ON CONSISTENCY OF KERNEL DENSITY ESTIMATORS FOR RANDOMLY CENSORED DATA: RATES HOLDING UNIFORMLY OVER ADAPTIVE INTERVALS

Evarist GINÉ\textsuperscript{a,1}, Armelle GUILLOU\textsuperscript{b}

\textsuperscript{a}Departments of Mathematics and Statistics, University of Connecticut, Storrs, CT 06269, USA
\textsuperscript{b}Université Paris VI, L.S.T.A., Tour 45-55 E3, 4 Place Jussieu, 75252 Paris Cedex 05, France

Received 17 November 1999, revised 12 January 2001

ABSTRACT. – In the usual right-censored data situation, let $f_n, n \in \mathbb{N}$, denote the convolution of the Kaplan–Meier product limit estimator with the kernels $a_n^{-1} K(\cdot/a_n)$, where $K$ is a smooth probability density with bounded support and $a_n \to 0$. That is, $f_n$ is the usual kernel density estimator based on Kaplan–Meier. Let $\bar{f}_n$ denote the convolution of the distribution of the uncensored data, which is assumed to have a bounded density, with the same kernels. For each $n$, let $J_n$ denote the half line with right end point $Z_n(1-\varepsilon_n)/a_n - a_n$, where $\varepsilon_n \to 0$ and, for each $m$, $Z_{m,n}$ is the $m$th order statistic of the censored data. It is shown that, under some mild conditions on $a_n$ and $\varepsilon_n$, $\sup_{J_n} |f_n(t) - \bar{f}_n(t)|$ converges a.s. to zero as $n \to \infty$ at least as fast as $\sqrt{\log(a_n \wedge \varepsilon_n)}/(na_n\varepsilon_n)$. For $\varepsilon_n$ constant, this rate compares, up to constants, with the exact rate for fixed intervals. © 2001 Éditions scientifiques et médicales Elsevier SAS

AMS classification: Primary 62G07; 62G20; 60F15, Secondary 62G30

Keywords: Censored data; Uniform almost sure rates; Product limit estimator; Non-parametric density estimation; Kernel density estimators; Exponential inequalities

RÉSUMÉ. – Dans le cas des données censurées à droite, on note $f_n, n \in \mathbb{N}$, la convolution de l’estimateur de Kaplan–Meier avec les noyaux $a_n^{-1} K(\cdot/a_n)$, où $K$ est une densité de probabilité à support borné et $a_n \to 0$. En d’autres termes, $f_n$ représente l’estimateur à noyau de la densité basé sur Kaplan–Meier. De façon analogue, on note $\bar{f}_n$ la convolution de la distribution des données non censurées, qui sera supposée avoir une densité bornée, avec ces mêmes noyaux. Pour tout $n$, soit $J_n$ la demi-droite délimitée supérieurement par $Z_n(1-\varepsilon_n)/a_n - a_n$, où $\varepsilon_n \to 0$ et, pour tout $m$, $Z_{m,n}$ désigne la $m$-ème statistique d’ordre des données censurées. Nous démontrons dans cet article que, sous des conditions minimales sur $a_n$ et $\varepsilon_n$, $\sup_{J_n} |f_n(t) - \bar{f}_n(t)|$ converge p.s. vers zéro quand $n \to \infty$ à une vitesse au moins égale à $\sqrt{\log(a_n \wedge \varepsilon_n)}/(na_n\varepsilon_n)$. Pour $\varepsilon_n$ constant, cette vitesse est comparable, à une constante multiplicative près, à la vitesse exacte dans le cas d’intervalles fixes. © 2001 Éditions scientifiques et médicales Elsevier SAS

E-mail addresses: gine@uconnvm.uconn.edu (E. Giné), guillou@ccr.jussieu.fr (A. Guillou).

1 Research supported in part by NSF Grants No. DMS-9625457 and No. DMS-0070382.
1. Introduction

Let $X_i, i \in \mathbb{N}$, be independent and identically distributed (i.i.d.) random variables with common distribution function (cdf) $F$, which we assume differentiable, with density $f$. Let $Y_i, i \in \mathbb{N}$, be a second i.i.d. sequence independent of the first, with cdf $G$, and let $Z_i = X_i \wedge Y_i$ and $\delta_i = I_{X_i \leq Y_i}$, $i \in \mathbb{N}$. We denote by $H$ the cdf of $Z$, by $\tau_H = \inf\{x: H(x) = 1\}$ the supremum of the support of $H$ and by $H_n$ and $H_n^{-1}$ respectively the empirical cdf and the empirical quantile function corresponding to $Z_1, \ldots, Z_n, n \in \mathbb{N}$. Let $\hat{F}_n(x), -\infty < x < \tau_H$, be the Kaplan–Meier [10] product limit estimator of $F(x)$. (See Section 2 for definitions.) A natural nonparametric estimator of $f$ is

$$fn(t) = \frac{1}{a_n} \int_{-\infty}^{\infty} K\left(\frac{t-x}{a_n}\right) \hat{F}_n(x), \quad n \in \mathbb{N}, \quad (1.1)$$

where $K$ is a probability kernel and $a_n$ is a sequence of positive constants tending to zero. In this article we are interested in the general problem of understanding how well does $fn$ estimate $f$. Diehl and Stute [7] contains an exact law of the iterated logarithm (LIL) for the variable

$$\sup_{t \leq T} |fn(t) - \bar{fn}(t)|,$$

where $T < \tau_H$ is fixed and

$$\bar{fn}(t) = \frac{1}{a_n} \int_{-\infty}^{\infty} K\left(\frac{t-x}{a_n}\right) dF(x) \quad (1.2)$$

is the convolution of $f$ with the approximate identity $a_n^{-1}K(x/a_n) dx$. (The ‘bias’ part, $\bar{fn} - f$, is ignored as it can always be balanced with the term $fn - \bar{fn}$ by calibrating the normalizing sequence $\{a_n\}$, provided enough regularity for $K$ is assumed.) Stute [14] introduced a.s. bounds for $|\hat{F}_n - F|$ uniform over varying data driven intervals that asymptotically cover the full domain of $H$, $(-\infty, \tau_H)$, and his analysis was refined in Csörgő [5] and Giné and Guillou [9]. In view of these developments it is only natural to ask whether the same idea can be applied to kernel density estimation, that is, whether sensible rates of a.s. convergence to zero can be obtained for the random variables

$$\sup_{t \leq \tau_n} |fn(t) - \bar{fn}(t)| \quad (1.3)$$

with $\tau_n = H_n^{-1}(1 - \epsilon_n)$, for suitable $\epsilon_n \rightarrow 0$. The object of this article is to provide such rates, which are given in Theorem 3.3 and Corollary 3.4 below. When we take $\epsilon_n = \text{constant}$, our results nearly recover, except for multiplicative constants, the main result in Diehl and Stute [7], which is optimal.

We should mention here that the upper limit $\tau_n = H_n^{-1}(1 - \epsilon_n)$ is eventually a.s. dominated by $T_n = H^{-1}(1 - \epsilon_n/8)$ if $n \epsilon_n \geq \log n$ for all $n$ large enough (see e.g., any
of the last mentioned three articles). Then, in order to obtain an asymptotic upper bound for (1.3), it suffices to obtain one for

\[ \sup_{t \leq T_n} |f_n(t) - \tilde{f}_n(t)|. \] (1.4)

Hence, although our results apply to (1.3), they are stated for (1.4) (our assumptions imply that \( n \varepsilon_n \) is indeed asymptotically larger than \( \log n \)).

This article may be considered as an application to density estimation of the methods and results from Giné and Guillou [9]. As in this last article, the main innovation with respect to previous work consists in using sharp exponential bounds for the empirical process (Alexander [2], Massart [11] and Talagrand [16]). The exponential bound used here combines Talagrand’s [16] exponential bound with convenient estimates of expected values of suprema of empirical processes over Vapnik–Cervonenkis classes of functions and their squares, also provided by Talagrand [15]. (See Einmahl and Mason [8] for a similar exponential bound.) In order to perform ‘blocking’ as in the classical proofs of the law of the iterated logarithm (LIL) we found a maximal inequality of Montgomery-Smith [12] very useful, just as in our previous work.

We became recently aware, through a referee, of the article of Bitouzé et al. [3], published after ours had already been submitted. This article contains a LIL for the Kaplan–Meier estimator, uniform over the whole line, obtained from van der Laan’s identity and an exponential bound for the empirical process over classes of functions that are not necessarily Vapnik–Cervonenkis. The results from the present article do not seem to follow directly from those in Bitouzé et al. [3]; in fact they are more related to results in Giné and Guillou [9] (compare, for instance their LIL, (2), with our LIL, (4.17)). The methods in their article may well provide an alternative approach to density estimation in the Kaplan–Meier framework, possibly leading to results different from those presented here.

2. An exponential inequality and other preparatory material

The next two propositions, on empirical processes, follow from Talagrand [15,16] with very little additional elaboration. We refer to Giné and Guillou [9], Lemma 3 and the paragraph before it, for the definition of measurable classes of functions \( \mathcal{F} \), VC (Vapnik–Cervonenkis) with respect to an envelope \( F \), as well as for the covering numbers \( N(T, d, \varepsilon) \) of a metric space \( (T, d) \). If \( \mathcal{F} \) is VC with respect to \( F = \sup \{|f| : f \in \mathcal{F}| \) then we simply say that \( \mathcal{F} \) is a VC class of functions. The functions in \( \mathcal{F} \) are assumed to be measurable real functions on a measurable space \( (S, \mathcal{S}) \), \( P \) is a probability measure on \( (S, \mathcal{S}) \), and \( \xi_i, i \in \mathbb{N}, \) are the coordinate functions \( S^N \mapsto S \), in particular, they are i.i.d. with common law \( P \). \( P_n := \sum_{i=1}^n \delta_{\xi_i}/n \) are the empirical measures corresponding to the sequence \( \xi_i \). Finally, \( \{\eta_i\} \) is a sequence of independent Rademacher variables, independent of \( \{\xi_i\} \), in fact, defined on another factor of a common product probability space. (We recall that a Rademacher variable \( \eta \) is one that satisfies \( \Pr[\eta = 1] = \Pr[\eta = -1] = 1/2 \).) Also, \( \|\Phi\|_{\mathcal{F}} := \sup_{f \in \mathcal{F}} |\Phi(f)| \).
PROPOSITION 2.1. – Let $\mathcal{F}$ be a measurable uniformly bounded VC class of functions. Let $\sigma^2 \geq \sup_{f \in \mathcal{F}} E_P f^2$ and $U \geq \sup_{f \in \mathcal{F}} \|f\|_{\infty}$ be such that $0 < \sigma \leq U$. Then there exist $A, v$ dependent on $\mathcal{F}$ but not on $P$ or $n$, $A \geq 3\sqrt{e}$, $v \geq 1$, such that, for all $n \in \mathbb{N}$,

$$E \left| \sum_{i=1}^{n} \eta_i f(\xi_i) \right|_{\mathcal{F}} \leq C \left[ vU \log \frac{AU}{\sigma} + \sqrt{v} \sqrt{n} \sigma \sqrt{\log \frac{AU}{\sigma}} \right],$$

(2.1)

where $C$ is a universal constant.

Proof. – Since $\mathcal{F}$ is VC, there exists $A$ and $v$ positive such that, if $F := \sup_{f \in \mathcal{F}} |f|$, then for all probability measures $P$ on $(\mathcal{S}, \mathcal{S})$ and $0 < \tau < 1$,

$$N(\mathcal{F}, L_2(P), \tau \| F \|_{L_2(P)}) \leq \left( \frac{A}{\tau} \right)^v.$$  

(2.2)

We can assume $A \geq 3\sqrt{e}$ and $v \geq 1$. We can also assume that $0 \in \mathcal{F}$. Then, the usual entropy bound for Rademacher processes (e.g., Corollary 5.1.8 in de la Peña and Giné [6]) gives

$$E_{\eta} \left| \sum_{i=1}^{n} \eta_i f(\xi_i) / \sqrt{n} \right|_{\mathcal{F}} \leq C \int_0^1 \sqrt{\log N(\mathcal{F}, L_2(P_n), \tau)} \, d\tau,$$

(2.3)

where $E_{\eta}$ denotes integration with respect to the Rademacher variables only and $C$ is a universal constant. On the other hand, by Talagrand [15], Corollary 3.4,

$$E \left| \sum_{i=1}^{n} f^2(\xi_i) \right|_{\mathcal{F}} \leq n\sigma^2 + 8U E \left| \sum_{i=1}^{n} \eta_i f(\xi_i) \right|_{\mathcal{F}}.$$  

(2.4)

Then, combining (2.2) and (2.3) and changing variables, we have

$$E_{\eta} \left| \sum_{i=1}^{n} \eta_i f(\xi_i) / \sqrt{n} \right|_{\mathcal{F}} \leq ACU \int_{AU/\left| \sum_{i=1}^{n} f^2(\xi_i)/n \right|_{\mathcal{F}}^{1/2}} \frac{\sqrt{\log \tau}}{\tau^2} \, d\tau$$

$$\leq C' \sqrt{v} \left| \sum_{i=1}^{n} f^2(\xi_i)/n \right|_{\mathcal{F}}^{1/2} \sqrt{\log \frac{A^2U^2}{\| \sum_{i=1}^{n} f^2(\xi_i)/n \|_{\mathcal{F}}}},$$

for another universal constant $C'$. By Hölder’s inequality and concavity of the function $y = x \log \frac{a}{x}$, the above gives

$$E \left| \sum_{i=1}^{n} \eta_i f(\xi_i) / \sqrt{n} \right|_{\mathcal{F}} \leq C' \sqrt{v} \sqrt{E \left| \sum_{i=1}^{n} f^2(\xi_i)/n \right|_{\mathcal{F}}} \log \frac{A^2U^2}{E \| \sum_{i=1}^{n} f^2(\xi_i)/n \|_{\mathcal{F}}}.$$  

Then, inequality (2.4) and the fact that the function $y = x \log \frac{a}{x}$ is increasing for $0 < x < a/e$ yield
Thus, setting
\[ Z = E \left\| \frac{\sum_{i=1}^{n} \eta_i f(\xi_i)}{\sqrt{n}} \right\| \],

Z satisfies the inequation
\[ Z^2 \leq C v n \sigma^2 \log \left( \frac{AU}{\sigma} \right) + 8 C v Z U \log \left( \frac{AU}{\sigma} \right), \]

where C is a universal constant. Hence, Z is between the two roots of the corresponding equation and, since one is negative and Z ≥ 0, we conclude that
\[
Z \leq 4 C v U \log \frac{A U}{\sigma} + \sqrt{16 C^2 v^2 U^2 \left( \log \frac{A U}{\sigma} \right)^2 + C v n \sigma^2 \left( \log \frac{A U}{\sigma} \right)}
\]
\[
\leq 8 C v U \log \frac{A U}{\sigma} + \sqrt{C \sqrt{v \sqrt{n} \sigma}} \sqrt{\log \frac{A U}{\sigma}},
\]

proving Proposition 2.1. □

For a similar proposition with a different proof see Einmahl and Mason [8], Proposition A.1, and for the same result for indicator functions, with a similar proof, see Talagrand [15], Proposition 6.2.

Since
\[
E \left\| \sum_{i=1}^{n} (f(\xi_i) - Ef(\xi_i)) \right\| \leq 2 E \left\| \sum_{i=1}^{n} \eta_i (f(\xi_i) - Ef(\xi_i)) \right\|
\]
and since, if \( \mathcal{F} \) is a measurable uniformly bounded VC-type class of functions so is \( \tilde{\mathcal{F}} := \{ f - Ef(\xi) : f \in \mathcal{F} \} \) (as is well known and can be seen by a simple estimation of covering numbers), we can apply the previous proposition to \( \tilde{\mathcal{F}} \) with \( U \) replaced by \( 2U \) and \( \sigma^2 \) satisfying instead the requirements \( \sigma^2 \geq \sup_{f \in \mathcal{F}} \text{Var}_P f \) and \( 0 < \sigma \leq 2U \). We then conclude that there is a universal constant C such that
\[
E \left\| \sum_{i=1}^{n} (f(\xi_i) - Ef(\xi_i)) \right\| \leq C \left[ v U \log \frac{A U}{\sigma} + \sqrt{v \sqrt{n} \sigma} \sqrt{\log \frac{A U}{\sigma}} \right], \tag{2.5}
\]
(C may be different from the constant in Proposition 2.1 as it must absorb several numerical factors due to symmetrization and to the change from \( U \) to \( 2U \)). Moreover, this and inequality (2.4) applied to \( \tilde{F} \) yield

\[
E \left\| \sum_{i=1}^{n} \left( f(\xi_i) - E f(\xi_1) \right) \right\|_{\mathcal{F}} \leq n\sigma^2 + 8CvU^2 \log \frac{AU}{\sigma} + 8C \sqrt{v} U \sqrt{\log \frac{AU}{\sigma}} \sqrt{n} \sigma
\]

\[
\leq \left( \sqrt{n} \sigma + L \sqrt{v} U \sqrt{\log \frac{AU}{\sigma}} \right)^2,
\]

(2.6)

for some universal constant \( L \).

Talagrand [16] proved the following exponential inequality for any measurable, uniformly bounded class of functions \( \mathcal{F} \):

\[
\Pr \left\{ \left\| \sum_{i=1}^{n} f(\xi_i) - E \sum_{i=1}^{n} f(\xi_i) \right\|_{\mathcal{F}} > t \right\} \leq K \exp \left\{ -\frac{1}{KU} \log \left( 1 + \frac{tU}{V} \right) \right\},
\]

(2.7)

valid for all \( t > 0 \), and where \( K \) is a universal constant, \( U \) is as above and \( V \) is any number satisfying \( V \geq E \sup_{f \in \mathcal{F}} \sum_{i=1}^{n} f^2(\xi_i) \). This inequality applied to \( \tilde{F} \), together with the estimates (2.5) and (2.6), then gives the following:

**Proposition 2.2.** Let \( \mathcal{F} \) be a measurable uniformly bounded VC class of functions, and let \( \sigma^2 \) and \( U \) be any numbers such that \( \sigma^2 \geq \sup_{f \in \mathcal{F}} \text{Var}_f \), \( U \geq \sup_{f \in \mathcal{F}} \| f \|_{\infty} \) and \( 0 < \sigma \leq U \). Then, there exist constants \( C \) and \( K \), depending only on the VC characteristics \( A \) and \( v \) of the class \( \mathcal{F} \), such that the inequality

\[
\Pr \left\{ \left\| \sum_{i=1}^{n} \left( f(\xi_i) - E f(\xi_1) \right) \right\|_{\mathcal{F}} > t \right\} \leq K \exp \left\{ -\frac{1}{KU} \log \left( 1 + \frac{tU}{V} \right) \right\}
\]

is valid for all

\[
t \geq C \left[ U \log \frac{AU}{\sigma} + \sqrt{n} \sigma \sqrt{\log \frac{AU}{\sigma}} \right].
\]

(2.8)

(2.9)

If in Proposition 2.2 we assume \( 0 < \sigma \leq cU \) for some \( c < 1 \) then \( \log A \leq d \log(U/\sigma) \) for some \( d < \infty \) and we can replace \( \log(AU/\sigma) \) by \( \log(U/\sigma) \) in both (2.8) and (2.9) at the price of changing the constants \( K \) and \( C \) (that now depend on \( c \)).

The exponential bound in Proposition 2.2 is much less elaborate than the bounds in Talagrand [15] but it suits our purposes: whereas it is important for us that the multiplicative constant \( K \) outside the exponent in (2.8) should not depend on \( \sigma \), we are not presently interested in the optimal value of the multiplicative constant within the exponent of (2.8) (here, also \( K \)). On the other hand, there are parts of the range of \( t \) for which the bound (2.8) does apply but the bounds in Alexander [2], Massart [11] and Einmahl and Mason [8] do not.
We will use inequality (2.8) for $0 < \sigma \leq U/2$ and under the additional assumption

$$\sqrt{n} \sigma \geq C_1 U \sqrt{\log \frac{U}{\sigma}}$$

(2.10)

and then, with

$$t = C_2 \sqrt{n} \sigma \sqrt{\log \frac{U}{\sigma}},$$

(2.11)

for fixed $C_1$ and $C_2$, with $C_2$ large, in which case, it gives

$$\Pr\left\{ \left\| \sum_{i=1}^{n} (f(\xi_i) - E f(\xi_i)) \right\| > C_2 \sigma \sqrt{n} \sqrt{\log \frac{U}{\sigma}} \right\} \leq K \exp\left\{ -\frac{D}{K} \log \frac{U}{\sigma} \right\},$$

(2.12)

where $D$ can be taken to be $D = 4C_2(1 + C_1^{-1})^{-2} \log(1 + 4^{-1} C_2 K^{-1})$, as can be easily checked (and therefore, $D \to \infty$ as $C_2 \to \infty$ for each $C_1$ fixed). To see this, we may proceed as follows: (i) we take $K$ and $C$ in (2.8) and (2.9) for $c = 1/2$, so that $A = 1$, as indicated below (2.9), and then note that the value of $t$ prescribed by (2.11) satisfies (2.9) as long as $C_2 \geq C(1 + C_1^{-1})$; (ii) since $x^{-1} \log(1 + x)$ is monotone decreasing for $x > 0$ and $4xy \leq (x + y)^2$, so that $C_2 \sqrt{n} \sigma U \sqrt{\log(U/\sigma)}/[K(\sqrt{n} \sigma + U \sqrt{\log(U/\sigma)})^2] \leq C_2/(4K)$, we have

$$\log \left( 1 + \frac{C_2 \sqrt{n} \sigma U \sqrt{\log(U/\sigma)}}{K(\sqrt{n} \sigma + U \sqrt{\log(U/\sigma)})^2} \right) \geq \frac{4 \log(1 + C_2/(4K)) \sqrt{n} \sigma U \sqrt{\log(U/\sigma)}^2}{(\sqrt{n} \sigma + U \sqrt{\log(U/\sigma)})^2};$$

and (iii) inequality (2.12) follows directly from these observations together with condition (2.10).

As a first application of these inequalities we prove a lemma that will be useful throughout. We recall that the quantile function of $H$ is defined as $H^{-1}(x) = \inf\{z: H(z) \geq x\}$ for $x \in (0, 1)$, and that $H(H^{-1}(x)-) \leq x \leq H(H^{-1}(x))$. It is convenient to make the following definition: we say that a nonincreasing sequence of numbers $\{e_n\}$ is regular if there exists a positive constant $A$ such that $e_{2n} \geq Ae_n$ for all $n$.

**Lemma 2.3.** Let $\{e_n\}$ be a regular nonincreasing sequence such that $0 < e_n < 1$ and

$$\lim_{n \to \infty} \frac{n e_n}{\log e_n} = \infty.$$  

(2.13)

Let

$$T_n := H^{-1}(1 - e_n).$$

(2.14)

Then,

$$\sup_{x \leq T_n} \left| \frac{1 - H_n(x-)}{1 - H(x-)} - 1 \right| = O \left( \sqrt{\frac{(\log \frac{1}{e_n}) \vee \log \log n}{n e_n}} \right) \text{ a.s.}$$

(2.15)
and, in particular,

\[
\lim_{n \to \infty} \sup_{x \leq T_n} \frac{1 - H_n(x -)}{1 - H(x -)} = 1 \quad \text{a.s.} \quad (2.16)
\]

Hence, also

\[
\sup_{x \leq T_n} \frac{1 - H(x -)}{1 - H_n(x -)} - 1 = O\left(\sqrt{\log \frac{1}{\varepsilon_n} \vee \log \log n} \right) \quad \text{a.s.} \quad (2.15')
\]

and

\[
\lim_{n \to \infty} \sup_{x \leq T_n} \frac{1 - H(x -)}{1 - H_n(x -)} = 1 \quad \text{a.s.} \quad (2.16')
\]

Proof. – We have

\[
\frac{1 - H_n(x -)}{1 - H(x -)} - 1 = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{I_{Z_i \geq x}}{1 - H(x -)} - E \frac{I_{Z_i \geq x}}{1 - H(x -)} \right) \quad := \sum_{i=1}^{n} (f_{x,n}(Z_i) - Ef_{x,n}(Z)).
\]

For each \(n\), the family of functions \(\{f_{x,n}: x \leq T_n\}\) is obviously bounded, it is measurable because it is parametrized by a half line and \(f_{x,n}(t)\) is jointly measurable in \(x\) and \(t\), and it is a \(VC\) class because of its monotonicity properties (each function \(f_{x,n}\) is the difference of a constant \(c_x\) and a function \(g_{x,n}\) such that the functions \(g_{x,n}\) increase as \(x\) increases: see, e.g., Lemma 3, b) and c), in Giné and Guillou [9]). Thus, we can apply the exponential inequalities above, in this case, Talagrand’s inequality (2.7) to \(\tilde{F} := \{f_{x,n} - Ef_{x,n}\}\) in conjunction with the estimate (2.6) of \(V\). We can obviously take \(U_n = (n\varepsilon_n)^{-1}\) and, since \(Ef_{x,n}^2(Z) = (1 - H(x -))^{-1}/n^2, x \leq T_n\), we can take \(\sigma^2 = (n^2\varepsilon_n)^{-1}\). Hence,

\[
\Pr\left\{ \sup_{x \leq T_n} \left| \sum_{i=1}^{n} (f_{x,n}(Z_i) - Ef_{x,n}(Z)) \right| - E \left( \sup_{x \leq T_n} \left| \sum_{i=1}^{n} (f_{x,n}(Z_i) - Ef_{x,n}(Z)) \right| \right) > C \sqrt{\log \log n \over n \varepsilon_n} \right\} \leq K \exp\left\{ -{C^2 \over 2K} \log \log n \left(1 + \frac{C \sqrt{\log \log n / (n \varepsilon_n)^{3/2}}}{\sqrt{\log \varepsilon n} + 2L(1/(n \varepsilon_n)) \sqrt{\log \varepsilon n}^2} \right) \right\},
\]

which, by (2.13), is dominated by

\[
K \exp\left\{ -{C^2 \over 2K} \log \log n \right\}
\]

for all \(n\) large enough (since, as can be easily argued e.g. by contradiction, (2.13) implies \(\lim_{n \to \infty} n \varepsilon_n / \log n = \infty\)). Also, given the values assigned to \(U\) and \(\sigma^2\), (2.5) shows that the expected value of the sup of the process over \(\{x \leq T_n\}\) is of the order of
We then conclude that there exist $A$, $D$ and $n_0$ such that

$$\Pr \left\{ \sup_{x \leq T_n} \left| \frac{1 - H_n(x -)}{1 - H(x -)} - 1 \right| > A \sqrt{\frac{(\log \varepsilon_n^{-1} \vee \log \log n)}{n \varepsilon_n}} \right\} \leq D \exp\{-2 \log \log n\}$$

(2.17)

for all $n \geq n_0$. Setting, for ease of notation, $B_{2^n} := A^2 \left[ (\log \varepsilon_n^{-1} \vee \log \log n) / (n \varepsilon_n) \right] / (\log 2)^2 k^2$, the regularity of the sequence $\{\varepsilon_n\}$ implies that there exists $d > 0$ such that $B_{2^n} \leq dB_n$ for all $2^{k-1} < n \leq 2^k$, for all $k > \log n_0$. Then, by (2.17) and Montgomery-Smith’s [12] maximal inequality (see, e.g., de la Peña and Giné [6]), we have

$$\Pr \left\{ \max_{2^{k-1} < n \leq 2^k} B_{2^n}^{-1} \sup_{x \in [T_n]} \left| \frac{1 - H_n(x -)}{1 - H(x -)} - 1 \right| > 30d \right\} \leq 9 \Pr \left\{ \sup_{x \in [T_n]} \left| \frac{1 - H_n(x -)}{1 - H(x -)} - 1 \right| > B_{2^k} \right\} \leq \frac{9D}{(\log 2)^2 k^2}.$$ 

Now (2.15) follows by Borel–Cantelli. Condition (2.13) implies $n \varepsilon_n / \log \log n \to \infty$ since, as mentioned above, $n \varepsilon_n / \log n \to \infty$. Therefore the bound in (2.15) is $o(1)$, which implies (2.16) as well as (2.15′) and (2.16′). \qed

The scheme of proof of the previous lemma is used repeatedly throughout. We will refer to the above proof rather than reproduce repetitious arguments.

Next, following Csörgő [5], we describe a bound for the product limit estimator that follows from the classical expansion of Breslow and Crowley [4]. We need some additional notation, borrowed from Stute [14] and Csörgő [5]. We set $\tilde{H}(x) = \Pr\{Z \leq x, \delta = 1\}$, $-\infty < x \leq \tau_H$, and define $\tilde{H}_n$ to be its empirical counterpart, that is,

$$\tilde{H}_n(x) = \frac{1}{n} \sum_{i=1}^{n} I_{\{Z_i \leq x, \delta_i = 1\}}, \quad n \in \mathbb{N},$$

for $-\infty < x \leq \tau_H$. (We should recall from the introduction that $H$ is the cdf of $Z$ and that, for each $n$, $H_n(x) = \sum_{i=1}^{n} I_{\{Z_i \leq x\} / n}$.) The obvious facts that $d\tilde{H} \leq dH$ and $d\tilde{H}_n \leq dH_n$ will be used without further mention. We should recall that, with this notation and the notation set up in the introduction, the cumulative hazard function of $X$ is

$$\Lambda(x) = \int_{-\infty}^{x} \frac{dF(y)}{1 - F(y -)} = \int_{-\infty}^{x} \frac{d\tilde{H}(y)}{1 - H(y -)}, \quad x \in (-\infty, \tau_H).$$
and its Nelson–Aalen estimator (Nelson [13], and Aalen [1]) is
\[ \Lambda_n(x) = \int_{-\infty}^{x} \frac{d\tilde{H}_n(y)}{1 - H_n(y^-)}, \]
defined for \( x < \max_{i \leq n} Z_i := Z_{n,n} \). The product limit estimator (Kaplan and Meier [10]) is defined as
\[ 1 - \hat{F}_n(x) = \prod_{j=1}^{n} \left[ 1 - \frac{\delta_{j,n} I_{Z_{j,n} \leq x}}{n - j + 1} \right] \]
for all \( x < Z_{n,n} \), where \( Z_{j,n} \) are the order statistics of \( Z_1, \ldots, Z_n \) and \( \delta_{j,n} = \delta_k \) iff \( Z_{j,n} = Z_k \). Note that if we take \( T_n \) as in the introduction, \( T_n < Z_{n,n} \) a.s. If \( F \) is continuous, for any real function \( h \) and for all \( x < Z_{n,n} \), we have (Csörgő [5]):
\[ \left| \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - h(x) \right| \leq \left| (\Lambda_n(x) - \Lambda(x)) - h(x) \right| + \left| R_{n,6}(x) \right| \]  \tag{2.18}
where
\[ R_{n,6}(x) = \left[ \frac{1}{2} \right] \left| \Lambda_n(x) - \Lambda(x) \right|^2 + \left| \epsilon_n(x) \right| \exp\left( |\epsilon_n(x)| \right) \exp\left( |\Lambda_n(x) - \Lambda(x)| \right) \]  \tag{2.19}
and
\[ \epsilon_n(x) = - \log(1 - \hat{F}_n(x)) - \Lambda_n(x). \]  \tag{2.20}
\( \Lambda_n - \Lambda \) further decomposes as:
\[ \Lambda_n(x) - \Lambda(x) = \int_{-\infty}^{x} \frac{d(\tilde{H}_n - H)(y)}{1 - H(y^-)} \]
\[ + \int_{-\infty}^{x} \frac{H_n(y^-) - H(y^-)}{(1 - H_n(y^-))(1 - H(y^-))} d\tilde{H}_n(y) \]
\[ := L_{n,1}(x) + R_n(x). \]  \tag{2.21}
We note that \( L_{n,1} \) is only part of the linearization \( L_n \) of \( \Lambda_n - \Lambda \) considered, e.g., in Giné and Guillou [9].

The probability kernels \( K \) we will consider satisfy the following condition:
\[ K \text{ is differentiable with bounded derivative and vanishes on } [-1, 1]^c. \]  \tag{2.22}
The case of \( K \) vanishing outside \([r, s] \), \(-\infty < r < s < \infty \), is not more general as it reduces to \( K \) vanishing outside \([-1, 1] \) by translation and dilation. We take the limits \(-1 \) and \( 1 \) just for convenience. We then have (by (1.1), (1.2), (2.21), (2.22) and integration by parts):
\[ f_n(t) - \tilde{f}_n(t) = \frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) \left( \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - \left( \Lambda_n(x) - \Lambda(x) \right) \right) dK \left( \frac{t-x}{a_n} \right) \]

\[ - \frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) L_{n,1}(x) dK \left( \frac{t-x}{a_n} \right) \]

\[ - \frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) R_n(x) dK \left( \frac{t-x}{a_n} \right). \]  

The proof of the main result in the next section consists in estimating the sizes of these three terms. We anticipate that the second term dominates.

3. The order of magnitude of \( f_n - \tilde{f}_n \)

In what follows we assume that \( \{\varepsilon_n\} \) and \( \{a_n\} \) are nonincreasing regular sequences such that \( a_n \to 0 \) and both,

\[ \lim_{n \to \infty} \frac{na_n \varepsilon_n}{\log a_n \varepsilon_n} = \infty \]  

(3.1)

and

\[ \lim_{n \to \infty} \frac{\log \left( \frac{1}{a_n \varepsilon_n} \right)}{\log \log n} = \infty. \]  

(3.2)

In particular, \( \{\varepsilon_n\} \) satisfies hypothesis (2.13) from Lemma 2.3. We assume, in addition, two conditions which may seem less natural but that nevertheless are not too restrictive, namely, that

\[ \lim \inf \frac{ne_n}{d_n \log n} > 0 \]  

(3.3)

where \( d_n \downarrow \infty \) is such that \( \sum [kd_n \log k]^{-1} < \infty \) (such as, for instance, \( d_n = (\log \log \log n)^{1+\delta} \) for some \( \delta > 0 \), and that

\[ a_n \left( \log \frac{1}{\varepsilon_n} \right)^2 \to 0. \]  

(3.4)

Condition (3.4) is obviously satisfied if \( a_n \leq \varepsilon_n \), and, since by (3.3) \( \log e_n^{-1} < \log n \) for all \( n \) large enough, it also holds if \( a_n (\log n)^2 \to 0 \).

We also set

\[ T_n = H^{-1}(1 - e_n). \]  

(2.14)

Then, condition (3.3) implies that \( ne_n / \log n \to \infty \) and therefore, as mentioned in the introduction, it follows, e.g., by Remark 3 in Giné and Guillou [9], that \( T_n \) dominates \( H_n^{-1}(1 - 3e_n) \) eventually a.s., hence also \( Z_{n(1 - 3e_n),n} \) if the numbers \( 3n\varepsilon_n \) are integers. So, although the results that follow are stated in terms of \( T_n \) they are really results on the
sup of $|f_n - \tilde{f}_n|$ over adaptive random intervals that tend to $\tau_H$. Finally, we also assume that

$$F \text{ and } G \text{ are differentiable and } F' := f \text{ is uniformly bounded.} \quad (3.5)$$

**Lemma 3.1.** Let $F$ and $G$ be cdf’s satisfying condition (3.5), let $K$ be a probability kernel satisfying condition (2.22), and let $\{\varepsilon_n\}$ and $\{a_n\}$ be two nonincreasing regular sequences satisfying $a_n \to 0$ and conditions (3.1) and (3.2). We assume also that $\{\varepsilon_n\}$ satisfies (3.3). Then,

$$\sup_{t \leq T_n - a_n} \frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) \left( \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - (\Lambda_n(x) - \Lambda(x)) \right) \frac{1 - x}{a_n} \, dK \left( t - x \right)$$

$$= o \left( \frac{\log \frac{1}{a_n} \wedge \varepsilon_n}{n a_n \varepsilon_n} \right) \text{ a.s.} \quad (3.6)$$

**Proof.** Taking $h(x) = \Lambda_n(x) - \Lambda(x)$ in (2.18) gives

$$\left| \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - (\Lambda_n(x) - \Lambda(x)) \right| \leq |R_{n,6}(x)|,$$

and therefore,

$$\sup_{t \leq T_n - a_n} \frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) \left( \frac{\hat{F}_n(x) - F(x)}{1 - F(x)} - (\Lambda_n(x) - \Lambda(x)) \right) \frac{1 - x}{a_n} \, dK \left( t - x \right)$$

$$\leq 2 \|K'\|_{\infty} \sup_{t \leq T_n} |R_{n,6}(t)|.$$

Now, Theorem 6 in Giné and Guillou [9], gives

$$\sup_{t \leq T_n} (\Lambda_n(t) - \Lambda(t))^2 = O \left( \frac{\log \log n}{n \varepsilon_n} \right) \text{ a.s.}$$

on account of (3.3) and the regularity of the sequence $\{\varepsilon_n\}$. Moreover, by Lemma 1 in Breslow and Crowley [4], if $x \leq T_n$,

$$0 \leq \ell_n(t) = - \log (1 - \hat{F}_n(t)) - \Lambda_n(t) \leq \frac{H_n(t -)}{n(1 - H_n(t -))} \text{ a.s.}$$

and, by Lemma 2.3,

$$\sup_{t \leq T_n} \frac{H_n(t -)}{n(1 - H_n(t -))} \leq \frac{1}{n(1 - H(T_n))} \frac{1 - H(T_n)}{1 - H_n(T_n -)} = O \left( \frac{1}{n \varepsilon_n} \right) \text{ a.s.}$$

Then, since these bounds tend to zero by (3.1) and (3.2), combining them with (2.19) and (2.20) yields

$$\sup_{t \leq T_n} |R_{n,6}(t)| = O \left( \frac{\log \log n}{n \varepsilon_n} \right) \text{ a.s.}$$
Now the result follows from condition (3.2). □

**Lemma 3.2.** Under the assumptions of Lemma 3.1 plus condition (3.4) for \( \{a_n\} \), we have

\[
\sup_{t \leq T_n - a_n} \frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) R_n(x) \, dK \left( \frac{t - x}{a_n} \right) = o \left( \sqrt{\frac{\log(a_n \wedge \varepsilon_n)}{na_n \varepsilon_n}} \right) \quad \text{a.s.} \quad (3.7)
\]

**Proof.** Changing variables, we can write

\[
\frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) R_n(x) \, dK \left( \frac{t - x}{a_n} \right) = \frac{1}{a_n} \int_{-1}^{1} (1 - F(t - a_nu)) (R_n(t - a_nu) - R_n(t)) K'(u) \, du + \frac{1}{a_n} R_n(t) \int_{-1}^{1} (1 - F(t - a_nu)) K'(u) \, du
\]

\[
= (I_n) + (II_n).
\]

**Order of magnitude of** \((II_n)\). Since \( K \) is a probability kernel on \([-1, 1] \) and, by (2.22), \( K(1) = K(-1) = 0 \), it follows by integration by parts that

\[
\sup_{t \leq T_n - a_n} |(II_n)| \leq \frac{1}{a_n} \left( \sup_{t \leq T_n} |R_n(t)| \right) \left( \sup_{t \leq T_n - a_n} \left| \int_{-1}^{1} K(u) \, d(1 - F(t - a_nu)) \right| \right)
\]

\[
\leq \|f\| \sup_{t \leq T_n} |R_n(t)|. \quad (3.8)
\]

We consider two cases according as to whether \( a_n \geq \varepsilon_n \) or \( a_n < \varepsilon_n \). First, we assume \( a_n \geq \varepsilon_n \) for all \( n \) (strictly, we should just consider the subsequence of those integers \( n \) for which \( a_n \geq \varepsilon_n \) but, for ease of notation, we will assume that this subsequence is \( N \) as the changes in the proof if it is not all of \( N \) are only formal). In this case it is convenient to use the bound

\[
\sup_{t \leq T_n} |R_n(t)| \leq \sup_{t \leq T_n} \left| \frac{H_n(t) - H(t)}{1 - H(t)} \right| \int_{-\infty}^{T_n} \frac{d\tilde{H}_n(y)}{1 - H_n(y)}. \quad (3.9)
\]

By Lemma 2.3,

\[
\sup_{t \leq T_n} \left| \frac{H_n(t) - H(t)}{1 - H(t)} \right| = O \left( \sqrt{\frac{(\log \frac{1}{\varepsilon_n}) \vee \log \log n}{n \varepsilon_n}} \right) \quad \text{a.s.} \quad (3.10)
\]
By (2.16) and (2.16') in Lemma 2.3, we also have

\[ \int_{-\infty}^{T_n} \frac{d\tilde{H}_n(y)}{1 - H_n(y^-)} \leq \int_{-\infty}^{T_n} \frac{d\tilde{H}_n(y)}{1 - H(y^-)} \text{ a.s.,} \tag{3.11} \]

where \( A_n \asymp B_n \) means that \( A_n/B_n \) and \( B_n/A_n \) are \( O(1) \) a.s. We will estimate the right hand side of (3.11) using Prohorov’s inequality (for convenience, Talagrand’s applied to a single function) and then will proceed as in the proof of Lemma 2.3. The last integral in (3.11) is dominated as follows:

\[ T_n \int_{-\infty}^{-H(Z_i)} d\tilde{H}_n(y) \leq T_n \int_{-\infty}^{-H(Z)} dH_n(y) = \frac{1}{n} \sum_{i=1}^{n} I_{Z_i \leq T_n}. \]

The expected value of this average is

\[ E \left( \frac{I_{Z \leq T_n}}{1 - H(Z)} \right) = \int_{-\infty}^{T_n} \frac{dH}{1 - H} = \log \frac{1}{\varepsilon_n}, \]

and the parameters \( U \) and \( V \) in Talagrand’s inequality (2.7) can be taken to be respectively \( U_n := (n\varepsilon_n)^{-1} \) and

\[ \frac{1}{n} E \left( I_{Z \leq T_n} \right)^2 = \frac{1}{n} \int_{-\infty}^{T_n} \frac{dH}{(1 - H)^2} \leq \frac{1}{n\varepsilon_n} =: V_n. \]

Since condition (2.13) holds (as noted above, (2.13) is implied by (3.1)), Talagrand’s (or Prohorov’s) inequality shows that there exists \( C < \infty \) such that

\[
\begin{align*}
\Pr \left\{ \max_{2^k-1 \leq n \leq 2^k} \int_{-\infty}^{T_n} \frac{dH_n(y)}{1 - H(y^-)} > 2 \log \frac{1}{\varepsilon_n} + 2C \sqrt{\log \frac{\varepsilon_n^{-1}}{2k\varepsilon_n}} \right\} &
\leq \Pr \left\{ \int_{-\infty}^{T_n} \frac{dH_n(y)}{1 - H(y^-)} > \log \frac{1}{\varepsilon_n} + C \sqrt{\log \frac{\varepsilon_n^{-1}}{2k\varepsilon_n}} \right\} \\
&
\leq \Pr \left\{ \int_{-\infty}^{T_n} \frac{d(H_n - H)(y)}{1 - H(y^-)} > C \sqrt{\log \frac{\varepsilon_n^{-1}}{2k\varepsilon_n}} \right\} \\
&
\leq K \exp \left\{ -\log \frac{1}{\varepsilon_n} \right\}
\end{align*}
\]

for all \( k \). By hypothesis (3.2) and the assumption \( a_n \geq \varepsilon_n \),

\[ \sum \exp \left\{ -\log \frac{1}{\varepsilon_n} \right\} < \infty, \]
and therefore, Borel–Cantelli and the regularity of the sequence \( \{e_n\} \) imply that

\[
\int_{-\infty}^{T_n} \frac{dH_n(y)}{1 - H(y^-)} = O \left( \log \frac{1}{e_n} + C \sqrt{\log \frac{e_n^{-1}}{n e_n}} \right) \quad \text{a.s.}
\]

The term \( \log e_n^{-1} \) dominates (as \( e_n \leq a_n \to 0 \) and, by (3.1), \( n e_n \to \infty \)) and we have

\[
\int_{-\infty}^{T_n} \frac{d\tilde{H}_n(y)}{1 - H(y^-)} \leq \int_{-\infty}^{T_n} \frac{dH_n(y)}{1 - H(y^-)} = O \left( \log \frac{1}{e_n} \right) \quad \text{a.s.} \tag{3.12}
\]

Combining (3.8)–(3.12) with (3.4), proves that

\[
\sup_{t \leq T_n - a_n} |(II_n)| = o \left( \sqrt{\log \log \frac{n}{n}} \right) \quad \text{a.s.} \tag{3.13}
\]

assuming \( a_n \geq e_n \). Let now \( a_n < e_n \) (again, without real loss of generality, we assume this holds for all \( n \in \mathbb{N} \)). Then, we write

\[
\sup_{t \leq T_n - a_n} |(II_n)| = o \left( \sqrt{\log \frac{a_n \wedge e_n}{n a_n e_n}} \right) \quad \text{a.s.} \tag{3.13}
\]

By Lemma 2.3, we can replace \( 1 - H(y^-) \) in the integral (3.14) by \( 1 - H_n(y^-) \) (as in (3.11), where the opposite replacement is made) and then we can apply Lemma 2.1 from Stute [14] to the effect that

\[
\int_{-\infty}^{T_n} \frac{d\tilde{H}_n(y)}{(1 - H(y^-))^2} \leq \int_{-\infty}^{T_n} \frac{dH_n(y)}{(1 - H(y^-))^2} \leq \frac{2}{1 - H_n(T_n^-)}.
\]

By (2.16) in Lemma 2.3 and the definition of \( T_n \), this random variable is \( O(e_n^{-1}) \) a.s. Hence,

\[
\int_{-\infty}^{T_n} \frac{d\tilde{H}_n(y)}{(1 - H_n(y^-))(1 - H(y^-))} = O \left( \frac{1}{e_n} \right) \quad \text{a.s.} \tag{3.16}
\]

Since \( e_n \geq \sqrt{a_n e_n} \), combining (3.8), (3.14)–(3.16) with condition (3.2) yields

\[
\sup_{t \leq T_n - a_n} |(II_n)| = o \left( \sqrt{\log \frac{a_n \wedge e_n}{n a_n e_n}} \right) \quad \text{a.s.} \tag{3.17}
\]
for $a_n < \varepsilon_n$. Hence, by (3.13) and (3.17), we have that in all cases, the lemma is proved for the component $(I_n)$ of the left side variable in (3.7).

Order of magnitude of $(I_n)$. If in $(I_n)$ we replace the factor $1 - F(t - a_n u)$ by $1 - F(t + a_n)$, the difference is dominated by

$$
\frac{4}{a_n} \sup_{t \leq T_n} |R_n(t)| \sup_{R, |u| \leq 1} |F(t - a_n u) - F(t + a_n)|
$$

$$
\leq 8 \|f\|_\infty \sup_{t \leq T_n} |R_n(t)|,
$$

which is $o\left(\sqrt{(na_n \varepsilon_n)^{-1} \log (a_n \wedge \varepsilon_n)^{-1}}\right)$ a.s. by the first part of this proof. Hence, $K'$ being bounded, it suffices to prove that

$$
\frac{1}{a_n} \sup_{t \leq T_n - a_n, -1 \leq u \leq 1} \left| (1 - F(t + a_n)) (R_n(t - a_n u) - R_n(t)) \right| = o\left(\frac{\log (a_n \wedge \varepsilon_n)^{-1}}{na_n \varepsilon_n}\right). \quad (3.18)
$$

So, we must look at the process

$$
\frac{1}{a_n} (1 - F(t + a_n)) \int_{t - a_n u}^{t + a_n u} \frac{H_n(y) - H(y)}{(1 - H_n(y))(1 - H(y))} \, d\tilde{H}_n(y)
$$

on the parameter set $-\infty < t \leq T_n - a_n$, $-1 \leq u \leq 1$. For ease of notation, we will only consider $-1 \leq u \leq 0$. By factoring out

$$
\sup_{y \leq T_n} \left| \frac{H_n(y) - H(y)}{1 - H(y)} \right| = O\left(\sqrt{\frac{(\log \varepsilon_n)^{-1} \vee \log n}{n \varepsilon_n}}\right) \quad \text{a.s.} \quad (3.19)
$$

(by Lemma 2.3) it suffices to consider the process

$$
\frac{1}{a_n} (1 - F(t + a_n)) \int_{t}^{t + a_n u} \frac{d\tilde{H}_n(y)}{1 - H_n(y)} - \infty < t \leq T_n - a_n, \ 0 \leq u \leq 1,
$$

which, again by Lemma 2.3, is of the same order as

$$
\frac{1}{a_n} (1 - F(t + a_n)) \int_{t}^{t + a_n u} \frac{d\tilde{H}_n(y)}{1 - H(y)} \leq \frac{1}{a_n} \int_{t}^{t + a_n u} \frac{d\tilde{H}_n(y)}{1 - H(y)}
$$

$$
= \frac{1}{na_n} \sum_{i=1}^{n} \frac{I_{X_i \in [t, t + a_n u]} I_{X_i \leq Y_i}}{1 - G(X_i)}
$$

(on the same parameter set). The expected value of this process for each $t$ and $u$ satisfies

$$
\left| \frac{1}{na_n} \sum_{i=1}^{n} \frac{I_{X_i \in [t, t + a_n u]} I_{X_i \leq Y_i}}{1 - G(X_i)} \right| = \left| \frac{1}{a_n} \int_{t}^{t + a_n u} \frac{1 - G(x)}{1 - G(x)} \, dF(x) \right| \leq \| f \|_\infty,
$$
and we will apply the proof of Lemma 2.3 to show that, in fact, the sup of the difference between the process and its expected value is asymptotically negligible. The corresponding class of functions is a bounded measurable VC class and we can take
\[ U_n = \frac{1}{n a_n \varepsilon_n}, \quad \sigma_n^2 = \|f\|_\infty \frac{1}{n^2 a_n \varepsilon_n} \geq \frac{1}{n^2 a_n^2} E \left( \frac{I_{X(t) \leq a n}}{1 - G(X(t))} \right). \]

Condition (2.10) holds for these parameters so that we can use inequality (2.12) and proceed as in the proof of Lemma 2.3. Here are the details: Combining Montgomery-Smith maximal inequality and inequality (2.12), and setting
\[ B_n = C \sqrt{\log \frac{1}{na_n \varepsilon_n}} \]
for a conveniently chosen large constant C (note \( n a_n B_n \geq c 2^k a_2 B_2^k \) for some \( c > 0 \) and \( 2^{k-1} < n \leq 2^k \)), we obtain
\[ \Pr \left\{ \max_{2^{k-1} \leq n \leq 2^k} \sup_{0 \leq v \leq 1} \left| \frac{1}{n a_n B_n} \sum_{i=1}^n \frac{I_{X_i(t) \leq a n} I_{X_i \leq Y_i}}{1 - G(X_i)} - \int t^+ dF(x) \right| > 1 \right\} \]
\[ \leq \Pr \left\{ \max_{2^{k-1} \leq n \leq 2^k} \sup_{\{t,v\}: t \leq \tau_{k-1}, 0 \leq v \leq a_{k-1}, t+v \leq \tau_k} \left| \frac{1}{2^k a_2 B_2^k} \right| > c \right\} \]
\[ \times \sum_{i=1}^n \left( \frac{I_{X_i(t) \leq a n} I_{X_i \leq Y_i}}{1 - G(X_i)} - \int t^+ dF(x) \right) > c \]
\[ \leq 9 \Pr \left\{ \sup_{\{t,v\}: t \leq \tau_{k-1}, 0 \leq v \leq a_{k-1}, t+v \leq \tau_k} \left| \frac{1}{2^k a_2 B_2^k} \sum_{i=1}^n \frac{I_{X_i(t) \leq a n} I_{X_i \leq Y_i}}{1 - G(X_i)} - \int t^+ dF(x) \right| > c \right\} \]
\[ \leq K \exp \left\{ - \log \frac{1}{a_2 E x^2} \right\}. \] (3.20)

Since \( \log(a_n \varepsilon_n)^{-1} \simeq \log(a_n \wedge \varepsilon_n)^{-1} \), condition (3.2) implies that this is the general term of a convergent series, hence, by Borel–Cantelli, the process under consideration is a.s. of the order of the sup of its expected values (bounded by \( \|f\|_\infty \), which is finite), that is,
\[ \sup_{t \leq \tau_{k-1} - a_n} \frac{1}{a_n} \left( 1 - F(t + a_n) \right) \frac{t a_n^+}{1 - H_n(y-)} = O(1) \text{ a.s.} \] (3.21)

Now, (3.18) follows from (3.19) and (3.21) because, by (3.2), \( (\log \varepsilon_n^{-1}) \vee \log \log n \) is asymptotically smaller than \( \log(a_n \wedge \varepsilon_n)^{-1} \) and \( a_n \to 0 \). This shows that the component \( (J_n) \) of the left hand side of (3.7) is of the prescribed order, which completes the proof of the lemma.
We can now proceed to prove the LIL result for densities.

**Theorem 3.3.** – Assuming:
(a) \( F \) and \( G \) satisfy condition (3.5),
(b) the probability kernel \( K \) satisfies condition (2.22) and
(c) the sequences \( \{ \varepsilon_n \} \) and \( \{ a_n \} \) are regular,
nonincreasing, satisfy conditions (3.1)–(3.4) and \( a_n \to 0 \); letting \( T_n = H^{-1}(1 - \varepsilon_n) \) and
letting \( f_n \) and \( \bar{f}_n \) be as defined by Eqs. (1.1) and (1.2), we have

\[
\sup_{t \leq T_n-a_n} |f_n(t) - \bar{f}_n(t)| = O \left( \sqrt{\frac{\log (a_n \wedge \varepsilon_n)^{-1}}{n a_n \varepsilon_n}} \right) \quad \text{a.s.} \quad (3.22)
\]

**Proof.** – By the decomposition (2.23) and Lemmas 3.1 and 3.2, it suffices to show

\[
\sup_{t \leq T_n-a_n} \frac{1}{a_n} \left| \int_{t-a_n}^{t+a_n} (1 - F(x)) L_{n,1}(x) dK \left( \frac{t - x}{a_n} \right) \right| = O \left( \sqrt{\frac{\log (a_n \wedge \varepsilon_n)^{-1}}{n a_n \varepsilon_n}} \right) \quad (3.23)
\]

We decompose this integral as in the proof of Lemma 3.2:

\[
\frac{1}{a_n} \int_{t-a_n}^{t+a_n} (1 - F(x)) L_{n,1}(x) dK \left( \frac{t - x}{a_n} \right) = \frac{1}{a_n} \int_{1-\varepsilon_n}^{1} (1 - F(t - a_n u)) \left( L_{n,1}(t - a_n) - L_{n,1}(t) \right) \right) dU
\]

\[
+ \frac{1}{a_n} L_{n,1}(t) \int_{1-\varepsilon_n}^{1} (1 - F(t - a_n u)) \right) dU
\]

\[
= (I_n) + (II_n),
\]

and proceed to bound the two resulting terms.

By integration by parts, we see that the absolute value of the last term is dominated by

\[
|L_{n,1}(t)||f|_\infty. \quad \text{By Theorem 5 in Giné and Guillou [9] (actually by its proof since}
\]

\[
L_{n,1} \text{ there, each treated separately), we then obtain, owing}
\]

\[
to the regularity of \{ \varepsilon_n \} \text{ and to (3.3), that}\sup_{t \leq T_n-a_n} |(II_n)| = O(\sqrt{\log \log n/(a_n \varepsilon_n)}) \quad \text{a.s.}
\]

Hence, since \( a_n \to 0 \) and (3.2) holds,

\[
\sup_{t \leq T_n-a_n} |(II_n)| = O \left( \sqrt{\frac{\log (a_n \wedge \varepsilon_n)^{-1}}{n a_n \varepsilon_n}} \right) \quad \text{a.s.} \quad (3.24)
\]

If in \((I_n)\) we replace \( 1 - F(t - a_n u) \) by \( 1 - F(t + a_n) \), the difference is dominated by

\[
\frac{4}{a_n} \sup_{t \leq T_n} |L_{n,1}(t)||K'|_\infty \sup_{t \in \mathbb{R}} |F(t - a_n u) - F(t + a_n)|
\]

\[
\leq 8 ||f|_\infty||K'|_\infty \sup_{t \leq T_n} |L_{n,1}(t)| = o \left( \sqrt{\frac{\log (a_n \wedge \varepsilon_n)^{-1}}{n a_n \varepsilon_n}} \right) \quad \text{a.s.} \quad (3.25)
\]
as in (3.24). Hence, we need only prove

\[
\frac{1}{a_n} \sup_{t \leq T_n - a_n} \left( 1 - F(t + a_n) \right) \left| L_{n,1}(t - a_n u) - L_{n,1}(t) \right| = O \left( \frac{\log (a_n \wedge \varepsilon_n)^{-1}}{n a_n \varepsilon_n} a.s. \right)
\]

(3.26)

Again, for ease of notation, we restrict to \( u \in [0, 1] \) (as the part of this sup corresponding to \( u \in [-1, 0) \) can be dealt with in the same way). If we define

\[
W_n(t, u) := \frac{1}{a_n} \sum_{i=1}^{n} I_{X_i \in [t - a_n u, t]} I_{X_i \leq Y_i} \left( 1 - F(t + a_n) \right) \left( 1 - F(X_i) \right) \left( 1 - G(X_i) \right),
\]

then the left hand side of (3.26), with the restriction to \( u \in [0, 1] \), just becomes

\[
\sup_{t \leq T_n - a_n, 0 \leq u \leq 1} \left| W_n(t, u) - EW_n(t, u) \right| = O \left( \sqrt{\log (a_n \wedge \varepsilon_n)^{-1}} - \frac{1}{n a_n \varepsilon_n} \right).
\]

The corresponding class \( \mathcal{F} \) is bounded measurable \( VC \) and we have

\[
\sup_{t \leq T_n - a_n, 0 \leq u \leq 1} \frac{1}{a_n} \frac{1}{n} \sum_{i=1}^{n} I_{X_i \in [t - a_n u, t]} I_{X_i \leq Y_i} \left( 1 - F(t + a_n) \right) \left( 1 - F(X_i) \right) \left( 1 - G(X_i) \right) \leq \frac{1}{n a_n (1 - G(T_n))} \leq \frac{1}{n a_n \varepsilon_n}
\]

and

\[
\frac{1}{a_n^2 n^2} E \left[ \frac{I_{X \in [t - a_n u, t]} I_{X \leq Y} \left( 1 - F(t + a_n) \right)}{\left( 1 - F(X) \right) \left( 1 - G(X) \right)} \right]^2 \leq \frac{1}{a_n^2 n^2} E \left[ \frac{I_{X \in [t - a_n u, t]} I_{X \leq Y}}{1 - G(X)} \right] \leq \| f \|_{\infty} \frac{1}{n^2 a_n \varepsilon_n}
\]

Hence the parameters \( U_n \) and \( \sigma_n^2 \) can be taken to be \( U_n = 1/n a_n \varepsilon_n, \sigma_n^2 = \| f \|_{\infty} (1/n^2 a_n \varepsilon_n) \), and they satisfy inequality (2.10). Then we can apply inequality (2.12) and proceed as in the last part of the proof of Lemma 2.3 (see also (3.20) for more details on how to apply Montgomery-Smith’s maximal inequality) to obtain

\[
\sup_{t \leq T_n - a_n, 0 \leq u \leq 1} \left| W_n(t, u) - EW_n(t, u) \right| = O \left( \sqrt{\log (a_n \wedge \varepsilon_n)^{-1}} - \frac{1}{n a_n \varepsilon_n} \right).
\]

The same applies to \(-1 \leq u \leq 0\), proving (3.26) and, therefore, the theorem.  \( \square \)

The previous proof and the Kolmogorov 0–1 law show the following:

**Corollary 3.4.** Under the hypotheses of Theorem 3.3, there exists a finite constant \( C \) such that

\[
\limsup_{n \to \infty} \sqrt{\frac{n a_n \varepsilon_n}{\log \left( \frac{1}{a_n \wedge \varepsilon_n} \right)}} \sup_{t \leq T_n - a_n} \left| f_n(t) - \tilde{f}_n(t) \right| = \limsup_{n \to \infty} \sqrt{\frac{n a_n \varepsilon_n}{\log \left( \frac{1}{a_n \wedge \varepsilon_n} \right)}} \sup_{t \leq T_n - a_n} \frac{1}{a_n} \left( 1 - F(t + a_n) \right)
\]


\[
\times \left| \int_{-1}^{1} (L_{n,1}(t - a_n u) - L_{n,1}(t)) K'(u) \, du \right| = C \quad \text{a.s.} \tag{3.27}
\]

Corollary 2 in Diehl and Stute [7] shows that the constant \(C\) is not zero if \(\varepsilon_n\) is a constant independent of \(n\) and \(f\) is bounded away from zero on an interval with right end strictly larger than \(H^{-1}(1 - \varepsilon_n)\). We do not know if \(C \neq 0\) for \(\varepsilon_n \to 0\) as well, although we believe this to be the case in general (or at least if \(1 - G(t)\) is of the same order as \(1 - H(t)\) for large \(t\) and if \(\varepsilon_n\) is eventually larger than \(a_n\)), this remains an open question.

REFERENCES