Math-Net.Ru

Общероссийский математический портал

S. S. Ralescu, A Bahadur–Kiefer law for the Nadaraya empiric-quantile processes, *Теория вероятн. и ее примен.*, 1996, том 41, выпуск 2, 380–392

DOI: <https://doi.org/10.4213/tvp2946>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 34.229.45.116

26 апреля 2023 г., 13:26:21

© 1996 г.

RALESCU S. S.*

A BAHADUR-KIEFER LAW FOR
THE NADARAYA EMPIRIC-QUANTILE PROCESSES¹⁾

Изучается поведение п.н. эмпирических квантильных процессов Надарая $\hat{R}_n(\cdot)$. Устанавливается, используя инвариантность и стохастические свойства $\|\hat{R}_n\|$, усиленный закон Бахадура-Кифера для таких процессов, доказывающий робастность относительно класса возмущенных ядерных эмпирических ф.р. Также, при выведении этого результата, получен закон повторного логарифма в форме Штрассена, который обобщает теорему Финкельштейна и, возможно, представляет самостоятельный интерес. Кроме того, приведен краткий обзор приложений.

Ключевые слова и фразы: возмущенные ядерные эмпирические квантильные процессы, закон Бахадура-Кифера, закон повторного логарифма в форме Штрассена.

1. Introduction

Suppose $X_1, X_2, \dots$ are independent identically distributed (i.i.d.) random variables with common distribution function (d.f.) F . A fundamental question of interest in nonparametric statistics is the estimation of F by a random distribution F_n^* based on a sample $X_1, \dots, X_n$ of size n . A substantial proportion of the theory in statistics adopts F_n^* to be the traditional empirical d.f. (e.d.f.) F_n (where $F_n(x) = n^{-1} \# \{1 \leq i \leq n: X_i \leq x\}$) and uses the Kolmogorov test as a successful practical tool when one does not have any prior information about the distribution except that it is continuous. In the case when further prior information is available (such as: the density function exists or the density function is differentiable, etc.), clearly F_n is no longer suitable for F must be estimated by a discontinuous step random function. Indeed, it is well known (see [19]) that in smooth statistical models F_n is inadmissible with respect to integrated square loss as an estimator of F . The common wisdom among proponents of various statistical schools is that whenever one learns to recognize that the data points (after being collected and analyzed) reflect a possible smooth distribution, it is most appropriate

*Department of Mathematics, Queens College of the City University of New York, 65-30 Kissena Boulevard Flushing, NY 11367, USA.

¹⁾Research partially supported by PSC-CUNY, Grant 661374.

that this source of knowledge should be put to efficient use in constructing good estimators. Consequently, we will argue that in such situations it is more natural and desirable to use a smooth random estimator rather than the traditional F_n .

In this paper, for eliminating the undesirable properties of F_n associated with nonsmoothness, we use the data to make judicious choices by working directly with smooth kernel perturbations $\hat{F}_n$ of the e.d.f. defined by

$$\hat{F}_n(x) = n^{-1} \sum_{i=1}^n G\left(\frac{x - X_i}{\lambda_n}\right) = \int G\left(\frac{x - t}{\lambda_n}\right) dF_n(t) \quad (1.1)$$

where the bandwidth $\lambda_n > 0$ tend to zero as $n \rightarrow \infty$ and $G(\cdot)$ is a kernel d.f. associated with a kernel $k = G' \geq 0$, $\int k = 1$. Some examples of kernels generating (1.1) are: (i) $k(x) = \frac{1}{2}$ if $|x| \leq 1$, 0 if $|x| > 1$. This is the uniform kernel corresponding to the «running» histogram density estimator and played an important role in the initial stages of density estimation (see [22]). Even though the uniform is not the best kernel for density estimation, it turns out that it is optimal for kernel estimation of the d.f. or its inverse (see [6]); (ii) $k(x) = 1 - |x|$ if $|x| \leq 1$, 0 if $|x| > 1$. This is the triangular kernel which, although suboptimal, was shown to perform nearly as well as the uniform one in kernel distribution estimation ([6]); (iii) $k(x) = \frac{3}{4}(1 - x^2)$ if $|x| \leq 1$, 0 if $|x| > 1$. This is the Bartlett-Epanechnikov kernel of which the integral is $G(x) = \frac{1}{2} + 3x/4 - x^3/4$ if $|x| \leq 1$, 0 if $x < -1$ and 1 if $x > 1$. It is bounded, compactly supported and Hölder continuous. However, associated with this kernel there are some technical problems due to its non-differentiability at the end-points; (iv) $k(x) = \frac{15}{16}(1 - x^2)^2$ if $|x| \leq 1$, 0 if $|x| > 1$. This biweight kernel (called $K3$ by Tapia and Thompson ([25, p. 60]) is designed as a desirable alternative to the Bartlett-Epanechnikov kernel as it enjoys smoothness properties. It is noteworthy that other possibly more general smoothing methods (such as nearest neighbor smoothing, spline smoothing, etc.) may be considered. Indeed, the broad class of kernel estimators (1.1) is a member of an even larger class of so called perturbed e.d.f.'s which are defined as

$$F_n^*(x) = n^{-1} \sum_{i=1}^n G_n(x - X_i),$$

where $\{G_n\}$ is a sequence of smooth d.f.'s converging weakly to the distribution $u(x) = 1$ if $x \geq 0$, 0 if $x < 0$. General perturbed e.d.f.'s are studied by Winter [27], and Nixdorf [15] but mainly for reasons of analytical tractability and simplicity we concentrate only on the easy-to-use kernel based method. It should be understood that the resulting estimator $\hat{F}_n$ given by (1.1) has a more complicated structure than that of F_n and carrying out the analysis of the proposed kernel modification is a delicate and subtle affair and no routine analysis is possible.

The kernel-type estimator (1.1) was introduced by Nadaraya [14] who also considered the corresponding smooth quantile estimator $\hat{F}_n^{-1}(t) = \inf\{x: \hat{F}_n(x) \geq t\}$, $0 < t < 1$. Various large sample results have been obtained by several authors including Winter [26], [27], Yamato [28], Azzalini [1], Reiss [20], [21], Falk [4], Puri and Ralescu [16], Mack [13], Jones [6], Lio and Padgett [12], Ralescu [17] and Ralescu and Sun [18].

The theme that motivates this study and makes it interesting is connected to the question: when do the smooth estimators $\hat{F}_n$ (and $\hat{F}_n^{-1}$) differ substantially or improve substantially on the usual e.d.f. and empirical quantile estimator? In identifying the key differences and in order to stress the usefulness of processes based on perturbed kernel estimators, we present the following statistically supported arguments:

(i) (Asymptotic deficiency.) When the theoretical performance of $\hat{F}_n$ is addressed, it turns out that the Nadaraya perturbed kernel estimator behaves rather better than the ordinary F_n in terms of asymptotic relative deficiency for appropriately chosen kernels and sufficiently smooth F 's (see [20], [4]). Likewise, perturbed kernel quantiles of a data set which arise chiefly in certain problems connected with preliminary medical diagnosis as natural competitors to central order statistics have the tendency to outperform their counterparts in terms of deficiency, at least asymptotically. Indeed, several authors (see e.g. [1], [7]) have recognized the need to implement $\hat{F}_n^{-1}$ and their extensive simulation studies revealed success in applying $\hat{F}_n^{-1}$ to estimate F^{-1} in models governed by moderate to heavy tailed smooth distributions (in [21] displays examples which show considerable practical potential for this method).

(ii) (Bootstrap.) Recently, Falk and Reiss have considered a bootstrap estimator based on $\hat{F}_n$ and showed that under appropriate regularity conditions it has a better performance than the corresponding bootstrap estimator based on F_n . More specifically, in studying the weak convergence of a smooth bootstrap quantile estimator, they noted that the striking advantage in smoothing F_n before bootstrapping is that the rate of convergence of the resulting estimator is, roughly speaking of order $O(n^{-2/5})$ (for appropriate choices of λ_n) which is considerably better than in the nonsmooth case.

Much effort in probability and statistics has been expended in proving limit theorems for uniform empirical and quantile processes defined in terms of F_n . A pivotal role in the theory and applications is played by the classical Bahadur-Kiefer (BK) process defined by

$$D_n(t) = \sqrt{n}(F_n(t) - t) + \sqrt{n}(F_n^{-1}(t) - t), \quad 0 \leq t \leq 1,$$

where F_n corresponds to a uniform $[0, 1]$ i.i.d. sequence and $F_n^{-1}(t) = \inf\{x: F_n(x) \geq t\}$. The asymptotic study of the process D_n has represented a major focus of interest ever since the pioneering papers of Bahadur [2] and Kiefer [8], [9]. In [3] one can find many results and references on the subject. Here

we select to state the fundamental result of Kiefer [9] according to which

$$\limsup_{n \rightarrow \infty} n^{1/4} (\log n)^{-1/2} (\log \log n)^{-1/4} \|D_n\| = 2^{-1/4} \quad \text{a.s.} \quad (1.2)$$

where $\|D_n\| = \sup_{0 \leq t \leq 1} |D_n(t)|$.

With a desire to robustify the Bahadur-Kiefer theory and in view of the noted advantages of $\hat{F}_n$ over F_n in smooth statistical models, we ask *if the same rate of convergence is achievable in the estimation of F by $\hat{F}_n$ and the consideration of the Nadaraya smooth empiric-quantile process $\hat{R}_n$ defined by*

$$\begin{aligned} \hat{R}_n(t) = & \sqrt{n} \left[\hat{F}_n(F^{-1}(t)) - t \right] + \sqrt{n} f(F^{-1}(t)) \\ & \times [\hat{F}_n^{-1}(t) - F^{-1}(t)], \quad 0 < t < 1. \end{aligned} \quad (1.3)$$

This question is of interest since when applying asymptotic theory, one wants a certain stability: that is, the asymptotic result should be preserved when the e.d.f. is only slightly perturbed. In the present work a complete affirmative answer to this question is provided. One of the main objectives here is to gain theoretical insight into the asymptotic behavior of the BK-type process $\hat{R}_n$ and complement the recent result of Ralescu [17] who obtained the order embodied in (1.2) for the a.s. convergence of $\|\hat{R}_n\|$. An important point to be made is that one may be faced with a nonrobust situation, so that a choice must be carefully made with respect to the appropriate assumptions on λ_n . For purposes of applications, a possibly systematic way of choosing the parameter λ_n which determines the degree of smoothing may be based on cross-validation techniques.

The principal contributions of this paper are two-fold. (1) to derive, as a tool for arriving at the main result the relative compactness of smoothed empirical processes (in the sense of Strassen's law of the iterated logarithm), thus extending the well-known Finklestein result [5] to the broad class of perturbed kernel e.d.f.'s; and (2) to provide the exact limit theorem controlling the almost sure convergence of $\|\hat{R}_n\|$ thus showing that the result encapsulated in Theorem 2.2 of [17] is sharp under the stated restrictions on f and λ_n . Our highly desirable large sample result complements and extends the existing asymptotic properties of perturbed kernel estimators and provides a BK-type analogue of the corresponding result (1.2). It is hoped that variants of our arguments may work successfully with other smooth estimators.

The paper is organized as follows. In Section 2 we obtain the relative compactness of the Nadaraya perturbed kernel empirical process which opens the way to carry out the asymptotic analysis of $\hat{R}_n$. After this preliminary work, the main result of this paper, the Bahadur-Kiefer type law for the process $\hat{R}_n$ is presented as Theorem 3.1 in Section 3. Finally, in Section 4 we concentrate on potential applications of the main result which, for the sake of brevity are outlined without proof. It should be noted that our

theorems are new and go beyond the asymptotic results of related papers which are listed in the references.

2. Limit points of perturbed kernel empirical process

We write the perturbed kernel empirical process as

$$\hat{Z}_n(x) = n^{1/2} \{ \hat{F}_n(x) - F(x) \}, \quad -\infty < x < \infty,$$

and also set

$$\tilde{Z}_n(x) = n^{1/2} \{ \hat{F}_n(x) - E\hat{F}_n(x) \}, \quad -\infty < x < \infty.$$

Throughout the rest of the paper, we shall let $\theta_n = (2 \log \log n)^{1/2}$, $n \geq 3$.

To handle the Bahadur-Kiefer result for $\hat{F}_n$ we need to address the issue of relative compactness (in the sense of Strassen's LIL) for $\hat{Z}_n$. The celebrated Strassen law asserts that if $\{W(t), t \geq 0\}$ is a standard Wiener process, then the sequence $\{W(nx)/\theta_n, 0 \leq x \leq 1\}$ is almost surely (a.s.) relatively compact in $C[0, 1]$ (= the space of continuous real-valued functions on $[0, 1]$ with the supremum norm) and the set of its limit points equals L , where L consists of those functions h in $C[0, 1]$ which can be written in the Lebesgue-integral form

$$h(t) = \int_0^t g(s) ds, \quad \text{where} \quad \int_0^1 g^2(s) ds \leq 1.$$

Inspired by Strassen's work, Finkelstein [5] showed that if F is a continuous d.f. and

$$Z_n(x) = n^{1/2} \{ F_n(x) - F(x) \}, \quad -\infty < x < \infty,$$

denotes the standard empirical process, then the sequence $\{Z_n(F^{-1}(t))/\theta_n, 0 \leq t \leq 1\}$ is a.s. relatively compact in the space $D[0, 1]$ of all real-valued right-continuous functions on $[0, 1]$ which have a finite limit from the left, with limit set $H = L \cap \{h: h(1) = 0\}$. The similarity of these results hints that versions of the above LIL may exist in the case of the smooth empirical process $\hat{Z}_n$. Indeed, as a prelude of deriving (3.3) we shall show that $\hat{Z}_n$ satisfies a Strassen-type LIL by combining the following steps which give the flavor of our approach:

1) If $\{K(x, y), 0 \leq x \leq 1, 0 \leq y\}$ denotes a Kiefer process, it follows that the sequence $\{n^{-1/2} K(F(\cdot), n)/\theta_n\}$ is a.s. relatively compact in the space $B(-\infty, \infty)$ of bounded real-valued functions on $(-\infty, \infty)$ endowed with the topology induced by the sup-norm $\|\phi\| = \sup_{-\infty < x < \infty} |\phi(x)|$.

2) On account of Stute's theorem [24] on global oscillations of Z_n , an application of the strong approximation result of Komlós, Major and Tusnády (KMT) [11] will produce the estimate

$$\|\tilde{Z}_n(\cdot) - n^{-1/2} K(F(\cdot), n)\| = O(n^{-1/2} \log^2 n + \varepsilon_n) \quad \text{a.s.}$$

for some $\varepsilon_n \rightarrow 0$ as $n \rightarrow \infty$.

The following conditions which we will use are not very restrictive, requiring appropriate rates for the bandwidth and some regularity of the kernel.

Condition 2.1. $\lambda_n = O(a_n)$ where the sequence $\{a_n\}$ satisfies

- (i) $a_n \downarrow 0$, $na_n \uparrow \infty$, $\log(1/a_n)/(na_n) \rightarrow 0$,
- (ii) $\log(1/a_n)/(\log \log n) \rightarrow \infty$,
- (iii) $(na_n^4)/(\log \log n) \rightarrow 0$.

Condition 2.2.

- (i) k is compactly supported with support $\subset [-c, c]$, $c > 0$,
- (ii) $\int xk(x)dx = 0$ and $\int x^2k(x)dx < \infty$.

We note that the above assumptions are not mutually exclusive and we emphasize that only subsets of them will be used at different stages in the sequel. Assumptions (i) and (ii) of Condition 2.1 correspond to the Stute conditions [24] which are needed to ensure the proper order of a.s. convergence for the modulus of continuity of Z_n while assumption (iii) is required to handle the bias term $n^{1/2}\{\mathbf{E}\hat{F}_n(\cdot) - F(\cdot)\}$. It is clear that (iii) implies (ii). Assumptions (i) and (ii) of Condition 2.2 are the «usual» ones encountered in kernel estimation and they are both satisfied by the examples given in the Introduction.

In the following theorem we describe the almost sure behavior of the Nadaraya perturbed kernel empirical processes $\tilde{Z}_n$ and $\hat{Z}_n$.

Theorem 2.1.

(a) Assume that F is differentiable on its support and that k satisfies (i) of Condition 2.2. Then, under (i) and (ii) of Condition 2.1, the sequence $\{\tilde{Z}_n(\cdot)/\theta_n\}$ is a.s. relatively compact in $B(-\infty, \infty)$ with limit set equal to $H_F = \{h(F), h \in H\}$.

(b) If in addition F is twice differentiable on its support with bounded F'' , (i) and (iii) of Condition 2.1 hold and k satisfies also (ii) of Condition 2.2, then $\{\hat{Z}_n(\cdot)/\theta_n\}$ is a.s. relatively compact in $B(-\infty, \infty)$ with limit set equal to H_F .

In order to prove the theorem, it is of interest to investigate how well $\tilde{Z}_n$ can be approximated by appropriate Gaussian processes. According to the KMT [11] strong approximation result, without loss of generality we can assume that we are on a suitable probability space where sits a two-parameter Wiener process $\{W(x, y), x \geq 0, y \geq 0\}$ such that if we set $K(x, y) = W(x, y) - xW(1, y)$ for $0 \leq x \leq 1$ and $y \geq 0$, we have almost surely as $n \rightarrow \infty$

$$\|Z_n(\cdot) - n^{-1/2}K(F(\cdot), n)\| = O(n^{-1/2} \log^2 n). \quad (2.1)$$

We now proceed to approximate $\tilde{Z}_n$ in the following lemma.

Lemma 2.1. *Under the conditions of part (a) of Theorem 2.1 we have*

$$\|\tilde{Z}_n - S_n\| = O\left(\left(a_n \log \frac{1}{a_n}\right)^{1/2} + n^{-1/2} \log^2 n\right) \quad \text{a.s.}, \quad (2.2)$$

where $S_n(\cdot) = n^{-1/2} K(F(\cdot), n)$.

P r o o f. In view of (2.1) it is natural to couple $\tilde{Z}_n$ with Z_n and estimate $\|\tilde{Z}_n - Z_n\|$. This expression can be bounded as follows. By assumption, there exists a constant $d > 0$ such that $\lambda_n \leq da_n$ for all $n \geq 1$. Set $a^* = cda_n$. Then, observing that

$$\tilde{Z}_n(x) = \int_{-\infty}^{\infty} Z_n(x-t) dG_n(t),$$

where

$$G_n(x) = \int_{-\infty}^x \lambda_n^{-1} k(\lambda_n^{-1} t) dt,$$

we have with $\Gamma_n(x, t) = |Z_n(x-t) - Z_n(x)|$

$$\begin{aligned} \|\tilde{Z}_n - Z_n\| &\leq \left\| \int_{|t| \leq a_n^*} \Gamma_n(\cdot, t) dG_n(t) \right\| \\ &\quad + \left\| \int_{|t| > a_n^*} \Gamma_n(\cdot, t) dG_n(t) \right\| = \alpha_n + r_n. \end{aligned} \quad (2.3)$$

First, notice that uniformly in x

$$\int_{|t| > a_n^*} \Gamma_n(x, t) dG_n(t) \leq 2 \|Z_n\| \int_{|s| > \lambda_n^{-1} a_n^*} k(s) ds \quad (2.4)$$

and since $\lambda_n^{-1} a_n^* \geq c$ and the kernel k has support confined in $[-c, c]$ we see from (2.4) that $r_n = 0$.

To handle α_n , since $\{a_n^*\}$ obviously satisfies (i) and (ii) of Condition 2.1, on account of inequality (2.13) of [24] we have almost surely

$$\alpha_n \leq \left\| \sup_{|t| \leq a_n^*} \Gamma_n(\cdot, t) \right\| = O\left(\left(a_n^* \log \frac{1}{a_n^*}\right)^{1/2}\right) = O\left(\left(a_n \log \frac{1}{a_n}\right)^{1/2}\right). \quad (2.5)$$

The above estimates in combination with (2.1) yield the assertion (2.2).

P r o o f o f T h e o r e m 2.1. (a) First we notice that

$$\|\theta_n^{-1} S_n(\cdot) - (h \circ F)(\cdot)\| \leq \|(n^{1/2} \theta_n)^{-1} K(\cdot, n) - h(\cdot)\| \quad (2.6)$$

holds everywhere for each $h \in H$. Now (2.6) and the fact that the sequence $\{(n^{1/2} \theta_n)^{-1} K(\cdot, n)\}$ is a.s. relatively compact in $C[0, 1]$ with limit set equal to H (cf. [3, Theorem 1.15.1]) jointly imply that $\{\theta_n^{-1} S_n(\cdot)\}$ is a.s. relatively compact in $B(-\infty, \infty)$ with limit set equal to H_F . This together with an application of Lemma 2.1 entails the validity of statement (a) since

the conditions imposed on $\{a_n\}$ imply that $a_n \log(1/a_n)/\log \log n \rightarrow 0$ as $n \rightarrow \infty$.

For (b) it suffices to note that under the smoothness of F and the stated assumptions on k we have (cf. [18])

$$\|\mathbf{E}\hat{F}_n(\cdot) - F(\cdot)\| = O(\lambda_n^2) \quad \text{as } n \rightarrow \infty. \quad (2.7)$$

Combining the bias estimate with (2.2) and in view of (iii) of Condition 2.1, the conclusion of (b) follows. This completes the proof of Theorem 2.1.

Remark 1. Condition 2.1 (iii) is a small price one pays to control the bias term above in order to derive the relative compactness of $\hat{Z}_n$. For example, we may let $\lambda_n = O(a_n)$ and allow the choice $a_n = n^{-\alpha} \log n$ with $0 < \alpha < 1$. In this case (i) and (ii) of Condition 2.1 are clearly satisfied and if in addition $\alpha > \frac{1}{4}$ then (iii) also holds entailing the Strassen-type LIL for $\hat{Z}_n$.

Remark 2. From Theorem 2.1 a whole variety of results on functionals of $\hat{Z}_n$ may be quickly derived and we state

Corollary 2.1. *Under the assumptions of part (b) of Theorem 2.1 we have almost surely*

$$\begin{aligned} \text{(i)} \quad & \limsup_{n \rightarrow \infty} \theta_n^{-1} \|\hat{Z}_n\| = 2^{-1} \quad (\text{Chung-Smirnov property}), \\ \text{(ii)} \quad & \limsup_{n \rightarrow \infty} n\theta_n^{-2} \int_{-\infty}^{\infty} (\hat{F}_n - F)^2 dF = \pi^{-2}. \end{aligned} \quad (2.8)$$

3. The main result

In this section we are going to present our main theorem which gives the exact rate of a.s. convergence of $\|\hat{R}_n\|$. In order to do so, we need to assume somewhat stronger regularity conditions on F and $\{\lambda_n\}$.

Condition 3.1. *The distribution F has a density f which is positive on its support, has a bounded derivative f' and*

$$\inf_t f(F^{-1}(t)) = A > 0. \quad (3.1)$$

Condition 3.2. $\lambda_n = O(a_n)$ for some sequence which satisfies (i) of Condition 2.1 and

$$a_n \log \frac{1}{a_n} = o((n^{-1} \log \log n)^{1/2} \log n) \quad \text{as } n \rightarrow \infty. \quad (3.2)$$

Relation (3.1) of Condition 3.1 is needed to prevent $\hat{F}_n^{-1}(t)$ from being too much smaller than t while (3.2) amounts to assuming a smaller upper range of variation for $\{a_n\}$ as compared to (iii) of Condition 2.1. Note that (3.2) is assuredly satisfied if $a_n = n^{-\alpha} \log n$ with $\frac{1}{2} < \alpha < 1$.

Our main result is now formulated in the next theorem which gives the BK law.

Theorem 3.1. *Under Conditions 3.1 and 3.2 we have almost surely*

$$\limsup_{n \rightarrow \infty} \gamma_n \|\hat{R}_n\| = 2^{-1/2}, \quad (3.3)$$

where $\gamma_n = (n^{-1/2} \theta_n \log n)^{-1/2}$.

P r o o f. For notational convenience in the sequel we introduce the following kernel quantile processes

$$Q_n^*(t) = n^{1/2} \left[F(\hat{F}_n^{-1}(t)) - t \right], \quad 0 < t < 1,$$

and

$$\hat{Q}_n(t) = n^{1/2} f(F^{-1}(t)) [\hat{F}_n^{-1}(t) - F^{-1}(t)], \quad 0 < t < 1.$$

The proof of (3.3) may be better understood if it is decomposed into two parts as follows.

(i) **O b t a i n i n g t h e u p p e r b o u n d.** By using (i) of (2.8) and $\|Q_n^*\| \leq \|\hat{Z}_n\|$ we have

$$\limsup_{n \rightarrow \infty} \theta_n^{-1} \|Q_n^*\| \leq 2^{-1} \quad \text{a.s.} \quad (3.4)$$

To establish the upper bound in (3.3) we require the following statement.

Lemma 3.1. *There exists a constant $C > 0$ such that under Condition 3.1, we have*

$$\|\hat{Q}_n - Q_n^*\| \leq C n^{-1/2} \|Q_n^*\|^2. \quad (3.5)$$

P r o o f. In view of Condition 3.1, F is strictly increasing on its support and we can write

$$\hat{Q}_n(t) = Q_n^*(t) \frac{\sqrt{n} [F^{-1}(F(\hat{F}_n^{-1}(t))) - F^{-1}(t)]}{\sqrt{n} [F(\hat{F}_n^{-1}(t)) - t]} g(t),$$

where $g = f \circ F^{-1}$. Therefore, a one-term Taylor expansion yields

$$\hat{Q}_n(t) - Q_n^*(t) = Q_n^*(t) \left[\frac{g(t)}{g(\xi_n^*(t))} - 1 \right], \quad (3.6)$$

where $\xi_n^* = \xi_n(t)$ is a random quantity such that

$$|\xi_n^* - t| \leq |F(\hat{F}_n^{-1}(t)) - t|. \quad (3.7)$$

Concerning the bracketed term on the right-side of (3.6), by another application of Taylor's formula we have for some $\rho (= \rho_n(t))$ between t and ξ_n^*

$$|\varepsilon_n(t)| = \left| \frac{g(t)}{g(\xi_n^*)} - 1 \right| \leq A^{-1} |\xi_n^* - t| \cdot |g'(\rho)|. \quad (3.8)$$

Hence, since $|g'(s)| = |f'(F^{-1}(s))/g(s)| \leq A^{-1} \|f'\|$, putting together (3.6) through (3.8) we get

$$\|\hat{Q}_n - Q_n^*\| \leq \|Q_n^*\| \cdot \|\varepsilon_n\| \leq (A^{-2} \|f'\|) n^{-1/2} \|Q_n^*\|^2.$$

This proves (3.5) with $C = A^{-2} \|f'\|$.

Continuing to focus on the upper bound in (3.3) we note that $\hat{R}_n(t) = \hat{Z}_n(F^{-1}(t)) + \hat{Q}_n(t)$. Then, in view of Lemma 2.1 and (2.7) we can write uniformly in t

$$\begin{aligned} \gamma_n \hat{R}_n(t) &= \gamma_n [B_n(t) - B_n(F(\hat{F}_n^{-1}(t)))] \\ &\quad + \gamma_n [\hat{Q}_n(t) - Q_n^*(t)] + O(\gamma_n \delta_n) \quad \text{a.s.}, \end{aligned} \quad (3.9)$$

where

$$B_n(t) = n^{-1/2} K(t, n) \quad \text{and} \quad \delta_n = \left(a_n \log \frac{1}{a_n} \right)^{1/2} + n^{1/2} \lambda_n^2 + n^{-1/2} \log^2 n.$$

From (3.4), there exists Ω_0 with $P(\Omega_0) = 1$ such that for each $\omega_0 \in \Omega_0$ and every $\eta > 1$, for all $n \geq n_0 (= n(\omega_0, \eta))$ we have uniformly in t

$$|F(\hat{F}_n^{-1}(t)) - t| \leq b_n, \quad (3.10)$$

where $b_n = \frac{1}{2} \eta \theta_n n^{-1/2}$.

Now, the following property of the modulus of continuity of the Brownian bridges generated by the Kiefer process is decisive and can be found in [3]:

$$\lim_{n \rightarrow \infty} \delta_n \sup_{t-s \leq h_n} |B_n(t) - B_n(s)| = 1 \quad \text{a.s.}, \quad (3.11)$$

where $\delta_n = (2h_n \log h_n^{-1})^{-1/2}$ and h_n is a sequence of positive numbers satisfying $h_n \downarrow 0$, $nh_n \uparrow$ and $(\log h_n^{-1})(\log \log n)^{-1} \rightarrow \infty$ as $n \rightarrow \infty$.

On letting $h_n = b_n$ we can assume without loss of generality that (3.10) and (3.11) hold on Ω_0 uniformly in t . Then, tying the estimates (3.4), (3.5) and (3.9)–(3.11) together, since $\gamma_n \delta_n \rightarrow 0$ as $n \rightarrow \infty$ it follows for ω_0 that

$$\limsup_{n \rightarrow \infty} \gamma_n \|\hat{R}_n\| \leq \limsup_{n \rightarrow \infty} \frac{\sup_{t-s \leq b_n} |B_n(t) - B_n(s)|}{(2b_n \log b_n^{-1})^{1/2}} \cdot \frac{\sqrt{\eta/2}}{\omega_n} \leq \sqrt{\frac{\eta}{2}}, \quad (3.12)$$

where $\omega_n = (2 \log b_n^{-1} / \log n)^{-1/2} \rightarrow 1$ as $n \rightarrow \infty$. By letting η converge down to 1, we see that we have proven almost surely

$$\limsup_{n \rightarrow \infty} \gamma_n \|\hat{R}_n\| \leq 2^{-1/2}. \quad (3.13)$$

(ii) Obtaining the lower bound. Proceeding as in the derivation of (3.13) it can be shown that

$$\limsup_{n \rightarrow \infty} \gamma_n \|R_n^*\| \leq 2^{-1/2} \quad \text{a.s.}, \quad (3.14)$$

where $R_n^*(t) = \hat{Z}_n(F^{-1}(t)) + Q_n^*(t) = \hat{Z}_n(F^{-1}(t)) - \hat{Z}_n(\hat{F}_n^{-1}(t))$.

Now, by an application of (b) of Theorem 2.1 and (3.14) it follows that the sequence $\{Q_n^*(\cdot)/\theta_n\}$ is a.s. relatively compact with limit set equal

to H . In order to pass to a more convenient asymptotic representation of $\hat{R}_n$, choose any $0 < \varepsilon < \frac{1}{4}$ and set $d_n = (\frac{1}{2} - 2\varepsilon)n^{-1/2}\theta_n$. Then, from (3.4), (3.5), and (3.9) we have almost surely uniformly in t

$$\gamma_n \hat{R}_n(t) = \frac{[B_n(t) - B_n(F(\hat{F}_n^{-1}(t)))]}{(2d_n \log d_n^{-1})^{1/2}} \cdot \frac{\sqrt{\frac{1}{2} - 2\varepsilon}}{\tau_n} + o(1), \quad (3.15)$$

where $\tau_n = (2 \log d_n^{-1} / \log n)^{-1/2} \rightarrow 1$ as $n \rightarrow \infty$.

If we let $h_\varepsilon \in H$ denote the Shorack function [23, p. 370], from the relative compactness of $\{Q_n^*(\cdot)/\theta_n\}$ it follows that for almost every ω we can find a subsequence $m = m(\omega, h_\varepsilon)$ such that

$$\varphi_m(t) = \theta_m^{-1} Q_m^*(t) - h_\varepsilon(t), \quad 0 < t < 1,$$

satisfies $\|\varphi_m\| \rightarrow 0$ as $m \rightarrow \infty$ along this subsequence. Then, since

$$F(\hat{F}_m^{-1}(t)) = t + m^{-1/2} \theta_m [h_\varepsilon(t) + \varphi_m(t)]$$

it can be shown that

$$\limsup_{n \rightarrow \infty} \frac{\sup_{0 < t < 1} [B_n(t) - B_n(F(\hat{F}_n^{-1}(t)))]}{(2d_n \log d_n^{-1})^{1/2}} \geq 1 \quad \text{a.s.}$$

This in combination with (3.15) implies by the arbitrary choice of $0 < \varepsilon < \frac{1}{4}$ that almost surely

$$\limsup_{n \rightarrow \infty} \gamma_n \|\hat{R}_n\| \geq 2^{-1/2} \quad (3.16)$$

By (3.13) and (3.16) we have the theorem.

4. A brief profile of applications

In this section we will present the main points of potential applications of Theorem 3.1 by briefly explaining and discussing some useful asymptotic results which follow from (3.3).

(i) **Strassen-type LIL for perturbed kernel quantiles.** In deriving the set of limit points of the sequence $\{\hat{Q}_n(\cdot)/\theta_n\}$, Theorem 3.1 is a key result. Indeed, a combination of (3.3) with (b) of Theorem 2.2 (or the estimate (3.5) with the relative compactness of $\{Q_n^*(\cdot)/\theta_n\}$) immediately yields that $\{\hat{Q}_n(\cdot)/\theta_n\}$ is a.s. relatively compact with limit set H .

(ii) **Upper and lower class results.** Here we describe the domain of usefulness of our main result in studying the characterization of upper and lower class functions for the sequence $\|\hat{Q}_n\|$. Instead of working directly with (3.3), a helpful weaker variant will now be discussed. In Theorem 3.1 we insisted to obtain the exact limsup conclusion for $\|\hat{R}_n\|$ and this led to imposing the restriction (3.2) on the growth of $\{\lambda_n\}$. In many applications (including the one presented her) a weaker a.s. rate suffices and this allows a wider range of variability for the bandsequence $\{\lambda_n\}$. Indeed, introducing the process $\tilde{R}_n(t) = Z_n(F^{-1}(t)) + \hat{Q}_n(t)$, $0 < t < 1$, and by imposing

appropriate restrictions (i.e., assuming the conditions of Theorem 3.1 where (3.2) is exchanged for the weaker $n^{1/4}a_n(\log \log n)^{-1} \rightarrow 0$ as $n \rightarrow \infty$) one can obtain with $\nu_n = n^{-1/8}(\log \log n)^{1/2} + n^{1/2}\lambda_n^2$ that there exists a finite positive constant c^* such that

$$\limsup_{n \rightarrow \infty} \nu_n^{-1} \|\tilde{R}_n\| \leq c^* \quad \text{a.s.} \quad (4.1)$$

The details of showing this claim are omitted, the proof (based on modifications of techniques used in the previous section) being somewhat similar to that of (3.13). The advantage of this asymptotic variant is that it allows us to proceed by a strong invariance argument and transfer asymptotic laws from Z_n to $\hat{Q}_n$. In particular, if $a_n = o((n \log \log n)^{-1/4})$ as $n \rightarrow \infty$ and the other condition of Theorem 3.1 are satisfied, we can characterize $P\{\|\hat{Q}_n\| > \psi_n \text{ infinitely often}\}$ as being equal to 0 or 1 according to

$$\sum_{n=1}^{\infty} n^{-1} \psi_n^2 \exp(-2\psi_n^2) < \infty \quad \text{or} \quad = \infty,$$

provided the numerical sequence $\{\psi_n\}$ satisfies $\psi_n \uparrow \infty$.

(iii) **A LIL for $\hat{Q}_n(t)$ when $t = t_n \downarrow 0$.** With the aid of (4.1) we can produce a nice application in the asymptotic theory of the lower tail of the process $\{\hat{Q}_n(t), 0 < t < 1\}$. Indeed, when focusing on a sequence $t = t_n \downarrow 0$ in such a way that $nt_n(\log \log n)^{-1} \rightarrow \infty$ as $n \rightarrow \infty$, statement (4.1) turns out to be a valuable tool in establishing a LIL for $\hat{Q}_n(t_n)$ thus extending the work of Kiefer [10]. It should be noted, however, that the conditions imposed on $\{\lambda_n\}$ and F are somewhat restrictive, leaving certain statistically interesting cases uncovered and for this reason a more direct approach of such a result including the problem of finding the order of $\tilde{R}_n(t_n)$ is needed. The detailed treatment of these problems will be pursued elsewhere.

(iv) **Approximations by Brownian bridges and a Kiefer process.** It is also clear that as a by-product of our techniques and in the spirit of the work of [3] one may produce strong approximations of $\hat{Z}_n$ and $\hat{Q}_n$ by appropriate sequences of Brownian bridges and a Kiefer process via the KMT result.

REFERENCES

1. Azzalini A. A note on the estimation of a distribution function and quantiles by a kernel method. — *Biometrika*, 1981, v. 68, p. 326–328.
2. Bahadur R. R. A note on quantiles in large samples. — *Ann. Math. Statist.*, 1966, v. 37, p. 577–580.
3. Csörgő M., Révész P. Strong Approximations in Probability and Statistics. — New York: Academic Press, 1981.
4. Falk M. Relative efficiency and deficiency of kernel type estimators of smooth distribution functions. — *Statist. Neerlandica*, 1983, v. 37, p. 73–83.

5. *Finkelstein H.* The law of the iterated logarithm for empirical distributions. — *Ann. Math. Statist.*, 1971, v. 42, p. 607–615.
6. *Jones M. C.* The performance of kernel density functions in kernel distribution function estimation. — *Statist. and Probab. Letters*, 1990, v. 9, p. 129–132.
7. *Kappenman R. F.* Improved distribution quantile estimation. — *Commun. Statist.-Simula.*, 1987, v. 16, p. 307–320.
8. *Kiefer J.* On Bahadur's representation of sample quantiles. — *Ann. Math. Statist.*, 1967, v. 38, p. 1323–1342.
9. *Kiefer J.* Deviations between the sample quantile process and the sample d.f. — In: *Nonparametric Techniques in Statistical Inference*/Ed. by M. L. Puri, London: Cambridge Univ. Press, 1970, p. 299–319.
10. *Kiefer J.* Iterated logarithm analogues for sample quantiles when $p_n \rightarrow 0$. — *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 1972, Vol 1, Berkeley: Univ. of California Press, p. 227–244.
11. *Komlós J., Major P., Tusnády G.* An approximation of partial sums of independent r.v.'s and the sample d.f.I. — *Z. Wahrsch. verw. Geb.*, 1975, B. 32, S. 111–131.
12. *Lio Y. L., Padgett W. J.* A note on the asymptotically optimal bandwidth for Nadaraya's quantile estimator. — *Statist. and Probab. Letters*, 1991, v. 11, p. 243–249.
13. *Mack Y. P.* Bahadur's representation of sample quantiles based on smoothed estimates of a distribution function. — *Probab. Math. Statist.*, 1987, v. 8, p. 183–189.
14. *Nadaraya E. A.* Some new estimates for distribution function. — *Theory Probab. Appl.*, 1964, v. 9, p. 497–502.
15. *Nizdorf R.* Central limit theorems in $C[0, 1]$ for a class of estimators of a distribution function. — *Statist. Neerlandica*, 1985, v. 39, p. 251–260.
16. *Puri M. L., Ralescu S. S.* Central limit theorem for perturbed empirical distribution functions evaluated at a random point. — *J. Multivar. Anal.*, 1986, v. 19, p. 273–279.
17. *Ralescu S. S.* Asymptotic deviations between perturbed empirical and quantile processes. — *J. Statist. Plann. Inference*, 1992, v. 32, p. 243–258.
18. *Ralescu S. S., Sun S.* Necessary and sufficient conditions for the asymptotic normality of perturbed sample quantiles. — *J. Statist. Plann. Inference*, 1993, v. 35, p. 55–64.
19. *Read R. R.* The asymptotic inadmissibility of the sample distribution function. — *Ann. Math. Statist.*, 1972, v. 43, p. 89–95.
20. *Reiss R. D.* Nonparametric estimation of smooth distribution functions. — *Scand. J. Statist.*, 1981, v. 8, p. 116–119.
21. *Reiss R. D.* *Approximate Distributions of Order Statistics*. — New York: Wiley, 1989.
22. *Rosenblatt M.* Remarks on some nonparametric estimation of a density function. — *Ann. Math. Statist.*, 1956, v. 27, p. 832–837.
23. *Shorack G.* Kiefer's theorem via the Hungarian construction. — *Z. Wahrsch. verw. Geb.*, 1982, B. 61, S. 369–373.
24. *Stute W.* The oscillation behavior of empirical processes. — *Ann. Probab.*, 1982, v. 10, p. 86–107.
25. *Tapia R. A., Thompson J. R.* *Nonparametric Probability Density Estimation*. — Baltimore: John Hopkins Univ. Press, 1978.
26. *Winter B. B.* Strong uniform consistency of integrals of density estimators. — *Canad. J. Statist.*, 1973, v. 1, p. 247–250.
27. *Winter B. B.* Convergence rates of perturbed empirical distribution functions. — *Appl. Probab.*, 1979, v. 16, p. 163–173.
28. *Yamato H.* Uniform convergence of an estimator of a distribution function. — *Bull. Math. Statist.*, 1973, v. 15, p. 69–78.

Поступила в редакцию
8.VI.1993