The Annals of Statistics
1991, Vol. 19, No. 4, 1813–1831

CENTRAL LIMIT THEOREMS FOR L_p DISTANCES OF KERNEL ESTIMATORS OF DENSITIES UNDER RANDOM CENSORSHIP

BY MIKLÓS CSÖRGŐ,[1] EDIT GOMBAY[1] AND LAJOS HORVÁTH

Carleton University, University of Alberta, and University of Utah and Szeged University

A sequence of independent nonnegative random variables with common distribution function F is censored on the right by another sequence of independent identically distributed random variables. These two sequences are also assumed to be independent. We estimate the density function f of F by a sequence of kernel estimators $f_n(t) = (\int_{-\infty}^{\infty} K((t-x)/h(n))\, d\hat{F}_n(x))/h(n)$, where $h(n)$ is a sequence of numbers, K is kernel density function and $\hat{F}_n$ is the product-limit estimator of F. We prove central limit theorems for $\int_0^T |f_n(t) - f(t)|^p \, d\mu(t)$, $1 \le p < \infty$, $0 < T \le \infty$, where μ is a measure on the Borel sets of the real line. The result is tested in Monte Carlo trials and applied for goodness of fit.

1. Introduction and results. Let $X_1^0, X_2^0, \dots$ be a sequence of independent, *nonnegative* random variables with common distribution function F. Another sequence (independent of the $\{X_i^0,\ i \ge 1\}$) $Y_1, Y_2, \dots$ of independent random variables with common continuous distribution function R censors the preceding one on the right, so that the observations available to us at the nth stage consist of the pairs (X_j, δ_j), $1 \le j \le n$, where $X_j = X_j^0 \wedge Y_j$ [$a \wedge b = \min(a, b)$] and δ_j is the indicator of the event $\{X_j = X_j^0\}$. Survival data in clinical trials or failure time data in reliability studies, for example, are often subject to such censoring [cf., e.g., Kalbfleisch and Prentice (1980)]. Based on these randomly censored data, the nonparametric maximum likelihood estimator of F is the product-limit estimator $\hat{F}_n$, first introduced by Kaplan and Meier (1958) and defined by

$$1 - \hat{F}_n(t) = \begin{cases} \displaystyle\prod_{\{1 \le j \le n:\, X_j \le t\}} \left(\frac{n - N_{j,n} - 1}{n - N_{j,n}} \right)^{\delta_j}, & \text{if } t < X_{n,n}, \\ 0, & \text{if } t \ge X_{n,n}, \end{cases}$$

where $X_{n,n} = X_1 \vee \cdots \vee X_n$ and $N_{j,n} = \#\{k: 1 \le k \le n: X_k < X_j\}$, with $a \vee b = \max(a, b)$.

Received October 1989; revised November 1990.
[1]Research supported by an NSERC Canada grant.
AMS 1980 *subject classifications*. Primary 60F05, 62G10; secondary 60G15.
Key words and phrases. Censored data, kernel estimator, L_1 norm, Wiener process, strong approximation.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Statistics*.
www.jstor.org

In survival and failure time data analysis, the random censorship model is one of the most accepted and popular models in use. The asymptotic and finite sample properties of $\hat{F}_n$ and the product-limit process $\alpha_n(t) = n^{1/2}(\hat{F}_n(t) - F(t))$ have been investigated in a series of papers. For surveys on this topic, we refer to Chapter 8 in M. Csörgő (1983) and Chapter 7 in Shorack and Wellner (1986), as well as to Gu and Lai (1990), where a number of new results are also proven of course.

The estimation of the density function $f(t) = F'(t)$ is also an important tool in stochastic analysis. For recent advances in this area, we refer to Devroye (1987), Devroye and Györfi (1985) and Györfi, Härdle, Sarda and Vieu (1989). In his preface, Devroye (1987) explains where and how density estimates are applied. He, for example, lists exploratory data analysis, probability theory, detection problems, pattern recognition, estimating a tail probability, clustering analysis and simulation as areas of application. In this paper, we are interested in estimating the density function f under random censorship by a sequence of kernel estimators f_n defined by

$$f_n(t) = \frac{1}{h(n)} \int_{-\infty}^{\infty} K\left(\frac{t-x}{h(n)}\right) d\hat{F}_n(x), \tag{1.1}$$

where $h(n)$ is a sequence of positive numbers and K is a density function (kernel). Such kernel estimates of a density, based on complete samples, were introduced by Rosenblatt (1956) and have been extensively studied in the literature, for example, in the just quoted three books, since the appearance of Parzen (1962). Limit theorems for the L_∞ (i.e., sup-norm) and L_2 deviations between f_n and f under random censorship can be found in Blum and Susarla (1980), Burke (1983), Burke and Horváth (1984), Yandell (1983) and in the references of these papers.

In this exposition we study the L_p distance

$$I_n(T,p) = \int_0^T |f_n(t) - f(t)|^p \, d\mu(t), \qquad 1 \le p < \infty, \tag{1.2}$$

where $0 < T \le \infty$ and μ is a measure on the Borel sets of $\mathbb{R}$. The books of Devroye and Györfi (1985) and Devroye (1987) contain several results on the L_1 consistency of density estimators based on complete samples. A remarkable central limit theorem for the L_1 distance of Grenander's maximum likelihood estimate [Grenander (1956)] for monotone densities concentrated on a bounded interval $[0, B]$ is due to Groeneboom (1985). It is discussed in Devroye and Györfi (1985) and studied in-depth by Devroye (1987). However, the asymptotic distribution of the L_1 deviation between the density function and its kernel estimator has been an open problem. Recently, M. Csörgő and Horváth (1988) obtained a central limit theorem for L_p distances $(1 \le p < \infty)$ of kernel estimators based on complete samples.

The main aim of this paper is to prove a similar result for $I_n(T, p)$. The most important implications of our general results are for the L_1 error, $I_n(T, 1)$, which is always between 0 and 2 if $\mu(t) = t$ and K is a density itself.

It also has many further desirable properties. For example, monotone transformations of the coordinate axes leave the L_1 distance unaffected. This means that $I_n(T, 1)$ with $\mu(t) = t$ is a universal measure of closeness between densities. In contrast, the L_p distance with $p \neq 1$ is not even invariant under a simple rescaling of the axes. Also, the total variation of the measures induced by the densities f_n and f is just half of the L_1 distance between f_n and f. For details and further discussion on distances between densities, we refer to Devroye and Györfi (1985), Chapter 1, and Devroye (1987), Chapter 1. The L_1 distance is not only easily interpreted, it is also easily visualized. The visual impression of the distance between the plots of f_n and f is precisely $I_n(T, 1)$, the area between the curves. These and further arguments in Devroye and Györfi (1985) and Devroye (1987) lead us most convincingly to conclude that L_1 is the natural place for studying densities. This remains true in random censorship models as well, and for the very same reasons. However, when estimating a density under random censorship, the presence of the usually unknown distribution function R of the censoring r.v.'s inevitably results in creating additional difficulties. Nevertheless, our Theorems 1 and 3 also cover the optimal case of $nh^5(n) \to$ a constant when T of $I_n(T, p)$ is finite, and possibly infinite respectively. Theorem 2 can be viewed as a generalization of Bickel and Rosenblatt (1973).

Theorems 1, 2 and 3 can, for example, be used to test for goodness of fit whenever both F and R are completely specified. However, Theorems 1 and 2, as well as a special case of Theorem 3, remain valid (cf. Corollary 2 and Remark 1 respectively) if we replace the unknown R by its product-limit estimator $\hat{R}_n$ in their respective statements.

The use of weighted L_p distances [$d\mu(t)$ instead of dt in the L_p integrals] enables us to transform Theorem 2 into such a form where the asymptotic mean and variance of $I_n(T, p)$ do not depend on the unknown distribution function R of the censoring r.v.'s (cf. Corollary 1 and Section 2 for application of the latter to goodness of fit). In a similar vein, in L_1 the asymptotic variance of Theorem 3 can be made independent of not only R, but also of F as well in a special case (cf. Remark 1).

Before stating our result we introduce further notations and then list all the assumptions used in the first part of this paper. Let $G(t) = P\{X_1 \leq t\}$, $\tilde{F}(t) = P\{X_1 \leq t \text{ and } \delta_1 = 1\}$ and $d(t) = \int_0^t (1 - G(s))^{-2}\, d\tilde{F}(s)$. Then, by independence, we have

$$1 - G(t) = (1 - F(t))(1 - R(t)), \tag{1.3}$$

$$\tilde{F}(t) = \int_0^t (1 - R(s))\, dF(s) \tag{1.4}$$

and

$$d(t) = \int_0^t (1 - F(s))^{-2} (1 - R(s))^{-1}\, dF(s). \tag{1.5}$$

When studying $I_n(T, p)$ of (1.2) in the case of $0 < T < \infty$, we assume that, concerning the distributions F and R, the following conditions hold true:

C.1. $G(T^*) < 1$, $0 < T < T^* < \infty$.
C.2. (i) f is continuous $[0, T^*]$;
(ii) $\sup_{0 < t \le T^*} |f^{(1)}(t)/f^{1/2}(t)| < \infty$;
(iii) $\sup_{0 < t \le T^*} r(t) < \infty$;
(iv) $\sup_{0 < t \le T^*} |f^{(2)}(t)| < \infty$, where $f^{(1)}$ and $f^{(2)}$ are the first and second derivatives of f, $r = R'$ and $T < T^*$, $G(T^*) < 1$.
C.3. $d\mu(t) = w(t)\,dt$, where $w(t) \ge 0$ and continuous on $[0, T^*]$, where $T < T^*$ and $G(T^*) < 1$.

The kernel function satisfies:

C.4. K is bounded and vanishes outside of a finite interval;
C.5. $\int_{-\infty}^{\infty} K(t)\,dt = 1$;
C.6. $\int_{-\infty}^{\infty} K^2(t)\,dt > 0$;
C.7. $\int_{-\infty}^{\infty} |t|\,d|K(t)| < \infty$;
C.8. $\int_{-\infty}^{\infty} tK(t)\,dt = 0$;

Throughout this paper $N = N(0,1)$ stands for a standard normal r.v. Let

$$\bar{f}_{(n)}(t) = \frac{1}{h(n)} \int_{-\infty}^{\infty} K\left(\frac{t-x}{h(n)}\right) dF(x),$$

$$\phi(x) = (2\pi)^{-1/2} \exp(-x^2/2),$$

$$\psi(u, x, y) = (2\pi)^{-1}(1-u^2)^{-1/2} \exp\left(-(2(1-u^2))^{-1}(x^2 - 2uxy + y^2)\right),$$

$$D = \left(\int_{-\infty}^{\infty} K^2(t)\,dt\right)^{1/2}, \qquad L = \frac{1}{2}\int_{-\infty}^{\infty} u^2 K(u)\,du,$$

$$l(t) = (1-F(t))(d'(t))^{1/2} = (1-R(t))^{-1/2}(f(t))^{1/2},$$

$$m_n(T,p) = \int_{-\infty}^{\infty}\int_0^T \left|l(t)Dx + (nh(n))^{1/2}\left(\bar{f}_{(n)}(t) - f(t)\right)\right|^p w(t)\phi(x)\,dt\,dx,$$

$$\eta(t) = \frac{\int_{-\infty}^{\infty} K(u)K(t+u)\,du}{\int_{-\infty}^{\infty} K^2(u)\,du},$$

$$\sigma_1^2(p) = (2\pi)^{-1}\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |xy|^p (1-\eta^2(u))^{-1/2} \times \exp\left(-\frac{1}{2(1-\eta^2(u))}(x^2 - 2xy\eta(u) + y^2)\right) dx\,dy - (E|N|^p)^2\right\} du,$$

$$\sigma^2(T,p) = \sigma_1^2(p)\int_0^T (1-R(t))^{-p} f^p(t) w^2(t)\,dt \left(\int_{-\infty}^{\infty} K^2(u)\,du\right)^p$$

and

$$\theta^2(T,p) = \int_0^T \int_{\mathbb{R}^3} \left| l^2(t) D^2 xy + l(t) DLf^{(2)}(t)(x+y)C_0 + C_0^2 L^2 \left(f^{(2)}(t)\right)^2 \right|^p$$
$$\times w^2(t)\left(\psi(\eta(u), x, y) - \phi(x)\phi(y)\right) du\, dy\, dx\, dt.$$

Now we can state the main result of our paper when $0 < T < \infty$.

THEOREM 1. *Let* $1 \le p < \infty$. *We assume that* C.1–C.8 *hold and*

$$h(n) \to 0, \qquad n^{-1/2}(h(n))^{-1} \log n \to 0, \qquad n \to \infty. \tag{1.6}$$

(i) *If* $nh^5(n) \to 0$, $n \to \infty$, *then, as* $n \to \infty$, *we have*

$$\left(h(n)\sigma^2(T,p)\right)^{-1/2}\left\{(nh(n))^{p/2} I_n(T,p) - m_n(T,p)\right\} \to_{\mathscr{D}} N(0,1).$$

(ii) *If* $nh^5(n) \to C_0^2 > 0$, $n \to \infty$, *then, as* $n \to \infty$, *we have*

$$\left(h(n)\theta^2(T,p)\right)^{-1/2}\left\{(nh(n))^{p/2} I_n(T,p) - m_n(T,p)\right\} \to_{\mathscr{D}} N(0,1).$$

We say that the estimation is undersmoothed when $nh^5(n) \to 0$. A popular choice of $h(n)$ is the value which minimizes $EI_n(T,2)$ or $EI_n(T,1)$. The "optimal" choice of $h(n)$ satisfies $nh^5(n) \to C_0^2 > 0$, where C_0 depends on F and R in a very intricate way.

In Theorem 1 the centralizing sequence depends on the numerical error $\bar{f}_{(n)}(t) - f(t)$. Adding a little bit more to our conditions on h in (1.6), we get the following theorem which is a generalization of an earlier result of Bickel and Rosenblatt (1973). Let

$$m(T,p) = E|N|^p \left(\int_{-\infty}^{\infty} K^2(t)\, dt\right)^{p/2} \int_0^T (1 - F(t))^p (d'(t))^{p/2}\, d\mu(t)$$
$$= E|N|^p \left(\int_{-\infty}^{\infty} K^2(t)\, dt\right)^{p/2} \int_0^T (1 - H(t))^{-p/2} (f(t))^{p/2}\, d\mu(t).$$

THEOREM 2. *Let* $1 \le p < \infty$. *Assume that* C.1–C.8 *hold and*

$$h(n) \to 0, \qquad n^{-1/2}(h(n))^{-1}\log n \to 0,$$
$$n(h(n))^4 \to 0, \qquad n \to \infty. \tag{1.7}$$

Then, as $n \to \infty$, *we have*

$$\left(h(n)\sigma^2(T,p)\right)^{-1/2}\left\{(nh(n))^{p/2} I_n(T,p) - m(T,p)\right\} \to_{\mathscr{D}} N(0,1).$$

We note that if $h(n) = n^{-\alpha}$ with $\frac{1}{4} < \alpha < \frac{1}{2}$, then (1.7) holds.

A natural choice for $w(t)\, dt = d\mu(t)$ is

$$w(t)\, dt = d\mu(t) = (1 - R(t))^{p/2}\, dt, \tag{1.8}$$

for then the normalizing constants $m(T,p)$ and $\sigma(T,p)$ do not depend on R, which is usually unknown. Namely by (1.8) and the above definitions of $m(T,p)$ and $\sigma^2(T,p)$ respectively, we have

$$(1.9)\qquad m(T,p) = E|N|^p\left(\int_{-\infty}^{\infty} K^2(t)\,dt\right)^{p/2} \int_0^T (f(t))^{p/2}\,dt$$

and, with $\sigma_1^2(p)$ as before,

$$(1.10)\qquad \sigma^2(T,p) = \sigma_1^2(p)\int_0^T f^p(t)\,dt\left(\int_{-\infty}^{\infty} K^2(u)\,du\right)^p.$$

These normalizing constants are like those of the noncensored case on taking $\mu(t) = t$ there [cf. Theorem of M. Csörgő and Horváth (1988)]. But now we have $(1 - R(t))^{p/2}\,dt$ for $d\mu(t)$ in the definition of $I_n(T,p)$. It is, however, easy to estimate R from the sample by its product-limit estimator $\hat{R}_n$. Namely, under condition C.1, the embedding theorem of Burke, S. Csörgő and Horváth (1981, 1988) (cf. Lemma 4 in Section 3 here) implies

$$(1.11)\qquad \sup_{0\le t\le T}\left|\hat{R}_n(t) - R(t)\right| = O_P(n^{-1/2}).$$

Let [cf. (1.2) with $d\mu(t)$ as in (1.8)]

$$I_n(T,p,R) = \int_0^T |f_n(t) - f(t)|^p (1 - R(t))^{p/2}\,dt$$

and

$$\hat{I}_n(T,p,\hat{R}_n) = \int_0^T |f_n(t) - f(t)|^p \left(1 - \hat{R}_n(t)\right)^{p/2}\,dt.$$

Then

$$(1.12)\qquad \begin{aligned} &\left|\hat{I}_n(T,p,\hat{R}_n) - I_n(t,p,R)\right| \\ &\le \int_0^T |f_n(t) - f(t)|^p \frac{\left|\left(1-\hat{R}_n(t)\right)^p - (1-R(t))^p\right|}{\left(1-\hat{R}_n(t)\right)^{p/2} + (1-R(t))^{p/2}}\,dt \\ &\le \frac{O_P(n^{-1/2})}{\left(1-\hat{R}_n(T)\right)^{p/2} + (1-R(T))^{p/2}} \int_0^T |f_n(t) - f(t)|^p\,dt \\ &= O_P(n^{-1/2})\int_0^T |f_n(t) - f(t)|^p p\,dt, \end{aligned}$$

where we have used the inequality $|\,|a|^p - |b|^p| \le p2^{p-1}|a-b|^p + p2^{p-1}|b|^{p-1}|a-b|$, $p \ge 1$, in combination with (1.11). Our Theorem 2 is true for $I_n(T,p) = \int_0^T |f_n(t) - f(t)|^p\,dt$, and the latter combined with (1.12) gives

$$\begin{aligned} &\left((nh(n))^{p/2}/(h(n))^{1/2}\right)\left|\hat{I}_n(T,p,\hat{R}_n) - I_n(T,p,R)\right| \\ &= O_P\left((nh(n))^{-1/2}\right) = o_P(1). \end{aligned}$$

Hence we have the following corollary.

COROLLARY 1. *Let* $1 \le p < \infty$, *assume* C.1–C.8 [C.3 *with* $d\mu(t)$ *as in* (1.8)] *and* (1.7). *Then with* $m(T,p)$ *and* $\sigma^2(T,p)$ *as in* (1.9) *and* (1.10) *respectively, we have*

$$(h(n)\sigma^2(T,p))^{-1/2}\left\{(nh(n))^{p/2}\hat{I}_n\left(T,p,\hat{R}_n\right) - m(T,p)\right\} \to_{\mathscr{D}} N(0,1).$$

When working with complete samples, the beauty of the mentioned central limit theorem of Groeneboom (1985) for the L_1 distance of Grenander's maximum likelihood estimate for monotone densities concentrated on a bounded interval [0, B], and also that of Theorem of M. Csörgő and Horváth (1988) in the case of $\mu(t) = t$ for the L_1 distance of kernel estimators for smooth densities over the whole real line, is that the respective asymptotic variances of these L_1 distances are constants which are independent of f. The comparable case under random censorship is summarized by Corollary 1 above. Here [cf. (1.10)]

$$\sigma^2(T,1) = \sigma_1^2(1)\int_0^T f(t)\,dt\int_{-\infty}^{\infty} K^2(u)\,du$$

would reduce to $\sigma_1^2(1)\int_{-\infty}^{\infty}K^2(u)\,du$, if we could only take $T = \infty$. In any case, the latter is of course an upper bound for the former. For a solution of this problem, we refer to Theorem 3 and Remark 1.

Another way of looking at the problem of dependence on R is to continue working with the definition of I_n as in (1.2) while estimating $m(T,p,R) := m(T,p)$ and $\sigma^2(T,p,R) := \sigma^2(T,p)$ in Theorem 2 by

$$\begin{aligned}\hat{m}_n(T,p) &:= m\left(T,p,\hat{R}_n\right)\\ &= E|N|^p\left(\int_{-\infty}^{\infty} K^2(t)\,dt\right)^{p/2}\int_0^T\left(1-\hat{R}_n(t)\right)^{-p/2}(f(t))^{p/2}\,d\mu(t)\end{aligned}$$

and

$$\begin{aligned}\sigma_n^2(T,p) &:= m\left(T,p,\hat{R}_n\right)\\ &= \sigma_1^2(p)\int_0^T\left(1-\hat{R}_n(t)\right)^{-p} f^p(t)w^2(t)\,dt\left(\int_{-\infty}^{\infty} K^2(u)\,du\right)\end{aligned}$$

respectively. Similarly, in the case of Theorem 1, we may want to continue working with $I_n(T,p)$ as in (1.2) while estimating $m_n(T,p,R) := m_n(T,p)$, $\sigma^2(T,p,R) := \sigma^2(T,p)$ and $\theta^2(T,p,R) := \theta^2(T,p)$ by $\hat{m}_n(T,p) := m_n(T,p,\hat{R}_n)$, $\sigma_n^2(T,p) := \sigma^2(T,p,\hat{R}_n)$ and $\theta_n^2(T,p) := \theta^2(T,p,\hat{R}_n)$ respectively. Now (1.11) and an argument like that preceding Corollary 1 yield

$$|\hat{m}_n(T,p) - m(T,p)| = O_P(n^{-1/2})$$

and

$$\left|\sigma_n^2(T,p) - \sigma^2(T,p)\right| = O_P(n^{-1/2}),$$

as well as

$$\left|\hat{\hat{m}}_n(T,p) - m_n(T,p)\right| = O_P(n^{-1/2})$$

and

$$\left|\theta_n^2(T,p) - \theta^2(T,p)\right| = O_P(n^{-1/2}).$$

Hence, as $n \to \infty$,

$$\sigma_n(T,p)/\sigma(T,p) \to_P 1, \qquad \theta_n(T,p)/\theta(T,p) \to_P 1,$$

$$|\hat{m}_n(T,p) - m(T,p)|/(h(n))^{1/2} = O_P\big((nh(n))^{-1/2}\big) = o_P(1)$$

and

$$\left|\hat{\hat{m}}_n(T,p) - m_n(T,p)\right|/(h(n))^{1/2} = O_P((nh(n))^{-1/2} = o_P(1).$$

Consequently, we have:

COROLLARY 2.

(i) *Theorem* 1 *remains true if we replace* $m_n(T,p)$, $\sigma^2(T,p)$ *and* $\theta^2(T,p)$ *by* $\hat{\hat{m}}_n(T,p)$, $\sigma_n^2(T,p)$ *and* $\theta_n^2(T,p)$ *respectively.*

(ii) *Theorem* 2 *remains true if we replace* $m(T,p)$ *and* $\sigma^2(T,p)$ *by* $\hat{m}_n(T,p)$ *and* $\sigma_n^2(T,p)$ *respectively.*

In Theorems 1 and 2 and also in Corollary 1, we studied the behavior of the density estimator $I_n(T,p)$ on $[0,T]$, where $G(T) < 1$. This means that we have noncensored observations after T with positive probability. Hence we let $T_F = \inf\{t\colon F(t) = 1\} \le \infty$, and we will consider the behavior of f_n on $[0, T_F)$ via $I_n(T_F, p)$. The estimation of f on its support $[0, T_F)$ requires further assumptions. It is well known [cf. S. Csörgő and Horváth (1983) and Gu and Lai (1990)] that the product limit estimator is not necessarily a strongly uniformly consistent estimator of F on its support. Thus, for the sake of studying $I_n(T_F, p)$, we must impose stronger conditions on F, R and μ. In fact, instead of C.1–C.3 we assume the following conditions:

C.9. $T_F \le T_R$, where $T_R = \inf\{t\colon R(t) = 1\}$.

C.10. $\limsup_{t \to T_F}(1 - F(t))^{\beta}/(1 - R(t)) < \infty$ for some $0 \le \beta < 1$.

C.11. (i) f is uniformly bounded on $[0, T_F)$ and is monotone in a neighborhood of T_F;

(ii) $\sup_{0 < t < T_F} r(t) < \infty$;

(iii) $\sup_{0 \le t \le T} |f^{(1)}(t)/f^{1/2}(t)| < \infty$ for all $T < T_F$;

(iv) $\sup_{0 \le t < T_F} |f^{(2)}(t)| < \infty$.

C.12. $d\mu(t) = w(t)\,dt$, where $w(t) \ge 0$, continuous on $[0, T_F)$, and there is an $\varepsilon > 0$ such that

$$\limsup_{t \to T_F} \frac{t^{1+\varepsilon} w(t)}{(1 - R(t))^{2p}} < \infty.$$

We note that under conditions C.4–C.12, $m_n(T_F, p)$, $\sigma^2(T_F, p)$ and $\theta^2(T_F, p)$ are finite. Also, if $R(T_F) < 1$, then C.10 holds with $\beta = 0$ and C.12 is a simple tail condition on $w(t)$.

THEOREM 3. *Let* $1 \le p < \infty$. *We assume that* C.4–C.12 *hold and*

$$h(n) \to 0, \qquad nh^4(n) \to \infty, \qquad n^{1/2+\kappa}/h(n) \to 0 \tag{1.13}$$

for some $\kappa > \beta/(\beta + 1)$.

(i) *If* $nh^5(n) \to 0$, $n \to \infty$, *then, as* $n \to \infty$, *we have*

$$\left(h(n)\sigma^2(T_F, p)\right)^{-1/2}\left\{(nh(n))^{p/2} I_n(T_F, p) - m_n(T_F, p)\right\} \to_{\mathscr{D}} N(0,1).$$

(ii) *if* $nh^5(n) \to C_0^2 > 0$, $n \to \infty$, *then, as* $n \to \infty$, *we have*

$$\left(h(n)\theta^2(T_F, p)\right)^{-1/2}\left\{(nh(n))^{p/2} I_n(T_F, p) - m_n(T_F, p)\right\} \to_{\mathscr{D}} N(0,1).$$

REMARK 1. If $T_F < T_R$, then C.10 holds with $\beta = 0$ and we can choose $w(t) = (1 - R(t))^{p/2}$ in C.12. In this special case we have

$$\sigma^2(T_F, p) = \sigma_1^2(p) \int_0^{T_F} f^p(t)\, dt \left(\int_{-\infty}^{\infty} K^2(u)\, du \right), \tag{1.14}$$

and hence also

$$\sigma^2(T_F, 1) = \sigma_1^2(1) \int_{-\infty}^{\infty} K^2(u)\, du. \tag{1.15}$$

Thus *the asymptotic variance of the* L_1 *distance is a constant which is not a function of the unknown distribution functions* F *and* R. As to the problem of now having $(1 - R(t))^{p/2}\, dt$ for $d\mu(t) = w(t)\, dt$ in the definition of $I_n(T_F, p)$, unfortunately it is not immediate here to prove a complete analogue of Corollary 1. Nevertheless, assuming further conditions as in Theorem 1 of Gu and Lai (1990), we may estimate R again from the sample by its product-limit estimator $\hat{R}_n$ and use $(1 - \hat{R}_n(t))^{p/2}$ instead of $w(t) = (1 - R(t))^{p/2}$ in the definition of $I_n(T_F, p)$.

2. Application. It is Corollary 1 to Theorem 2 with the special choice of weight function as in (1.8) that offers itself most readily for immediate applications. In this case the normalizing constants $m(T, p)$ and $\sigma(T, p)$ can be easily calculated for integers $p \ge 1$. Let $\eta(u)$ be as in Section 1. Since K is bounded and vanishes outside of a finite interval, $\eta(u)$ has similar properties. Let X, Y have bivariate normal distribution with correlation coefficient $\eta(u)$, where $\eta(u) = 0$ if $u \notin (a, b)$ for some constants a and b. Then we get

$$\sigma_1^2(p) = \int_a^b E|XY|^p\, du - (b - a)\left(E|X|^p\right)^2.$$

The joint absolute moments of X and Y were calculated by Kamat (1953) for integer values. When $p = 1$,

$$E|XY| = 2\{(1 - \eta(u))^{1/2} + \eta(u)\arcsin(\eta(u))\}/\pi, \tag{2.1}$$

and if we choose the so-called naive kernel

$$K(u) = \begin{cases} 1, & |u| \le 0.5, \\ 0, & |u| > 0.5, \end{cases} \tag{2.2}$$

then by integration

$$\sigma_1^2(1) = 2 + \pi/4.$$

If (2.1) is not analytically integrable for some kernel $K(u)$, then $\sigma_1^2(1)$ is a one-dimensional integral that is most likely smooth enough to cause no problems in numerical integration.

We performed Monte Carlo trials to see how Corollary 1 works for goodness of fit in the most important case of $p = 1$. Both (2.2) and

$$K(u) = \begin{cases} 1.5 - 6u^2, & |u| \le 0.5, \\ 0, & |u| > 0.5, \end{cases} \tag{2.3}$$

were used as kernels of density estimation in our experiments. The bandwidth was $h(n) = n^{-\alpha}$, where various α values were chosen in the interval $(\frac{1}{4}, \frac{1}{2})$. The results showed only minor sensitivity to the choice of these kernels and small bandwidth changes. The censoring distribution was exponential with mean value $\lambda > 0$, that is, $r(x) = (1/\lambda)\exp(-x/\lambda)$, $x > 0$, generated with various values of λ in the interval [3.5, 6]. The data subject to this random censorship were generated either from the exponential or from the Weibull distribution of density function $f(x) = (\gamma/\beta)(x/\beta)^{\gamma-1}\exp\{-(x/\beta)^\gamma\}$, $x > 0$, Weib(β, γ), $\beta, \gamma > 0$, using various values for the parameters. We will simply call anyone of these the *data-generating distribution*. The hypothetical densities under H_0 were completely specified members of either of these two families. We call any of these the *hypothetical distribution*. For numerical integration IMSL routines were used. Three types of Monte Carlo experiments were performed and on the basis of 30 to 50 runs of samples of size 200 to 2000, our findings can be summarized as follows.

Whenever the data-generating and hypothetical distributions were the same, for example, like Exp(4) or like Weib(4, 3), we obtained "perfectly" fitting values, packed around 0.

"Moderate deviations," such as a Weib(5, 5) data-generating distribution viewed as a Weib(4, 3) hypothetical distribution, or Weib(3, 4) viewed as Weib(4, 3) or Exp(6) viewed as Exp(4), were not shown to be significantly different shapes.

However, when the data-generating distributions Weib(4, 3), Exp(4), Exp(1) and Weib(1.5, 2) were tested as Exp(4), Weib(4, 3), Exp(4) and Weib(4, 3) hypothetical distributions respectively, then the observed values were found to be consistently far too large in absolute value for being viewed as $N(0,1)$ readings.

A survival experiment was performed on mice by Upton et al. (1969). The resulting data are censored due to serial sacrifice both in the treatment and in the control group. These have been examined by several authors. We refer to Yandell (1983) for references. Yandell rejected the hypothesis that the death rate is constant. We rescaled the data to have mean 4 and tested the hypothesis H_0 that the data follow exponential distribution. For the treated and control mice groups respectively, we got observed values 17.69 and 15.36 of random variables that are standard normal under H_0, which clearly indicates that H_0 is false. The number of observations were 1454 and 1080 respectively.

3. Proofs. We will assume throughout, without loss of generality, that all random variables and processes of the present paper are defined on the same probability space [cf., e.g., de Acosta (1982)]. Also, C stands for a generic constant whose value may differ from line to line. For the sake of notational simplicity, we assume $K(u) = 0$ if $u \notin (-1, 1)$. Let $\{W(t),\ -\infty < t < \infty\}$ be a Wiener process [cf. Doob (1953), page 97] and define the Gaussian process

$$\Gamma_n^{(1)}(t) = \int_{-\infty}^{\infty} K\left(\frac{t-x}{h(n)}\right) dW(x).$$

Let

$$\gamma^2(T,p) = \begin{cases} \sigma^2(T,p), & \text{if } nh^5(n) \to 0, \\ \theta^2(T,p), & \text{if } nh^5(n) \to C_0^2 > 0, \end{cases}$$

and

$$g_{(n)}(t) = (nh(n))^{1/2}\big(\bar{f}_{(n)}(t) - f(t)\big).$$

LEMMA 1. *We assume that the conditions of Theorem* 1 *hold. Then, as* $n \to \infty$, *we have*

$$(h(n)\gamma^2(T,p))^{-1/2}\left\{\int_0^T \left|l(t)h^{-1/2}(n)\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p d\mu(t) - m_n(T,p)\right\} \to_{\mathscr{D}} N(0,1). \tag{3.1}$$

PROOF. The proof of Lemma 1 goes along the lines of the proof of Lemmas 1 and 2 in M. Csörgő and Horváth (1988). Hence it is omitted. □

The following inequality will be useful later on. Let $1 \le p < \infty$. Then for functions q and u in L_p we have

$$\int_0^T \Big| |q(t)|^p - |u(t)|^p \Big| \, d\mu(t)$$

$$\le p2^{p-1} \int_0^T |q(t) - u(t)|^p \, d\mu(t) \tag{3.2}$$

$$+ p2^{p-1} \left(\int_0^T |u(t)|^p \, d\mu(t) \right)^{1-1/p} \left(\int_0^T |q(t) - u(t)|^p \, d\mu(t) \right)^{1/p}.$$

Let

$$\rho(t) = (d'(t))^{1/2} = (1 - F(t))^{-1} (1 - R(t))^{-1/2} (f(t))^{1/2}$$

and

$$\Gamma_n^{(2)}(t) = \int_{-\infty}^{\infty} K\left(\frac{t - x}{h(n)} \right) \rho(x) \, dW(x).$$

LEMMA 2. *We assume that the conditions of Theorem* 1 *hold. Then, as* $n \to \infty$, *we have*

$$(h(n)\gamma^2(T,p))^{-1/2} \left\{ \int_0^T |(1 - F(t)) h^{-1/2}(n) \Gamma_n^{(2)}(t) + g_{(n)}(t)|^p \, d\mu(t) - m_n(T,p) \right\} \to_{\mathscr{D}} N(0,1). \tag{3.3}$$

PROOF. Integration by parts gives

$$\begin{aligned}
\Gamma_n^{(2)}(t) &= \rho(t)\Gamma_n^{(1)}(t) \\
&\quad + \int_{-1}^{1} (W(t - yh(n)) - W(t))(\rho(t - yh(n)) - \rho(t)) \, dK(y) \\
&\quad + \int_{-1}^{1} (W(t - yh(n)) - W(t)) K(y) \, d\rho(t - yh(n)) \\
&= \rho(t)\Gamma_n^{(1)}(t) + A_n^{(1)}(t) + A_n^{(2)}(t).
\end{aligned}$$

Using now continuity of Wiener process and the mean value theorem, we get

$$\begin{aligned}
\sup_{0<t\le T} |A_n^{(1)}(t)| &\le Ch(n) \sup_{0<t\le T} \sup_{0<s\le h(n)} |W(t+s) - W(t)| \\
&\quad \times \sup_{0<t\le T+h(n)} |f^{(1)}(t)/f^{1/2}(t)| \\
&= o_P(1)h(n).
\end{aligned} \tag{3.4}$$

A similar argument gives

$$\sup_{0<t\le T} \left| A_n^{(2)}(t) \right| = o_P(1)h(n). \tag{3.5}$$

Hence, from (3.4) and (3.5), we obtain

$$\int_0^T \left| \Gamma_n^{(2)}(t) - \rho(t)\Gamma_n^{(1)}(t) \right|^p (1 - F(t))^p \, d\mu(t) = o_P(1)(h(n))^p. \tag{3.6}$$

Applying Lemma 1, (3.2) and (3.6), we have

$$\begin{aligned} &\int_0^T \left| \left|\Gamma_n^{(2)}(t)\right|^p - \rho^p(t)\left|\Gamma_n^{(1)}(t)\right|^p \right| (1 - F(t))^p \, d\mu(t) \\ &\quad = o_P(1)(h(n))^p \\ &\qquad + o_P(1)(h(n)) \left(\int_0^T l^p(t) \left|\Gamma_n^{(1)}(t)\right|^p d\mu(t) \right)^{1-1/p} \\ &\quad = o_P(1)(h(n))^{(p+1)/2}. \end{aligned} \tag{3.7}$$

Using (3.2), Lemma 1 and (3.7), we get

$$\begin{aligned} &\int_0^T \Big| \left|(1 - F(t))h^{-1/2}(n)\Gamma_n^{(2)}(t) + g_{(n)}(t)\right|^p \\ &\qquad - \left|l(t)h^{-1/2}(n)\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p \Big| \, d\mu(t) \\ &\quad \le p2^{p-1} \int_0^T \left|(1 - F(t))h^{-1/2}(n)\Gamma_n^{(2)}(t) - l(t)h^{-1/2}(n)\Gamma_n^{(1)}(t)\right|^p d\mu(t) \\ &\qquad + p2^{p-1} \left(\int_0^T \left|l(t)h^{-1/2}(n)\Gamma_n^{(1)}(t) + g_{(n)}(t)\right|^p d\mu(t) \right)^{1-1/p} \\ &\qquad \times \left(\int_0^T \left|(1 - F(t))h^{-1/2}(n)\Gamma_n^{(2)}(t) - l(t)h^{-1/2}(n)\Gamma_n^{(1)}(t)\right|^p d\mu(t) \right)^{1/p} \\ &\quad = o_P(1)h^{1/2}(n). \end{aligned}$$

By Lemma 1 the proof of Lemma 2 is complete. □

Let

$$\Gamma_n^{(3)}(t) = \int_{-\infty}^{\infty} (1 - F(x))W(d(x)) \, dK\left(\frac{t - x}{h(n)} \right).$$

LEMMA 3. *We assume that the conditions of Theorem 1 hold. Then, as $n \to \infty$, we have*

$$(\gamma^2(T,p)h(n))^{-1/2} \left\{ \int_0^T \left| h^{-1/2}(n)\Gamma_n^{(3)}(t) + g_{(n)}(t) \right|^p d\mu(t) - m_n(T,p) \right\} \to_{\mathscr{D}} N(0,1).$$

PROOF. It is easy to see that

$$\Gamma_n^{(3)}(t) = \int_{-1}^{1} (1 - F(t - uh(n)))W(d(t - uh(n)))\, dK(u).$$

Let

$$\Gamma_n^{(4)}(t) = (1 - F(t)) \int_{-\infty}^{\infty} W(d(t - uh(n)))\, dK(u).$$

A two-term Taylor expansion gives

$$\Gamma_n^{(3)}(t) - \Gamma_n^{(4)}(t) = h(n) f(t) \int_{-1}^{1} W(d(t - uh(n)))u\, dK(u)$$

$$+ \tfrac{1}{2} h^2(n) \int_{-1}^{1} f^{(1)}(\xi) W(d(t - uh(n)))u^2\, dK(u)$$

with $t \wedge (t - uh(n)) \le \xi \le t \vee (t - uh(n))$. We also have

$$\begin{aligned}
&\left|\Gamma_n^{(3)}(t) - \Gamma_n^{(4)}(t)\right|^p \\
&\quad\le 2^p (f(t))^p (h(n))^p \left| \int_{-1}^{1} W(d(t - uh(n)))u\, dK(u)\right|^p \\
&\qquad + C(h(n))^{2p} \sup_{0 \le t \le d(T + h(n))} |W(t)|^p \\
&\quad= A_n^{(3)}(t) + A_n^{(4)}.
\end{aligned}$$

Let

$$L(t) = \begin{cases} \int_{-1}^{1} u\, dK(u), & \text{if } -1 \le t \le 1, \\ 0, & \text{otherwise}. \end{cases}$$

Integration by parts gives

$$\left| \int_{-1}^{1} W(d(t - uh(n)))\, dL(u) \right| = \left| \int_{-\infty}^{\infty} L\left(\frac{t - x}{h(n)} \right) dW(d(x)) \right|,$$

and hence

$$\begin{aligned}
E\left| \int_{-1}^{1} W(d(t - uh(n)))\, dL(u) \right|^p &\le C \left(\int_{-\infty}^{\infty} L^2\left(\frac{t - x}{h(n)} \right) dd(x) \right)^{p/2} \\
&= C(h(n))^{p/2}.
\end{aligned}$$

Consequently,

$$EA_n^{(3)}(t) \le C(h(n))^{3p/2} \tag{3.8}$$

and we also have

$$EA_n^{(4)} \le C(h(n))^{2p}. \tag{3.9}$$

The Markov inequality with (3.8) and (3.9) implies

$$\left|\Gamma_n^{(3)}(t) - \Gamma_n^{(4)}(t)\right|^p = O_p(1)(h(n))^{3p/2}. \tag{3.10}$$

Using (3.2), Lemma 2 and (3.10), we get

$$\begin{aligned}
&\left|\int_0^T \left(\left|\Gamma_n^{(3)}(t)\right|^p - \left|\Gamma_n^{(4)}(t)\right|^p\right) d\mu(t)\right| \\
&\quad \le p2^{p-1}\int_0^T \left|\Gamma_n^{(3)}(t) - \Gamma_n^{(4)}(t)\right|^p d\mu(t) \\
&\qquad + p2^{p-1}\left(\int_0^T \left|\Gamma_n^{(4)}(t)\right|^p d\mu(t)\right)^{1-1/p}\left(\int_0^T \left|\Gamma_n^{(3)}(t) - \Gamma_n^{(4)}(t)\right|^p d\mu(t)\right)^{1/p} \\
&\quad = O_P(1)(h(n))^{3p/2} + O_P(1)(h(n))^{p/2+1} \\
&\quad = o_P(1)(h(n))^{(p+1)/2}.
\end{aligned}$$

Now we observe

$$\left\{\Gamma_n^{(4)}(t), 0 \le t \le T\right\} =_{\mathscr{D}} \left\{(1 - F(t))\Gamma_n^{(2)}(t), 0 \le t \le T\right\},$$

and hence Lemma 3 follows from Lemma 2 and (3.2). □

The following lemma is due to Burke, S. Csörgő and Horváth (1981, 1988).

LEMMA 4. *We assume that* C.1 *holds. Then we can define a sequence of Wiener processes* $\{W_n(x),\ x \ge 0\}$ *such that*

$$\sup_{0 \le t \le T^*} \left|\alpha_n(t) - (1 - F(t))W_n(d(t))\right| = O(n^{-1/2}\log n) \quad a.s.,$$

where

$$\alpha_n(t) = n^{1/2}\left(\hat{F}_n(t) - F(t)\right).$$

Let

$$\Gamma_n^{(5)}(t) = \int_{-\infty}^{\infty} (1 - F(x))W_n(d(x))\, dK\left(\frac{t-x}{h(n)}\right).$$

PROOF OF THEOREM 1. Lemmas 3, 4 and (3.2) imply

$$
\begin{aligned}
&\left|(nh(n))^{p/2}I_n(p) - \int_0^T \left|h^{-1/2}(n)\Gamma_n^{(5)}(t) + g_{(n)}(t)\right|^p d\mu(t)\right| \\
&\le \int_0^T \left| \left|h^{-1/2}(n)\int_{-\infty}^{\infty}\alpha_n(x)\,dK\left(\frac{t-x}{h(n)}\right) + g_{(n)}(t)\right|^p \right. \\
&\qquad\qquad \left. - \left|h^{-1/2}(n)\Gamma_n^{(5)}(t) + g_{(n)}(t)\right|^p \right| d\mu(t) \\
&\le p2^{p-1}h^{-p/2}(n)\int_0^T\left|\int_{-\infty}^{\infty}(\alpha_n(x) - (1-F(x))W_n(d(x)))\,dK \right. \\
&\qquad\qquad \left. \times\left(\frac{t-x}{h(n)}\right)\right|^p d\mu(t) \\
&\quad + p2^{p-1}h^{-1/2}(n)\left(\int_0^T\left|\int_{-\infty}^{\infty}(\alpha_n(x) - (1-F(x))W_n(d(x)))\,dK \right.\right. \\
&\qquad\qquad \left.\left. \times\left(\frac{t-x}{h(n)}\right)\right|^p d\mu(t)\right)^{1/p} \\
&\quad \times\left(\int_0^T\left|h^{-1/2}(n)\Gamma_n^{(5)}(t) + g_{(n)}(t)\right|^p d\mu(t)\right)^{1-1/p} \\
&= O_P(1)h^{p/2}(n)(n^{-1/2}h^{-1}(n)\log n)^p \\
&\quad + O_P(1)h^{1/2}(n)(n^{-1/2}h^{-1}(n)\log n).
\end{aligned}
$$

Now Theorem 1 follows immediately from Lemma 3. □

PROOF OF THEOREM 2. A two-term Taylor expansion gives

$$
\int_0^T \left|f(t) - \bar{f}_{(n)}(t)\right|^p d\mu(t) = O\left((h(n))^{2p}\right) \tag{3.11}
$$

as $n \to \infty$. Using the assumption $nh^4(n) \to 0$, we get from (3.2) that

$$
|m_n(T,p) - m(T,p)| = o(h^{1/2}(n)).
$$

Now Theorem 2 follows from Theorem 1. □

PROOF OF THEOREM 3. It is enough to show that

$$\lim_{T\to\infty}\lim_{n\to\infty} P\Big\{h^{-1/2}(n)\Big|(nh(n))^{p/2}\int_T^\infty |f_n(t)-f(t)|^p\,d\mu(t) \tag{3.12}$$
$$-\int_{-\infty}^{\infty}\int_T^\infty |l(t)Dx+g_{(n)}(t)|^p w(t)\varphi(x)\,dt\,dx\Big|>\delta\Big\}=0$$

for all $\delta>0$. Gu and Lai (1990), Corollary 3, constructed a sequence of independent processes $\{\xi_i(t),\ 0\le t\le T_F\}_{i=1}^{\infty}$ such that

$$\sup_{0\le t<T_F}\left|\alpha_n(t)-n^{-1/2}\sum_{i=1}^n \xi_i(t)\right| = O(n^{\kappa-1/2}) \quad \text{a.s.} \tag{3.13}$$

Let

$$\Gamma_n^{(6)}(t) = n^{-1/2}\sum_{i=1}^n \int_{-\infty}^{\infty}\xi_i(x)\,dK((t-x)/h(n)).$$

Using (3.13) instead of Lemma 4, we can argue as in the proof of Theorem 1 and get

$$\Big|(nh(n))^{p/2}\int_T^\infty |f_n(t)-f(t)|^p\,d\mu(t) \tag{3.14}$$
$$-\int_T^\infty \big|h^{-1/2}(n)\Gamma_n^{(6)}(t)+g_{(n)}(t)\big|^p\,d\mu(t)\Big|$$
$$=O_P(1)h^{1/2}(n)(n^{\kappa-1/2}/h(n)).$$

Next we show that

$$\lim_{T\to\infty}\limsup_{n\to\infty}\Big|E\int_T^\infty |h^{-1/2}(n)\Gamma_n^{(6)}(t)+g_{(n)}(t)|^p\,d\mu(t) \tag{3.15}$$
$$-\int_{-\infty}^{\infty}\int_T^\infty |l(t)Dx+g_{(n)}(t)|^p w(t)\varphi(x)\,dt\,dx\Big|\Big/h^{1/2}(n)=0$$

and

$$\lim_{T\to\infty}\limsup_{n\to\infty}\frac{1}{h(n)}\operatorname{var}\left(\int_T^\infty \big|h^{-1/2}(n)\Gamma_n^{(6)}(t)+g_{(n)}(t)\big|^p\,d\mu(t)\right)=0. \tag{3.16}$$

For every $t>0$, $\Gamma_n^{(6)}(t)$ is a sum of i.i.d. r.v.'s, so the computation of

$$E\big|h^{-1/2}(n)\Gamma_n^{(6)}(t)+g_{(n)}(t)\big|^p w(t)$$

is not too difficult. By condition C.12, $\Gamma_n^{(6)}(t)w^{1/p}(t)$ is a sum of i.i.d. bounded r.v.'s, so according to Petrov (1975), page 125, Theorem 13, the distribution of

$\Gamma_n^{(6)}(t)w^{1/p}(t)$ can be replaced by a normal distribution. Thus we have

$$\left|E\int_T^{\infty}\left|h^{-1/2}(n)\Gamma_n^{(6)}(t)+g_{(n)}(t)\right|^p w(t)\,dt - \int_{-\infty}^{\infty}\int_T^{\infty}\left|l(t)Dx+g_{(n)}(t)\right|^p w(t)\varphi(x)\,dt\,dx\right| = O\big(1/(nh(n))^{1/2}\big) = o(h^{1/2}(n)).$$

The proof of (3.16) is based on the same argument, but the calculations are more tedious. The computation of the variance in (3.16) requires the approximation of the distribution of $(w^{1/p}(t)\Gamma_n^{(6)}(t),\ w^{1/p}(s)\Gamma_n^{(6)}(s))$ via using multivariate normal distributions. In this case, instead of Petrov (1975), page 125, Theorem 13, we use Theorem 17.6 in Bhattacharya and Ranga Rao (1976), page 171, and obtain

$$\left|\operatorname{var}\left(\int_T^{\infty}\left|h^{-1/2}(n)\Gamma_n^{(6)}(t)+g_{(n)}(t)\right|^p d\mu(t)\right) - \frac{1}{h(n)}\int_T^{\infty}g_n^*(t)w(t)\,dt\right| = O\big(1/(nh^2(n))^{1/2}\big) = o(h(n)) \tag{3.17}$$

with some $g_n^*(t)$ such that under condition C.12 and (1.13)

$$\lim_{T\to\infty}\limsup_{n\to\infty}\int_T^{\infty}g_n^*(t)w(t)\,dt = 0.$$

Thus (3.17) implies (3.16), which completes the proof of Theorem 3. □

REFERENCES

DE ACOSTA, A. (1982). Invariance principles in probability for triangular arrays of B-valued random vectors and some applications. *Ann. Probab.* **10** 346–373.

BHATTACHARYA, R. N. and RANGA RAO, S. (1976). *Normal Approximation and Asymptotic Expansions.* Wiley, New York.

BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviation of density function estimates. *Ann. Statist.* **1** 1071–1095.

BLUM, J. R. and SUSARLA, V. (1980). Maximal deviation theory of density and failure rate function estimates based on censored data. *J. Multivariate Anal.* **5** 213–222.

BURKE, M. D. (1983). Approximations of some hazard rate estimators in a competing risks model. *Stochastic Process. Appl.* **14** 157–174.

BURKE, M. D., CSÖRGŐ, S. and HORVÁTH, L. (1981). Strong approximations of some biometric estimates under random censorship. *Z. Wahrsch. Verw. Gebiete* **56** 87–112.

BURKE, M. D., CSÖRGŐ, S. and HORVÁTH, L. (1988). A correction to and improvement of "Strong approximations of some biometric estimates under random censorship." *Probab. Theory Related Fields* **79** 51–57.

BURKE, M. D. and HORVÁTH, L. (1984). Density and failure rate estimation in a competing risks model. *Sankhyā Ser. A* **46** 135–154.

CSÖRGŐ, M. (1983). *Quantile Processes with Statistical Applications. CBMS-NSF Regional Conference Series in Applied Mathematics* **42**. SIAM, Philadelphia.

CSÖRGŐ, M. and HORVÁTH, L. (1988). Central limit theorems for L_p-norms of density estimators. *Probab. Theory Related Fields* **80** 269–291.

CSÖRGŐ, S. and HORVÁTH, L. (1983). The rate of strong uniform consistency for the product-limit estimator. *Z. Wahrsch. Verw. Gebiete* **62** 411–426.

DEVROYE, L. (1987). *A Course in Density Estimation*. Birkhäuser, Boston.

DEVROYE, L. and GYÖRFI, L. (1985). *Nonparametric Density Estimation: The L_1 View*. Wiley, New York.

DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

GRENANDER, U. (1956). On the theory of mortality measurement II. *Scand. Actuar. J.* **39** 125–153.

GROENEBOOM, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.) **2** 539–555. Wadsworth, Belmont, Calif.

GU, M. G. and LAI, T. L. (1990). Functional laws of the iterated logarithm for the product-limit estimator of a distribution function under random censorship or truncation. *Ann. Probab.* **18** 160–189.

GYÖRFI, L., HÄRDLE, W., SARDA, P. and VIEU, P. (1989). *Nonparametric Curve Estimation from Time Series. Lecture Notes in Statist.* **60**. Springer, Berlin.

KALBFLEISCH, J. G. and PRENTICE, R. L. (1980). *The Statistical Analysis of Failure Time Data*. Wiley, New York.

KAMAT, A. R. (1953). Incomplete and absolute moments of the multivariate normal distribution with some applications. *Biometrika* **40** 20–34.

KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

PARZEN, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York.

ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of density function. *Ann. Math. Statist.* **27** 832–837.

SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.

UPTON, A. C., ALLEN, R. C., BROWN, R. C., CLAPP, N. K., CONLIN, J. W., COSGROVE, G. E., DARDEN, E. B., JR., KASTENBAUM, M. A., O'DELL, P. T., JR., SERRANO, L. J., TYNDALL, R. L. and WALBURG, H. A., JR. (1969). Quantitative experimental study of low-level radiation carcinogenesis. In *Radiation Induced Cancer* 425–438. I.A.E.A., Vienna.

YANDELL, B. S. (1983). Nonparametric inference for rates with censored survival data. *Ann. Statist.* **11** 1119–1135.

MIKLÓS CSÖRGŐ
DEPARTMENT OF MATHEMATICS
AND STATISTICS
CARLETON UNIVERSITY
OTTAWA, ONTARIO
CANADA K1S 5B6

EDIT GOMBAY
DEPARTMENT OF STATISTICS
AND APPLIED PROBABILITY
UNIVERSITY OF ALBERTA
EDMONTON, ALBERTA
CANADA T6G 2G1

LAJOS HORVÁTH
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112