Comment b). The problem of establishing a LIL for $g_n(x)$ given by (3.13) has remained open (see also Mielniczuk, 1986, who deals with nn estimators in the context of a censoring mechanism). The next result provides a solution to this problem.

THEOREM 3. *If*

(a) *K is continuously differentiable probability kernel with finite support*

(b) *f is bounded and satisfies a Lipschitz condition in a neighborhood of x, and*

(c) $k_n \uparrow \infty$, $k_n/n \sim \delta_n \downarrow 0$, $(\log n)^4/n\,\delta_n(\log\log n) \to 0$ *and* $\delta_m/\delta_n \to 1$ as $m, n \to \infty$ with $m/n \to 1$,

then

$$\limsup_{n\to\infty} \pm \sqrt{\frac{k_n}{2\log\log n}}\,[g_n(x) - \bar{g}_n(x)]$$

$$= f(x)\sqrt{\frac{2}{1-G(x)}\int_{-\infty}^{\infty} K^2(z)\,dz}, \tag{3.14}$$

where $b_n = b_n(x)$ *is defined by* $H^*(x+b_n) - H^*(x-b_n) = k_n/n$ *and* $\bar{g}_n(x) = (1/b_n)\int_{-\infty}^{\infty} K((x-y)/b_n)\,dF(y)$.

Appendix

A.1. *Proof of Lemma* 2.

We establish (3.12) by making use of Theorem 1 of Hall (1981). For this, we set $K_h(y) = K_0(y/h)$ in Hall's notation with $h = h_n$. To simplify the arguments, we shall assume that K_0 has bounded support, although we stress that this condition is not necessary for the result to hold.

Let $\sigma_h^2 = \operatorname{Var} K_h(Y_1)$. It is clear that

$$\sigma_h^2 = L(h) - \left\{\int_0^{\infty} H(hz)\,dK_0(z)\right\}^2$$

where

$$L(h) = -2\int_0^{\infty} H(hz)\,K_0(z)\,dK_0(z).$$

Set $\psi_n(z) = H(hz)/h^p z^p - c_0 f(x)$ and note that $(\forall)\varepsilon > 0$, $(\exists)\, n_0 = n_0(\varepsilon)$ such that for all $n \geqslant n_0$ and $|z| \leqslant c$, $|\psi_n(z)| \leqslant \varepsilon$. Thus, since $K_0(\cdot)$ has finite support, writing

$$L(h) = c_0 f(x)\,h^p \int_0^{\infty} pz^{p-1} K_0^2(z)\,dz + 2h^p \int_0^{\infty} \psi_n(z)\,z^p K_0(z)\,dK_0(z)$$

and using $E\{K_h(Y_1)\} = -\int_0^{\infty} H(hz)\,dK_0(z) \sim (\text{const})\,h^p$ as $h \to 0$, it follows that

$$\sigma_h^2 \sim L(h) \sim c_0 f(x)\,h^p \int_0^{\infty} pz^{p-1} K_0^2(z)\,dz \qquad \text{as} \quad h \to 0. \tag{4.1}$$

In view of Hall's Theorem 1 and (4.1) we have to demonstrate that if $h, k \to 0$ such that $h/k \to 1$, then

$$\frac{1}{h^p} \operatorname{cov}\{K_h(Y_1), K_k(Y_1)\} \to a \tag{4.2}$$

where $a = c_0 f(x) \int_0^c p z^{p-1} K_0^2(z)\, dz$. But the above infinitesimal estimation shows that an account of (4.1), (4.2) is entailed by

$$D(h, k) = \frac{1}{h^p} \int_0^c \left[K_0\left(\frac{y}{h}\right) - K_0\left(\frac{y}{k}\right) \right]^2 dH(y) \to 0. \tag{4.3}$$

Looking at $D(h, k)$ we see that it equals

$$\int_0^c \left[K_0(z) - K_0\left(\frac{zh}{k}\right) \right]^2 z^{p-1} A(zh)\, dz$$

where $A(\cdot)$ was defined in the proof of Theorem 2. Thus, since K_0 is continuous a.e., $K_0(zh/k) \to K_0(z)$ a.e. and it follows that (4.3) holds. Lemma 2 is thus verified.

A.2. *Proof of Remark* 7.

We shall assume that K_0 is continuously differentiable, has bounded support, and that f is Lipschitz at x. Furthermore, suppose that $\{k_n\}$ satisfies (2.2). In view of the remark following (3.2), all the integrals below may be considered a.s. over $[0, c\varepsilon_n^{1/p}]$ (for n large enough) with $\varepsilon_n = k_n/n$ as before. We note that

$$|g_n(x) - \tilde{g}_n(x)| \leqslant [\sup_{0 \leqslant t \leqslant c\varepsilon_n^{1/p}} \Delta_n(t)] \int_0^{c\varepsilon_n^{1/p}} dH_n(t) \tag{4.4}$$

where

$$\Delta_n(t) = \left| \frac{1}{R_n^p} K_0\left(\frac{t}{R_n}\right) - \frac{1}{b_n^p} K_0\left(\frac{t}{b_n}\right) \right|.$$

Now

$$\begin{aligned} \Delta_n(t) &\leqslant \left| \left(\frac{1}{R_n^p} - \frac{1}{b_n^p}\right) K_0\left(\frac{t}{R_n}\right) \right| + \frac{1}{b_n^p} \left| K_0\left(\frac{t}{R_n}\right) - K_0\left(\frac{t}{b_n}\right) \right| \\ &= A_n(t) + B_n(t). \end{aligned}$$

Looking at $A_n(t)$, we see that uniformly in $0 \leqslant t \leqslant c\varepsilon_n^{1/p}$ it is not greater than (const) $|1/R_n^p - 1/b_n^p|$ and in view of the estimate preceding (3.8) it follows that

$$\sup_{0 \leqslant t \leqslant c\varepsilon_n^{1/p}} A_n(t) = O\left(\frac{n}{k_n}\sqrt{\frac{\log\log n}{k_n}}\right) \quad \text{a.s.} \tag{4.5}$$

To treat $B_n(t)$, from the assumption on K_0 we have

$$B_n(t) \leqslant (\text{const})\,\frac{|t|\,|b_n - R_n|}{b_n^{p+1} R_n}$$

and using (3.5) we see that uniformly in $0 \leqslant t \leqslant c\varepsilon_n^{1/p}$; with probability one,

$$\begin{aligned}B_n(t) &\leqslant O(\varepsilon_n^{1/p})\,O\left(\frac{n}{k_n}\right)O\left(\frac{\sqrt{k_n \log\log n}}{n}\right)O(\varepsilon_n^{(1-p)/p})\,O(\varepsilon_n^{-2/p})\\ &= O\left(\frac{n}{k_n}\sqrt{\frac{\log\log n}{k_n}}\right).\end{aligned} \tag{4.6}$$

On the other hand, we have a.s.

$$\int_0^{c\varepsilon_n^{1/p}} dH_n(t) \leqslant H(c\varepsilon_n^{1/p}) + (\text{const})\,\frac{\sqrt{k_n \log\log n}}{n}$$

and since $(n/k_n)\,H(c\varepsilon_n^{1/p}) \to c_0 c^p f(x)$ as $n \to \infty$ and $\sqrt{k_n \log\log n} \leqslant k_n$ for n sufficiently large, the previous estimate entails

$$\int_0^{c\varepsilon_n^{1/p}} dH_n(t) = O\left(\frac{k_n}{n}\right) \quad \text{a.s.} \tag{4.7}$$

Combining (4.4)–(4.7) demonstrates

$$|g_n(x) - \tilde{g}_n(x)| = O\left(\sqrt{\frac{\log\log n}{k_n}}\right) \quad \text{a.s.} \tag{4.8}$$

We conclude the proof Remark 7 with an analysis of the error in the approximation $\tilde{g}_n(x) \approx \bar{g}_n(x)$. Upon integration by parts we can express

$$\tilde{g}_n(x) - \bar{g}_n(x) = \int_0^{c\varepsilon_n^{1/p}} [H(t) - H_n(t)]\, dK_n(t)$$

where $K_n(t) = (1/b_n^p)\,K_0(t/b_n)$, whence

$$|\tilde{g}_n(x) - \bar{g}_n(x)| \leqslant \left[\sup_{0 \leqslant t \leqslant c\varepsilon_n^{1/p}} |H_n(t) - H(t)|\right] V_{J_n}(K_n)$$

where $J_n = [0, c\varepsilon_n^{1/p}]$ and $V_J(\varphi)$ denotes the total variation of a function φ on the interval J. From Lemma 1, if $\{k_n\}$ satisfies (2.2) it follows that

$$\sup_{t \in J_n} |H_n(t) - H(t)| = O\left(\frac{\sqrt{k_n \log \log n}}{n}\right) \qquad \text{a.s.}$$

Also, if Π denotes the set of all partitions π: $0 = t_0 < t_1 < \cdots < t_m = c\varepsilon_n^{1/p}$

$$V_{J_n}\left(t \to K_0\left(\frac{t}{b_n}\right)\right) = \sup_{\pi \in \Pi} \sum_{i=0}^{m-1} \left|K_0\left(\frac{t_{i+1}}{b_n}\right) - K_0\left(\frac{t_i}{b_n}\right)\right|$$

$$\leqslant (\text{const}) \sup_{\pi \in \Pi} \sum_{i=0}^{m-1} \frac{(t_{i+1} - t_i)}{b_n}$$

$$= (\text{const})(k_n/n)^{1/p}/b_n$$

Thus from the above estimate and since $1/b_n^p = O(n/k_n)$, we can assert that

$$|\tilde{g}_n(x) - \bar{g}_n(x)| = O\left(\sqrt{\frac{\log \log n}{k_n}}\right) \qquad \text{a.s.} \tag{4.9}$$

The conclusion stated in Remark 7 now follows from (4.8) and (4.9).

Remark A.2. Suppose that K is a positive real-valued function on R^p that integrates to 1. Then, if $\{k_n\}$ satisfies (2.6), we can estimate the deterministic bias $\bar{g}_n(x) - f(x)$ as follows: since $\int K(x)\,dx = 1$ is equivalent to

$$\int_0^1 c_0 p t^{p-1} K_0(t)\,dt = 1,$$

we have

$$|\bar{g}_n(x) - f(x)| \leqslant \frac{1}{b_n^p} \int_0^{c\varepsilon_n^{1/p}} |H'(t) - c_0 p f(x)\, t^{p-1}|\, K_0\left(\frac{t}{b_n}\right) dt.$$

Next, since f is Lipschitz at x it follows that

$$|H'(t) - c_0 p f(x)\, t^{p-1}| = \left|\int_{\|u-x\|=t} [f(u) - f(x)]\, d\sigma(u)\right| \leqslant (\text{const})\, t^p.$$

Therefore,

$$|\bar{g}_n(x) - f(x)| \leqslant (\text{const}) \frac{k_n}{n b_n^p} \int_0^{c\varepsilon_n^{1/p}} K_0\left(\frac{t}{b_n}\right) dt = O(b_n) \tag{4.10}$$

The "big oh" on the right-hand side of (4.10) is $O((k_n/n)^{1/p})$ and under the requirement (2.6) this is in turn $O(\sqrt{(\log\log n)/k_n})$. Thus

$$|\bar{g}_n(x) - f(x)| = O\left(\sqrt{\frac{\log\log n}{k_n}}\right).$$

A.3. *Proof of Theorem* 3.

Let

$$g_n^*(x) = \frac{1}{b_n}\int_{-\infty}^{\infty} K\left(\frac{x-y}{b_n}\right) d\Gamma_n(y).$$

Since the proof is similar to that of Theorem 2, it will only be sketched.

Let $W_n(z) = \sqrt{n}\,[\Gamma_n(z) - F(z)]$ be the empirical process pertaining to $\Gamma_n(\cdot)$. Then $g_n(x) - g_n^*(x) = \mathrm{I} + \mathrm{II} + \mathrm{III}$ where

$$\mathrm{I} = \frac{1}{\sqrt{n}}\left(\frac{1}{\rho_n} - \frac{1}{b_n}\right)\int_{-\infty}^{\infty} [W_n(x - b_n y) - W_n(x)]\, dK(y)$$

$$\mathrm{II} = \frac{1}{\sqrt{n}\,\rho_n}\int_{-\infty}^{\infty} [W_n(x - \rho_n y) - W_n(x - b_n y)]\, dK(y)$$

$$\mathrm{III} = \int_{-\infty}^{\infty} [f(x - \rho_n y) - f(x - b_n y)]\, K(y)\, dy.$$

The following result is from Schäfer (1986).

LEMMA 3. *If* $p_n \to 0$ *such that* $np_n/\log n \to \infty$, *then*

$$\sup_{|F(v) - F(u)| \leqslant p_n} |W_n(v) - W_n(u)| = O(\sqrt{p_n \log n}) \qquad \text{a.s.}$$

Denote by $\tilde{\omega}_n(\cdot)$ the global oscillation modulus corresponding to W_n. Using Lemma 3 we have

$$|\mathrm{I}| \leqslant \frac{1}{\sqrt{n}}\left|\frac{1}{\rho_n} - \frac{1}{b_n}\right| \tilde{\omega}_n(cb_n) = O\left(\frac{\sqrt{(\log n)(\log\log n)}}{k_n}\right) \qquad \text{a.s.}$$

As for II, since $|F(x - \rho_n y) - F(x - b_n y)| \leqslant (\text{const})\, |\rho_n - b_n|$ for all $|y| \leqslant C$, on account of Lemma 3 and the fact that $|\rho_n - b_n| = O(\sqrt{k_n \log\log n}/n)$ a.s., we find

$$\mathrm{II} = O\left(\frac{(\log\log n)^{1/4}\,(\log n)^{1/2}}{k_n^{3/4}}\right) \qquad \text{a.s.}$$

Finally,

$$|\mathrm{III}| \leqslant (\mathrm{const})\, |p_n - b_n| = O\left(\frac{\sqrt{k_n \log\log n}}{n}\right) \quad \text{a.s.}$$

These facts together with Corollary 1 of Diehl and Stute (1988) imply the validity of (3.14)

Acknowledgments

I express my sincere thanks to an anonymous referee for some very constructive criticism and suggestions which led to substantial improvement of the paper.

References

[1] Burman, P., and Nolan, D. (1992). Location-adaptive density estimation and nearest-neighbor distance. *J. Multivariate Anal.* **40** 132 157.

[2] Deheuvels, P., and Mason, D. M. (1992). Functional laws of the iterated logarithm for the increments of empirical and quantile processes. *Ann. Probab.* **20** 1248 1287.

[3] Devroye, L. (1985). A note on the L_1 consistency of variable kernel estimates. *Ann. Statist.* **13** 1041-1049.

[4] Devroye, L. (1987). *A Course in Density Estimation.* Birkhäuser, Boston Basel Stuttgart.

[5] Devroye, L., and Wagner, T. J. (1977). The strong uniform consistency of nearest-neighbor density estimates. *Ann. Statist.* **5** 536 540.

[6] Devroye, L., and Györfi, L. (1985). *Nonparametric Density Estimation: The L_1 View.* Wiley, New York.

[7] Diehl, S., and Stute, W. (1988). Kernel density and hazard function estimation in the presence of censoring. *J. Multivariate Anal.* **25** 299 310.

[8] Einmahl, J. H. J., and Mason, D. M. (1988). Laws of the iterated logarithm in the tails for weighted uniform empirical processes. *Ann. Probab.* **16** 126 141.

[9] Elkins, T. (1986). Cubical and spherical estimation of multivariate probability density. *J. Amer. Statist. Assoc.* **63** 1495 1513.

[10] Hall, P. (1981). Laws of the iterated logarithm for nonparametric density estimators. *Z. Wahrsch. verw. Gebiete* **56** 47-61.

[11] Hall, P. (1983). On near neighbor estimates of a multivariate density. *J. Multivariate Anal.* **13** 24-39.

[12] Härdle, W. (1984). A law of the iterated logarithm for nonparametric regression function estimators. *Ann. Statist.* **12** 624-635.

[13] Kiefer, J. (1972). Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$. In *Proc. Sixth Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, pp. 227-244. Univ. of California Press, Berkeley.

[14] Krieger, A. M., and Pickands, J., III (1981). Weak convergence and efficient density estimation at a point. *Ann. Statist.* **2** 1066-1078.

[15] Lo, S. H., Mack, Y. P., and Wang, J. L. (1989). Density and hazard rate estimation for censored data via strong representation of the Kaplan Meier estimator. *Probab. Theory Related Fields* **80** 461-473.

[16] LOFTSGAARDEN, D. O., AND QUESENBERRY, C. P. (1965). A nonparameteric estimate of a multivariate density function. *Ann. Math. Statist.* **5** 1049–1051.
[17] MACK, Y. P., AND ROSENBLATT, M. (1979). Multivariate k-nearest neighbor density estimates. *J. Multivariate Anal.* **9** 1–15.
[18] MACK, Y. P. (1983). Rate of strong uniform convergence of k-nn density estimates. *J. Statist. Plann. Inference* **9** 185–192.
[19] MIELNICZUK, J. (1986). Some asymptotic properties of kernel estimators of a density function in case of censored data. *Ann. Statist.* **13** 766–773.
[20] MOORE, D. S., AND HENRICHON, E. G. (1969). Uniform consistency of some estimates of a density function. *Ann. Math. Statist.* **40** 1499–1502.
[21] MOORE, D. S., AND YACKEL, J. W. (1977a). Consistency properties of nearest neighbor density estimate. *Ann. Statist.* **5** 143–154.
[22] MOORE, D. S., AND YACKEL, J. W. (1977b). Large sample properties of nearest neighbor density function estimators. In *Statistical Decision Theory and Related Topics* (S. S. Gupta and D. S. Moore, Eds.). Academic Press, New York.
[23] PRAKASA RAO, B. L. S. (1983). *Nonparametric Functional Estimation*. Academic Press, New York.
[24] SACKS, J., AND YLVISAKER, D. (1981). Asymptotically optimum kernels for density estimates at a point. *Ann. Statist.* **9** 334–346.
[25] SCHÄFER, H. (1986). Local convergence of empirical measures in the random censorship situation with applications to density and rate estimators. *Ann. Statist.* **14** 1240–1245.
[26] STUTE, W. (1982). The oscillation behavior of empirical processes. *Ann. Probab.* **10** 86–107.
[27] TERRELL, G., AND SCOTT, D. (1992). Variable kernel density estimator. *Ann. Statist.* **3** 1236–1265.
[28] WAGNER, T. J. (1973). Strong consistency of a nonparametric estimate of a density function. *IEEE Trans. Systems Man Cybernet.* **3** 289–290.
[29] WEGMAN, E. J., AND DAVIS, H. I. (1979). Remarks on some recursive estimators of a probability density. *Ann. Statist.* **7** 316–327.

Printed in Belgium
Uitgever: Academic Press, Inc.
Verantwoordelijke uitgever voor België:
Hubert Van Maele
Altenastraat 20, B-8310 Sint-Kruis