AN EXPONENTIAL INEQUALITY FOR A WEIGHTED APPROXIMATION TO THE UNIFORM EMPIRICAL PROCESS WITH APPLICATIONS

DAVID M. MASON[1]

University of Delaware

Mason and van Zwet (1987) obtained a refinement to the Komlós, Major, and Tusnády (1975) Brownian bridge approximation to the uniform empirical process. From this they derived a weighted approximation to this process, which has shown itself to have some important applications in large sample theory. We will show that their refinement, in fact, leads to a much stronger result, which should be even more useful than their original weighted approximation. We demonstrate its potential applications through several interesting examples. These include a useful new exponential inequality for Winsorized sums and results on the asymptotic equivalence of two sequences of local experiments.

AMS subject classifications: 60F99, 60F17.
Keywords and phrases: KMT, equivalence of experiments, Winsorized sums.

1 Introduction and statements of main results

Let $U, U_1, U_2, ...,$ be independent uniform $(0,1)$ random variables. For each integer $n \geq 1$ let

$$(1) \qquad G_n(t) = n^{-1} \sum_{i=1}^{n} 1\{U_i \leq t\}, \ -\infty < t < \infty,$$

denote the *empirical distribution function* based on $U_1, ..., U_n$, and

$$(2) \qquad \alpha_n(t) = \sqrt{n}\{G_n(t) - t\}, \ 0 \leq t \leq 1,$$

be the corresponding *uniform empirical process.* Mason and van Zwet (1987) proved the following refinement to the Komlós, Major, and Tusnády [KMT] (1975) Brownian bridge approximation to α_n.

Theorem 1.1 *There exists a probability space $(\Omega, \mathcal{A}, P)$ with independent uniform $(0,1)$ random variables $U_1, U_2, \ldots,$ and a sequence of Brownian bridges $B_1, B_2, ...,$ such that for all $n \geq 1$, $1 \leq d \leq n$ and $x \in \mathbb{R}$*

$$(3) \qquad P\left\{ \sup_{0 \leq t \leq d/n} |\alpha_n(t) - B_n(t)| \geq n^{-1/2}(a \log d + x) \right\} \leq b \exp(-cx)$$

[1]This research was partially supported by NSF Grant No. DMS–9803344.

and

$$(4)\quad P\left\{\sup_{1-d/n\le t\le 1}|\alpha_n(t)-B_n(t)|\ge n^{-1/2}(a\log d+x)\right\}\le b\exp(-cx),$$

where a, b and c are suitable positive constants.

Rio (1994) has obtained specific values for the constants a, b and c. Castelle and Laurent-Bonvalot (1998) have shown that (3) and (4) remain formally valid for $0 < d < 1$. However, in the regions $[0, d/n]$ and $[1-d/n, 1]$, where $0 < d < 1$, it is more appropriate then to approximate the uniform empirical process by a Poisson process than by a Brownian bridge.

Mason and van Zwet (1987) pointed out that their inequality leads to the following useful weighted approximation. For any $0 \le \nu < 1/2$, $n \ge 2$, and $1 \le d \le n-d \le n-1$ let

$$(5)\qquad \Delta_{n,\nu}^{(1)}(d) := \sup_{d/n\le t\le 1}\frac{n^{\nu}|\alpha_n(t)-B_n(t)|}{t^{1/2-\nu}},$$

$$(6)\qquad \Delta_{n,\nu}^{(2)}(d) := \sup_{0\le t\le 1-d/n}\frac{n^{\nu}|\alpha_n(t)-B_n(t)|}{(1-t)^{1/2-\nu}},$$

and

$$(7)\qquad \Delta_{n,\nu}(d) := \sup_{d/n\le t\le 1-d/n}\frac{n^{\nu}|\alpha_n(t)-B_n(t)|}{(t(1-t))^{1/2-\nu}}.$$

On the probability space of Theorem 1.1, one has

$$(8)\qquad \Delta_{n,\nu}(1) = O_p(1),$$

with the same holding with $\Delta_{n,\nu}(1)$ replaced by $\Delta_{n,\nu}^{(1)}(1)$ and $\Delta_{n,\nu}^{(2)}(1)$. Versions of these approximations were proved by M. Csörgő, S. Csörgő, Horváth and Mason [Cs-Cs-H-M] (1986) for the restricted range of $0 \le \nu < 1/4$. The Mason and van Zwet (1987) versions are the best possible in the sense that they are unimprovable with respect to the allowable range of $0 \le \nu < \frac{1}{2}$. These weighted approximations have found numerous and wide ranging applications in probability theory and statistics, see e.g. Part II of the proceedings volume edited by Hahn, Mason and Weiner (1991) and the monograph by M. Csörgő and Horváth (1993), along with the many references therein. The purpose of this paper is to demonstrate that, in fact, Theorem 1.1 readily yields the following much stronger version of (8) and to provide some examples of its potential use. Let $c > 0$ be as in Theorem 1.1.

Theorem 1.2 *On the probability space of Theorem 1.1 for every $0 \le \nu < 1/2$ there exist positive constants A_ν and C_ν such that for all $n \ge 2$, $1 \le d \le n-d \le n-1$ and $0 \le x < \infty$*

$$(9) \qquad P\left\{\Delta_{n,\nu}^{(1)}(d) \geq x\right\} \leq A_\nu \exp(d^{1/2-\nu}C_\nu)\exp(-d^{1/2-\nu}cx/2),$$

$$(10) \qquad P\left\{\Delta_{n,\nu}^{(2)}(d) \geq x\right\} \leq A_\nu \exp(d^{1/2-\nu}C_\nu)\ \exp(-d^{1/2-\nu}cx/2)$$

and

$$(11) \qquad P\{\Delta_{n,\nu}(d) \geq x\} \leq 2A_\nu\ \exp(d^{1/2-\nu}C_\nu)\ \exp(-d^{1/2-\nu}cx/4).$$

For each $n \geq 1$, let $U_{1,n} \leq \ldots \leq U_{n,n}$ denote the order statistics of $U_1, \ldots, U_n$. Introduce the uniform empirical quantile function on $[0,1]$

$$(12) \qquad U_n(t) = U_{k,n},\ (k-1)/n < t \leq k/n, \quad \text{for} \quad k = 1, \ldots, n,$$

and $U_n(0) = U_{1,n}$. Define the uniform quantile process

$$(13) \qquad \beta_n(t) = \sqrt{n}\{t - U_n(t)\}, \quad \text{for} \quad 0 \leq t \leq 1.$$

For any $n \geq 2$ and $0 \leq \nu < 1/4$ set

$$(14) \qquad K_{n,\nu} = \sup_{1/n \leq t \leq 1-1/n} \frac{n^\nu |\alpha_n(t) - \beta_n(t)|}{(t(1-t))^{1/2-\nu}}$$

and

$$(15) \qquad \Gamma_{n,\nu} = \sup_{1/n \leq t \leq 1-1/n} \frac{n^\nu |\beta_n(t) - B_n(t)|}{(t(1-t))^{1/2-\nu}}.$$

Cs-Cs-H-M (1986) (see also Mason (1991)) proved that for any $0 \leq \nu < 1/4$

$$(16) \qquad K_{n,\nu} = O_p(1),$$

Combining this with (7) we see that on the probability space of Theorem 1.1 one also has

$$(17) \qquad \Gamma_{n,\nu} = O_p(1).$$

We should point out here that on the probability space of Cs-Cs-H-M (1986) (17) holds for all $0 \leq \nu < 1/2$, with (8) being valid only for $0 \leq \nu < 1/4$. For completeness, we will provide an exponential inequality for the tail of the random variable $\Gamma_{n,\nu}$. This will be an easy consequence of the following exponential inequality for $K_{n,\nu}$.

Theorem 1.3 *For every $0 \leq \nu < 1/4$ there exist positive constants D_ν and d_ν such that for all $n \geq 2$ and $0 \leq x < \infty$*

$$(18) \qquad P\{K_{n,\nu} \geq x\} \leq D_\nu \exp(-d_\nu x).$$

Combining Theorems 1.2 and 1.3 we immediately conclude the following result.

Theorem 1.4 *On the probability space of Theorem 1.1 for every* $0 \le \nu < 1/4$ *there exist positive constants* E_ν *and* e_ν *such that for all* $n \ge 2$ *and* $0 \le x < \infty$

$$P\{\Gamma_{n,\nu} \ge x\} \le E_\nu \exp(-e_\nu x). \tag{19}$$

Remark 1.1 A dual to Theorem 1.2 exists for the uniform quantile process β_n. Inequalities (9), (10) and (11) hold on the probability space of Cs-Cs-H-M (1986), when α_n is replaced by β_n, with possibly different constants. The proof goes exactly like that of Theorem 1.2. However, at the step when one previously applied Theorem 1.1 in the proof of Theorem 1.2, one now makes use of Theorem 3.2.3 of M. Csörgő and Horváth (1993). The author is thankful to Sándor Csörgő for pointing this out to him.

2 Examples of how the inequality can be used

2.1 An exponential inequality for winsorized sums

Let $X, X_1, X_2, ...,$ be a sequence of i.i.d. nondegenerate random variables with common distribution function F with left continous inverse function Q. Choose $0 < a < 1 - b < 1$ and $n \ge 1$, and consider the Winsorized sum

$$W_n(a,b) :=$$

$$\sum_{i=1}^n \Big[\quad Q(a)1\{X_i \le Q(a)\} + X_i 1\{Q(a) < X_i \le Q(1-b)\}$$

$$+Q(1-b)1\{X_i > Q(1-b)\}\Big].$$

Now due to the fact that

$$(X_i)_{i\ge 1} \stackrel{d}{=} (Q(U_i))_{i\ge 1},$$

one sees after integrating by parts that

$$n^{-1/2}\{W_n(a,b) - EW_n(a,b)\} \stackrel{d}{=} -\int_a^{1-b} \alpha_n(s)dQ(s).$$

Set

$$\sigma^2(a,b) = \int_a^{1-b}\int_a^{1-b} (s \wedge t - st)dQ(s)dQ(t) = \text{Var } W_1(a,b).$$

It is known (cf. S. Csörgő, Haeusler and Mason (1988)) that for any two sequences a_n and b_n of positive constants such that $0 < a_n < 1/2 < 1 - b_n < 1$ for $n \ge 1$, and

$$a_n \to 0,\ na_n \to \infty,\ \ b_n \to 0 \text{ and } nb_n \to \infty, \tag{20}$$

as $n \to \infty$, that the sequence of random variables

$$(21) \qquad Z_n(a_n, b_n) := \int_{a_n}^{1-b_n} \alpha_n(s) dQ(s)/\sigma(a_n, b_n) \xrightarrow{d} Z,$$

as $n \to \infty$, where Z is a standard normal random variable. To see how this goes, note that on the probability space of Theorem 1.1

$$Z_n := \int_{a_n}^{1-b_n} B_n(s) dQ(s)/\sigma(a_n, b_n) \stackrel{d}{=} Z,$$

and

$$\begin{aligned} &|Z_n(a_n, b_n) - Z_n| \\ &\leq \int_{a_n}^{1/2} |\alpha_n(s) - B_n(s)| dQ(s)/\sigma(a_n, 1/2) \\ &\quad + \int_{1/2}^{1-b_n} |\alpha_n(s) - B_n(s)| dQ(s)/\sigma(1/2, b_n), \end{aligned}$$

which for any $0 < \nu < 1/2$ is

$$(22) \qquad \begin{aligned} \leq n^{-\nu} \{ \quad &\Delta_{n,\nu}(na_n) \textstyle\int_{a_n}^{1/2} (s(1-s))^{1/2-\nu} dQ(s)/\sigma(a_n, 1/2) \\ &+\Delta_{n,\nu}(nb_n) \textstyle\int_{1/2}^{1-b_n} (s(1-s))^{1/2-\nu} dQ(s)/\sigma(1/2, b_n) \}. \end{aligned}$$

Using the fact (e.g. Inequality 2.1 of Shorack (1997)) that for any $0 < c < 1 - d < 1$

$$(23) \qquad \int_c^{1-d} (s(1-s))^{1/2-\nu} dQ(s)/\sigma(c, d) \leq (3/\sqrt{\nu})(c \wedge d)^{-\nu},$$

we see that the bound in (22) is

$$(24) \qquad \leq (3/\sqrt{\nu})(na_n)^{-\nu} \Delta_{n,\nu}(na_n) + (3/\sqrt{\nu})(nb_n)^{-\nu} \Delta_{n,\nu}(nb_n).$$

Clearly from (24) and (11) we readily obtain that for any $\delta > 0$

$$(25) \qquad \begin{aligned} &P\{|Z_n(a_n, b_n) - Z_n| > \delta\} \\ &\leq 2A_\nu \ \exp((na_n)^{1/2-\nu} C_\nu) \exp(-(na_n)^{1/2} c\sqrt{\nu}\delta/24) \\ &+2A_\nu \ \exp((nb_n)^{1/2-\nu} C_\nu) \ \exp(-(nb_n)^{1/2} c\sqrt{\nu}\delta/24) \\ &=: P_n(a_n, b_n, \delta). \end{aligned}$$

This immediately yields the following uniform bounds on the distribution function of $Z_n(a_n, b_n)$.

Proposition 2.1. *Let a_n and b_n be sequences of positive constants such that for $n \geq 2$, $0 < a_n < 1/2 < 1 - b_n < 1$, $1 \leq na_n < n$ and $1 \leq nb_n < n$. Then for any $\delta > 0, n \geq 2$ and $z \in \mathbb{R}$*

$$\begin{aligned} P\{Z \leq z - \delta\} - P_n(a_n, b_n, \delta) \leq \; & P\{Z_n(a_n, b_n) \leq z\} \\ & \leq P\{Z \leq z + \delta\} + P_n(a_n, b_n, \delta). \end{aligned} \tag{26}$$

Notice that by choosing $\delta = \delta_n = C[\max(na_n, nb_n)]^{-\frac{1}{2}+\varepsilon}$ for a suitable constant $C > 0$ and any small $\varepsilon > 0$, to make $P_n(a_n, b_n, \delta_n) \approx \delta_n$ in (26), one easily obtains a bound on the Lévy distance between the distribution functions of $[W_n(a_n, b_n) - E(W_n(a_n, b_n))]/\sigma(a_n, b_n)$ and the standard normal distribution function. This yields a rate of convergence $O([\max(na_n, nb_n)]^{-\frac{1}{2}+\varepsilon})$ for Winsorized sums under no distributional assumptions. (The author thanks an anonymous referee for this observation.)

Let $1 \leq k_n < n$, $n \geq 3$, be a sequence of integers such that $k_n \sim na_n$ for some sequence $0 < a_n < 1 - a_n < 1$, $n \geq 1$, of positive constants satisfying as $n \to \infty$,

$$a_n \searrow 0, \; na_n \nearrow \text{ and } na_n/\log\log n \to \infty. \tag{27}$$

Haeusler and Mason (1987) showed that if F is in the domain of attraction of a stable law of index $0 < \alpha \leq 2$ then with probability 1

$$\limsup_{n\to\infty} \frac{\pm\left\{\sum_{i=k_n+1}^{n-k_n} X_{i,n} - n\int_{a_n}^{1-a_n} Q(s)ds\right\}}{\sigma(a_n, a_n)\sqrt{2n\log\log n}} = 1, \tag{28}$$

where $X_{1,n} \leq ... \leq X_{n,n}$ denote the order statistics of $X_1, ..., X_n$. The crux of their proof of (28) was to establish that

$$\limsup_{n\to\infty} \pm Z_n(a_n, a_n)/\sqrt{2\log\log n} = 1. \tag{29}$$

An essential step in the argument leading to (29) was to obtain inequalities like the following: For all $0 < \varepsilon < \sqrt{2}$,

$$\begin{aligned} & P\left\{Z_n(a_n, a_n) > (\sqrt{2} + \varepsilon)\sqrt{\log\log n}\right\} \\ & \leq P\left\{Z > (\sqrt{2} + \frac{\varepsilon}{2})\sqrt{\log\log n}\right\}(1 + o(1)) \end{aligned} \tag{30}$$

and

$$\begin{aligned} & P\left\{Z_n(a_n, a_n) > (\sqrt{2} - \varepsilon)\sqrt{\log\log n}\right\} \\ & \geq P\left\{Z > (\sqrt{2} - \frac{\varepsilon}{2})\sqrt{\log\log n}\right\}(1 + o(1)). \end{aligned} \tag{31}$$

Proposition 2.1 gives these inequalities immediately after taking into account the assumption that $na_n/\log\log n \to \infty$.

2.2 A moment bound for the weighted approximation

Clearly Theorem 1.2 yields immediately the following exponential moment result.

Proposition 2.2. *On the probability space of Theorem 1.1 for all* $0 \le \nu < 1/2$ *there exists a* $\gamma > 0$ *such that*

$$\sup_{n \ge 2} E \exp(\gamma \Delta_{n,\nu}(1)) < \infty, \tag{32}$$

with the same statement holding with $\Delta_{n,\nu}(1)$ *replaced by* $\Delta_{n,\nu}^{(1)}(1)$ *or* $\Delta_{n,\nu}^{(2)}(1)$.

Now for each integer $n \ge 2$ let $\mathcal{R}_n$ denote a class of nondecreasing left continuous functions r on $[1/n, 1-1/n]$. Assume there exists a sequence of positive constants c_n such that for some $0 \le \nu < 1/2$

$$\sup_{n \ge 2} \sup_{r \in \mathcal{R}_n} c_n^{-1} \int_{1/n}^{1-1/n} (s(1-s))^{1/2-\nu} dr(s) =: M < \infty. \tag{33}$$

From Proposition 2.2 we obtain

Proposition 2.3. *Let* $\{\mathcal{R}_n, n \ge 2\}$, *denote a sequence of classes of nondecreasing left continuous functions on* $[1/n, 1-1/n]$ *satisfying (33) for some* $0 \le \nu < 1/2$. *On the probability space of Theorem 1.1 there exists a* $\gamma > 0$ *such that*

$$\sup_{n \ge 2} E \exp(\gamma n^{\nu} I_n) < \infty, \tag{34}$$

where

$$I_n := \sup_{r \in \mathcal{R}_n} c_n^{-1} \int_{1/n}^{1-1/n} |\alpha_n(s) - B_n(s)| dr(s). \tag{35}$$

Proposition 2.3 follows trivially from Proposition 2.2 by observing that

$$I_n \le \Delta_{n,\nu}(1) M.$$

Moment bound results like (34) are useful in the study of central limit theorems for the Wasserstein distance between the empirical and the true distribution. Consult Barrio, Giné and Matrán (1999) for details, where they point out that they could have used our results in their analysis instead of a difficult inequality of Talagrand. We will soon see that they come in handy to obtain bounds on the deficiency distance between an experiment and its Gaussian approximation.

2.3 The local asymptotic equivalence of experiments

This example is motivated by the work of Nussbaum (1996) and we will use much of his basic setup.

Let $\mathcal{F}$ denote a class of densities on $\mathbb{R}$ with a common support. Fix an $f_0 \in \mathcal{F}$ and for any $f \in \mathcal{F}$ write the log ratio

$$\phi_{0,f} = \log(f(F_0^{-1})/f_0(F_0^{-1})), \tag{36}$$

where F_0^{-1} is the left continuous inverse of the distribution function F_0 of f_0 defined on $(0,1)$ and $0/0 := 1$. Introduce for each $n \geq 1$ the likelihood processes

$$\Lambda_{0,n}(f, f_0) = \exp(-n^{1/2} \int_0^1 \alpha_n(s) d\phi_{0,f}(s) + nE\phi_{0,f}(U)), \tag{37}$$

and

$$\Lambda_{1,n}(f, f_0) = \exp(-n^{1/2} \int_0^1 B_n(s) d\phi_{0,f}(s) - \frac{n\mathrm{Var}\phi_{0,f}(U)}{2}). \tag{38}$$

We call $\Lambda_{0,n}(f, f_0)$ a likelihood process since after integrating by parts we have

$$\begin{aligned} \Lambda_{0,n}(f, f_0) &= \exp(\sum_{i=1}^n \phi_{0,f}(U_i)) \\ &= \prod_{i=1}^n [f(F_0^{-1}(U_i))/f_0(F_0^{-1}(U_i))] \\ &\stackrel{d}{=} \prod_{i=1}^n (f(X_i)/f_0(X_i)), \end{aligned} \tag{39}$$

where $X_1, ..., X_n$ are i.i.d. with density f_0. Integrating by parts we also see that,

$$\Lambda_{1,n}(f, f_0) = \exp(n^{1/2} \int_0^1 \phi_{0,f}(s) dB_n(s) - \frac{n\mathrm{Var}\phi_{0,f}(U)}{2}) \tag{40}$$

is the likelihood process corresponding to n independent observations of the process

$$y(t) = \int_0^t \{\phi_{0,f}(s) - E\phi_{0,f}(U)\} ds + n^{-1/2} B_n(t), \ 0 \leq t \leq 1. \tag{41}$$

In fact, if one lets $Q_{0,f}^{(n)}$ and P_0 denote, respectively, the distribution induced by the process

$$Z_n(t) = n \int_0^t \{\phi_{0,f}(s) - E\phi_{0,f}(U)\} ds + n^{1/2} B_n(t), \ 0 \leq t \leq 1, \tag{42}$$

and by the Brownian bridge B_n, on $\mathcal{C}[0,1]$, then by applying the results of Hájek (1960), one obtains that

(43) $$\Lambda_{1,n}(f, f_0) = \frac{dQ_{0,f}^{(n)}}{dP_0}.$$

We introduce the following conditions and notation. Let K be a nondecreasing left continuous function on $(0,1)$ such that for some $p > 2$ and $\kappa < \infty$

(44) $$\int_0^1 (s(1-s))^{1/p} dK(s) =: \kappa < \infty.$$

For each $n \geq 1$ let $\mathcal{H}_n$ be a class of functions on $(0,1)$ such that each $h \in \mathcal{H}_n$ can be decomposed into the difference

(45) $$h = h_1 - h_2,$$

where h_1 and h_2 are nondecreasing left continuous functions on $(0,1)$ satisfying for all $0 < a \leq 1/2$

(46) $$\sup_{h\in\mathcal{H}_n} \int_0^a (s(1-s))^{1/p} d[h_1(s) + h_2(s)] \leq \int_0^a (s(1-s))^{1/p} dK(s)$$

and

(47) $$\sup_{h\in\mathcal{H}_n} \int_{1-a}^1 (s(1-s))^{1/p} d[h_1(s) + h_2(s)] \leq \int_{1-a}^1 (s(1-s))^{1/p} dK(s).$$

For each $n \geq 1$, let $\mathcal{F}_{0,n}$ denote the subclass of $\mathcal{F}$ such that $f_0 \in \mathcal{F}_{0,n}$ and for each $f \in \mathcal{F}_{0,n}$

(48) $$\phi_{0,f} = \gamma_n h,$$

where $h \in \mathcal{H}_n$ and

(49) $$\gamma_n = o(1).$$

Further assume that as $n \to \infty$

(50) $$\sup_{f\in\mathcal{F}_{0,n}} n|E\phi_{0,f}(U) + \text{Var}\phi_{0,f}(U)/2| \to 0$$

and for all large n for some $\eta > 0$

(51) $$\sup_{f\in\mathcal{F}_{0,n}} \text{Var}\phi_{0,f}(U) \leq \eta\gamma_n^2.$$

Define the following two sequences of local experiments around f_0 :

(52) $$(E_{0,n}(f_0))_{n\geq 1} = \left([0,1]^n, \mathcal{B}_{[0,1]}^n, (P_f^{\otimes n},\ f \in \mathcal{F}_{0,n})\right)_{n\geq 1},$$

where P_f is the measure induced on $[0,1]$ by the density $f(F_0^{-1})/f_0(F_0^{-1})$ and $P_f^{\otimes n} = P_f \times \cdots \times P_f$ is the corresponding product measure on $[0,1]^n$; and

$$(53) \qquad (E_{1,n}(f_0))_{n\geq 1} = \left(\mathcal{C}[0,1], \mathcal{B}_{\mathcal{C}[0,1]}, (Q_{0,f}^{(n)},\ f \in \mathcal{F}_{0,n})\right)_{n\geq 1}.$$

For any two experiments E_0 and E_1 let $\Delta(E_0, E_1)$ denote the deficiency distance between these two experiments. Refer to Le Cam and Yang (1990) for the definition of this distance.

Proposition 2.4. *Let $\gamma_n = n^{-1/2}$. Then, with the above assumptions and notation, the two sequences of experiments $(E_{0,n}(f_0))_{n\geq 1}$ and $(E_{1,n}(f_0))_{n\geq 1}$ are asymptotically equivalent, meaning that as $n \to \infty$*

$$(54) \qquad \Delta\left(E_{0,n}(f_0), E_{1,n}(f_0)\right) \to 0.$$

Now assume that there exists a sequence of classes $\mathcal{H}_n, n \geq 1$, of functions on $(0,1)$ such that each $h \in \mathcal{H}_n$ can be written $h = h_1 - h_2$, where h_1 and h_2 are nondecreasing left continuous functions satisfying for some finite positive constant κ

$$(55) \qquad \sup_{h\in\mathcal{H}_n} \int_0^1 d[h_1(s) + h_2(s)] \leq \kappa$$

and for each $f \in \mathcal{F}_{0,n}$ the representation (48) holds with a γ_n satisfying

$$(56) \qquad \gamma_n = o(n^{-1/3}).$$

Furthermore, assume (50) and (51) hold. We will see that a small modification of the proof of Proposition 2.4 leads to the following result closely related to the work of Nussbaum (1996).

Proposition 2.5. *Under the modified assumptions and notations just described (54) holds.*

If one assumes that $\mathcal{F}_{0,n}$, $n \geq 1$, is a sequence of classes of densities with $f_0 \in \mathcal{F}_{0,n}$ for each $n \geq 1$, such that for some sequence of positive constants $\gamma_n, n \geq 1$, converging to 0

$$(57) \qquad \sup_{f\in\mathcal{F}_{0,n}} \sup_{s\in(0,1)} \left|\frac{f}{f_0}(F_0^{-1}(s)) - 1\right| \leq \gamma_n,$$

then by using the fact that as $|u| \searrow 0$,

$$\varphi(u) = \log(1+u) - u + (\log(1+u))^2/2 = O(u^3)$$

and

$$(\log(1+u))^2 - u^2 = O(u^3),$$

one readily shows that

$$\sup_{f\in\mathcal{F}_{0,n}} E\varphi(\frac{f}{f_0}(F_0^{-1}(U)) - 1)$$

$$= \sup_{f\in\mathcal{F}_{0,n}} \{E\phi_{0,f}(U) + \frac{E\phi_{0,f}^2(U)}{2}\} = O(\gamma_n^3) \tag{58}$$

and

$$\sup_{f\in\mathcal{F}_{0,n}} E\phi_{0,f}^2(U) \le \gamma_n^2 + O(\gamma_n^3). \tag{59}$$

Now (58) and (59) imply that

$$\sup_{f\in\mathcal{F}_{0,n}} (E\phi_{0,f}(U))^2 = O(\gamma_n^4).$$

Thus we see that condition (50) holds for any γ_n satisfying $\gamma_n = o(n^{-1/3})$. Furthermore, we have for all large n

$$\sup_{f\in\mathcal{F}_{0,n}} \mathrm{Var}\phi_{0,f}(U) \le 2\gamma_n^2. \tag{60}$$

Moreover, if for some $\varepsilon > 0$

$$f_0 \ge \varepsilon \tag{61}$$

and some $A > 0$, uniformly for $s, t \in (0,1)$, $f \in \mathcal{F}_{0,n}$ and $n \ge 1$,

$$|f(s) - f(t)| \le A|s - t|, \tag{62}$$

then it is easily verified that (48) and (55) are satisfied. Therefore by Proposition 2.5 conclusion (54) holds. This is in correspondence with the remarks in the paragraph following Proposition 2.3 of Nussbaum (1996).

3 Proof of Theorems 1.2 and 1.3

3.1 Proof of Theorem 1.2

First consider (9). For any $1 \le i < i+1 \le n$ write

$$\delta_{i,n} = P\left\{\sup_{i/n\le t\le(i+1)/n} \frac{n^{\nu}|\alpha_n(t) - B_n(t)|}{t^{1/2-\nu}} \ge x\right\}.$$

Set $x = 2a_\nu + z$, where a_ν satisfies

$$a_\nu i^{1/2-\nu} > a\log(i+1) \text{ for all } i \ge 1$$

and the constant a is as in (3). We get then that

$$\delta_{i,n} \le P\left\{ \sup_{0\le t\le (i+1)/n} |\alpha_n(t) - B_n(t)| \ge n^{-1/2} i^{1/2-\nu} x \right\}$$

$$\le P\{ \sup_{0\le t\le (i+1)/n} |\alpha_n(t) - B_n(t)|$$

$$\ge n^{-1/2}(a\log(i+1) + i^{1/2-\nu}a_\nu + i^{1/2-\nu}z) \},$$

which by (3) is

$$\le b\exp(-i^{1/2-\nu}a_\nu c)\exp(-i^{1/2-\nu}cz).$$

We see then that for any $1 \le d < n$

$$P\left\{\Delta_{n,\nu}^{(1)}(d) \ge x\right\} \le \sum_{i=[d]}^{n-1} \delta_{i,n} \le b\sum_{i=[d]}^{\infty}\{\exp(-i^{1/2-\nu}a_\nu c)\exp(-i^{1/2-\nu}cz)\}$$

$$\le A_\nu \exp(-d^{1/2-\nu}cz/2) = A_\nu \ \exp(d^{1/2-\nu}C_\nu)\exp(-d^{1/2-\nu}cx/2),$$

where

$$A_\nu = b\sum_{i=1}^{\infty}\exp(-i^{1/2-\nu}a_\nu c) \text{ and } C_\nu = a_\nu c.$$

This proves inequality (9). Inequality (10) follows in the same way and inequality (11) is an immediate consequence of (9) and (10).□

3.2 Proof of Theorem 1.3

For any $n \ge 2$ and $0 \le \nu < 1/4$ set

$$K_{n,\nu}^{(1)} = \sup_{1/n\le t\le 1} \frac{n^\nu|\alpha_n(t) - \beta_n(t)|}{t^{1/2-\nu}}$$

and

$$K_{n,\nu}^{(2)} = \sup_{0\le t\le 1-1/n} \frac{n^\nu|\alpha_n(t) - \beta_n(t)|}{(1-t)^{1/2-\nu}}.$$

We shall first show

Proposition 3.1. *For every $0 \le \nu < 1/4$ there exist positive constants d_ν and k_ν such that for all $n \ge 2$ and $0 \le x < \infty$*

$$P\left\{K_{n,\nu}^{(1)} \ge x\right\} \le d_\nu \exp(-k_\nu x), \tag{63}$$

with the same inequality holding for $K_{n,\nu}^{(2)}$.

Before we can establish this we need to gather some facts.

For any $a > 0, 0 \le b < c \le 1$ and integer $n \ge 1$ set

$$\omega_n(a,b,c) = \sup\{|\alpha_n(s+h) - \alpha_n(s)| : 0 \le s+h \le 1, 0 \le |h| \le a,\ b \le s \le c\}.$$

The following inequality is stated in Mason (1991). Its proof is essentially contained in that of Inequality 1 of Mason, Shorack and Wellner (1983). Refer also to Inequality 1 of Einmahl and Mason (1988) where the ba^{-1} should be replaced by $(ba^{-1}) \vee 1$.

Fact 3.1. *For universal positive constants A and B for all $0 < a \le 1/2$, $0 \le b < c \le 1$, $n \ge 1$ and $\lambda > 0$*

$$(64) \qquad P\{\omega_n(a,b,c) > \lambda\sqrt{a}\} \le \{(c-b)a^{-1}\} A \exp(-B\lambda^2 \psi(\lambda/\sqrt{na})),$$

where for $x \ge 0$

$$(65) \qquad \psi(x) = 2x^{-2}\{(x+1)\log(x+1) - x\}.$$

For future reference we record the fact that for $x \ge 0$

$$(66) \qquad \psi(x) \downarrow \text{ as } x \uparrow .$$

For any integer $n \ge 1$ and $0 \le p \le 1$ let $B(n,p)$ denote a binomial random variable with parameters n and p. We will need the following special case of Bernstein's inequality (eg. Pollard (1984) or Shorack and Wellner (1986)).

Fact 3.2. *For any integer $n \ge 1$, $0 \le p \le 1$ and $x \ge p$*

$$(67) \qquad P\{B(n,p) \ge nx\}$$
$$\le \exp\left(\frac{-n(x-p)^2/2}{p(1-p) + (p \vee (1-p))(x-p)/3}\right).$$

We will also need the Dvorestzky, Kiefer and Wolfowitz (1956) inequality. See also Massart (1990) for the best possible constant.

Fact 3.3. *For any integer $n \ge 2$ and $x \ge 0$*

$$(68) \qquad P\{||\alpha_n|| > x\} \le 4\exp\left(-2x^2\right),$$

where

$$||\alpha_n|| = \sup_{0 \le t \le 1} |\alpha_n(t)|.$$

Choose $1/4 > \delta > \nu \ge 0$ and $\tau \ge 0$. For any $n \ge 1$ and $1 \le i \le n-1$, define

$$(69) \qquad \Delta_n(i,\tau) = \omega_n\left(\frac{\tau\, i^{1-2\delta}}{n}, \frac{i}{n}, \frac{i+1}{n}\right).$$

Lemma 3.1. *For universal positive constants A_1 and c_1 for all $\tau \geq 0$*

$$(70)\qquad P\left\{\max_{1\leq i\leq n-1} n^{\nu}\Delta_n(i,\tau)/(i/n)^{1/2-\nu} > \tau\right\} \leq A_1\exp(-c_1\tau).$$

Proof. First choose $\tau \geq 1$. We shall consider two cases.
Case 1. First assume $\tau\, i^{1-2\delta}/n \leq 1/2$. In this case, by Fact 3.1 we have

$$(71)\qquad P\{\Delta_n(i,\tau) > \tau n^{-1/2} i^{1/2-\nu}\} \leq A\exp(-Bi^{2(\delta-\nu)}\tau\psi(1)).$$

Case 2. Now assume $\tau\, i^{1-2\delta}/n > 1/2$. In this situation, by noting that

$$\Delta_n(i,\tau) \leq 2||\alpha_n||,$$

we get

$$P\{\Delta_n(i,\tau) > \tau n^{-1/2} i^{1/2-\nu}\} \leq P\{||\alpha_n|| > 2^{-1}\tau n^{-1/2} i^{1/2-\nu}\}$$

$$\leq 4\exp(-\frac{\tau^2}{2} n^{-1} i^{1-2\nu}) \leq 4\exp(-\frac{\tau}{4} i^{2\delta-2\nu}).$$

Clearly then with $\rho = 2\delta - 2\nu, \widetilde{A}_1 = \max\{A, 4\}$ and $c_1 = \min\{B\phi(1), 4^{-1}\}$ we have for $n \geq 2$, $1 \leq i \leq n-1$ and all $\tau \geq 1$

$$(72)\qquad P\{\Delta_n(i,\tau) > \tau n^{-1/2} i^{1/2-\nu}\} \leq \widetilde{A}_1\exp(-c_1 i^{\rho}\tau).$$

Therefore for all $\tau \geq 1$

$$P\Big\{\max_{1\leq i\leq n-1} n^{\nu}\Delta_n(i,\tau)/(i/n)^{1/2-\nu} > \tau\Big\} \leq \widetilde{A}_1\sum_{i=1}^{\infty}\exp(-c_1 i^{\rho}\tau)$$

$$\leq \widetilde{A}_1\exp(-c_1\tau)\sum_{i=1}^{\infty}\exp(-c_1(i^{\rho}-1)) =: \overline{A}_1\exp(-c_1\tau).$$

Now, by setting $A_1 = \max\{\overline{A}_1, \exp(c_1)\}$, we see that (70) holds for all $\tau \geq 0$. □

Set

$$(73)\ M_n(\delta) = \max_{2\leq i\leq n}\ \sup_{(i-1)/n<t\leq i/n} \frac{n^{2\delta}|U_n(t)-t|}{((i-1)/n)^{1-2\delta}} \vee \frac{n^{2\delta}|U_n(1/n)-1/n|}{(1/n)^{1-2\delta}}.$$

Lemma 3.2. *For a universal positive constant A_2 for all $\tau \geq 0$*

$$(74)\qquad P\{M_n(\delta) \geq \tau\} \leq A_2\exp(-12^{-1}\tau).$$

Proof. To begin, notice that for any $2 \leq i \leq n$

$$\sup_{(i-1)/n<t\leq i/n} \frac{n^{2\delta}|U_n(t)-t|}{((i-1)/n)^{1-2\delta}} \leq \frac{n^{2\delta}|U_{i,n}-\frac{i}{n}|}{((i-1)/n)^{1-2\delta}} + \left(\frac{1}{i-1}\right)^{1-2\delta}$$

$$\leq \frac{n^{2\delta}|U_{i,n} - \frac{i}{n}|}{(i/n)^{1-2\delta}} 2^{1-2\delta} + 1.$$

Using this string of inequalities we get that for any $2 < i \leq n$ and $z \geq 1$

$$(75)\ P\left\{\sup_{(i-1)/n<t\leq i/n} \frac{n^{2\delta}|U_n(t) - t|}{((i-1)/n)^{1-2\delta}} \geq 3z\right\} \leq P\left\{\frac{n^{2\delta}|U_{i,n} - \frac{i}{n}|}{(i/n)^{1-2\delta}} \geq z\right\}.$$

Now for any $2 \leq i \leq n$

$$P\left\{\frac{n^{2\delta}|U_{i,n} - \frac{i}{n}|}{(i/n)^{1-2\delta}} \geq z\right\}$$

$$(76) \qquad = P\{U_{i,n} \geq \frac{i}{n} + \frac{zi^{1-2\delta}}{n}\} + P\{U_{i,n} \leq \frac{i}{n} - \frac{zi^{1-2\delta}}{n}\}.$$

We will first show that for all $z \geq 1$

$$(77) \qquad P\{U_{i,n} \geq \frac{i}{n} + \frac{zi^{1-2\delta}}{n}\} \leq 2\exp(-6^{-1}zi^{1-4\delta}).$$

First assume $0 < 1 - \frac{i}{n} - \frac{zi^{1-2\delta}}{n} \leq 1$. Clearly,

$$P\{U_{i,n} \geq \frac{i}{n} + \frac{zi^{1-2\delta}}{n}\} \leq P\{B(n, \frac{i}{n} + \frac{zi^{1-2\delta}}{n}) \leq i\}$$

$$= P\{B(n, 1 - \frac{i}{n} - \frac{zi^{1-2\delta}}{n}) \geq n - i\}.$$

Applying Fact 3.2 we obtain after a little analysis the bound

$$P\Big\{B(n, 1 - \frac{i}{n} - \frac{zi^{1-2\delta}}{n}) \geq n - i\Big\} \leq \exp(-z^2 i^{2-4\delta}/(2i + 4zi^{1-2\delta}))$$

$$\leq \exp(-6^{-1}z^2 i^{1-4\delta}) + \exp(-6^{-1}zi^{1-2\delta}) \leq 2\exp(-6^{-1}zi^{1-4\delta}).$$

Thus (77) holds in this case. Since (77) is trivial when $1 - \frac{i}{n} - \frac{zi^{1-2\delta}}{n} \leq 0$, we conclude its validity for all $z \geq 1$. Thus we conclude (77).

Next observe that

$$(78) \qquad P\Big\{U_{i,n} \leq \frac{i}{n} - \frac{zi^{1-2\delta}}{n}\Big\} = P\Big\{B(n, \frac{i}{n} - \frac{zi^{1-2\delta}}{n}) \geq i\Big\},$$

which by an application of Fact 3.2, for all $z \geq 1$ such that $\frac{i}{n} - \frac{zi^{1-2\delta}}{n} > 0$, is

$$(79) \qquad \leq \exp(-z^2 i^{2-4\delta}/(2i - 4zi^{1-2\delta}/3)) \leq \exp(-2^{-1}zi^{1-4\delta}).$$

Note that this inequality holds trivially whenever $\frac{i}{n} - \frac{zi^{1-2\delta}}{n} \le 0$ and $z \ge 1$. Combining (76), (77), (78) and (79) we get

$$P\left\{\frac{n^{2\delta}|U_{i,n} - \frac{i}{n}|}{(i/n)^{1-2\delta}} \ge z\right\} \le 3\exp(-6^{-1}zi^{1-4\delta}). \tag{80}$$

This bound in conjunction with (75) yields for all $z \ge 1$ and $i \ge 2$

$$P\left\{\max_{2\le i\le n} \sup_{(i-1)/n<t\le i/n} \frac{n^{2\delta}|U_n(t) - t|}{((i-1)/n)^{1-2\delta}} \ge 3z\right\} \le 3\sum_{i=1}^{\infty} \exp(-6^{-1}zi^{1-4\delta}).$$

Notice that (80) also holds when $i = 1$. Thus

$$P\{M_n(\delta) \ge 3z\} \le 6\sum_{i=1}^{\infty} \exp(-6^{-1}zi^{1-4\delta})$$

$$\le 6\exp(-6^{-1}z)\sum_{i=1}^{\infty}\exp(-6^{-1}(i^{1-4\delta} - 1)) =: \overline{A}\exp(-6^{-1}z).$$

Now by changing variables to $\tau = 3z$ and setting $A_2 = \max\{\overline{A}, \exp(6)\}$ we obtain (74). □

We are almost ready to finish the proof of (63). First observe that for any $2 \le i \le n$

$$\sup_{(i-1)/n<t\le i/n} \frac{n^{\nu}|\alpha_n(U_n(t)) - \beta_n(t)|}{t^{1/2-\nu}} \le \frac{1}{(i-1)^{1/2-\nu}} \le 1$$

and $\alpha_n(U_n(1/n)) - \beta_n(1/n) = 0$. Thus for any $z \ge 1$

$$P\left\{K_{n,\nu}^{(1)} \ge 3z\right\} \le P\left\{\sup_{1/n\le t\le 1} \frac{n^{\nu}|\alpha_n(U_n(t)) - \beta_n(t)|}{t^{1/2-\nu}} > 1\right\}$$

$$+P\left\{\sup_{1/n\le t\le 1} \frac{n^{\nu}|\alpha_n(U_n(t)) - \alpha_n(t)|}{t^{1/2-\nu}} \ge z\right\}$$

$$= P\left\{\sup_{1/n\le t\le 1} \frac{n^{\nu}|\alpha_n(U_n(t)) - \alpha_n(t)|}{t^{1/2-\nu}} \ge z\right\},$$

which in turn is (recall (69) and (73))

$$\le P\left\{\max_{1\le i\le n-1} n^{\nu}\Delta_n(i,z)/(i/n)^{1/2-\nu} \ge z\right\} + P\{M_n(\delta) \ge z\}.$$

Applying Lemmas 3.1 and 3.2 we see that this last bound is

$$\le A_1\exp(-c_1 z) + A_2\exp(-12^{-1}z).$$

The rest of the proof of Theorem 1.3 is now straightforward.□

4 Proofs of Propositions 2.4 and 2.5

4.1 Proof of Proposition 2.4

Inequality 4.1. *Let f and g be densities with respect to a $\sigma-$finite measure μ on a measure space $(\Omega, \mathcal{F})$. Assuming f and g have common support, set $D = \log(f/g)/2$, where $0/0 := 1$. For all $\varepsilon > 0$*

$$0 \le 1 - \int_\Omega \sqrt{fg}d\mu \le 1 - e^{-\varepsilon} + \int_\Omega g1\{D < -\varepsilon\}d\mu. \tag{81}$$

Proof. Notice that

$$1 - \int_\Omega \sqrt{fg}d\mu = \int_\Omega (g - \sqrt{fg})d\mu$$

$$\le (1-e^{-\varepsilon}) \int_\Omega g\ 1\{D \ge -\varepsilon\}d\mu - \int_\Omega \sqrt{fg}\ 1\{D < -\varepsilon\}d\mu + \int_\Omega g\ 1\{D < -\varepsilon\}d\mu$$

$$\le 1 - e^{-\varepsilon} + \int_\Omega g\ 1\{D < -\varepsilon\}d\mu.$$

□

Consider the two sequences of experiments

$$(\mathcal{E}_{0,n})_{n\ge1} = (\Omega_{0,n}, \mathcal{A}_{0,n}, (P_{0,n,\theta}, \theta \in \Theta_n))_{n\ge1},$$

and

$$(\mathcal{E}_{1,n})_{n\ge1} = (\Omega_{1,n}, \mathcal{A}_{1,n}, (P_{1,n,\theta}, \theta \in \Theta_n))_{n\ge1}.$$

Lemma 4.1. *Suppose that for each $n \ge 1$ and $\theta \in \Theta_n$, $P_{i,n,\theta}$ is dominated by P_{i,n,θ_0}, $i = 0, 1$, where $\theta_0 \in \Theta_n$ and consider the likelihood processes for $i = 0, 1$*

$$\Lambda_{i,n}(\theta) = dP_{i,n,\theta}/dP_{i,n,\theta_0},\ \theta \in \Theta_n. \tag{82}$$

Assume that for each $n \ge 1$ and $\theta \in \Theta_n$ the processes $\Lambda_{i,n}(\theta)$, $i = 0, 1$, can be defined on the same probability space $((\Omega_{0,n} \times \Omega_{1,n}, \mathcal{A}_{0,n} \times \mathcal{A}_{1,n}), P_n)$, where each $\Lambda_{i,n}(\theta)$, $i = 1, 2$, is a density with respect to P_n such that as $n \to \infty$

$$\inf_{\theta\in\Theta_n} \int\int_{\Omega_{0,n}\times\Omega_{1,n}} \sqrt{\frac{dP_{0,n,\theta}}{dP_{0,n,\theta_0}}} \sqrt{\frac{dP_{1,n,\theta}}{dP_{1,n,\theta_0}}} dP_n \to 1. \tag{83}$$

Then as $n \to \infty$

$$\Delta(\mathcal{E}_{0,n}, \mathcal{E}_{1,n}) \to 0. \tag{84}$$

Proof. According to the remark on page 16 of Le Cam and Yang (1990) to establish (84), it suffices to show that (83) implies that as $n \to \infty$

$$\sup_{\theta\in\Theta_n} ||P_{0,n,\theta} - P_{1,n,\theta}|| = \frac{1}{2} \sup_{\theta\in\Theta_n} E_{P_n}|\Lambda_{0,n}(\theta) - \Lambda_{1,n}(\theta)| \to 0, \tag{85}$$

But this follows from the inequality

$$\frac{1}{2}||P_{0,n,\theta}-P_{1,n,\theta}|| \le \left(1-\left\{\int\int_{\Omega_{0,n}\times\Omega_{1,n}}\sqrt{\frac{dP_{0,n,\theta}}{dP_{0,n,\theta_0}}}\sqrt{\frac{dP_{1,n,\theta}}{dP_{1,n,\theta_0}}}dP_n\right\}^2\right)^{1/2}.$$

□

We are now ready to complete the proof of Proposition 2.4. Assume we are on the probability space of Theorem 1.1. First we will show that as $n \to \infty$

(86) $$\sup_{f\in\mathcal{F}_{0,n}} E|\log\Lambda_{0,n}(f,f_0)-\log\Lambda_{1,n}(f,f_0)| \to 0.$$

Clearly for each $f \in \mathcal{F}_{0,n}$

$$|\log\Lambda_{0,n}(f,f_0)-\log\Lambda_{1,n}(f,f_0)| \le$$

$$n^{1/2}|\int_0^1\{\alpha_n(s)-B_n(s)\}d\phi_{0,f}(s)| + n|E\phi_{0,f}(U)) + \frac{\mathrm{Var}\phi_{0,f}(U))}{2}|.$$

By assumption (50) to finish the proof of (86) it is enough to show that as $n \to \infty$

$$\sup_{f\in\mathcal{F}_{0,n}} n^{1/2}E|\int_0^1\{\alpha_n(s)-B_n(s)\}d\phi_{0,f}(s)| \to 0.$$

Now, in view of (45)-(48),

$$\sup_{f\in\mathcal{F}_{0,n}} n^{1/2}E|\int_{1/n}^{1-1/n}\{\alpha_n(s)-B_n(s)\}d\phi_{0,f}(s)| \le$$

$$n^{1/p}\gamma_n \sup_{h\in\mathcal{H}_n}\int_{1/n}^{1-1/n}(s(1-s))^{1/p}d[h_1(s)+h_2(s)]$$

$$\times E\left\{\sup_{1/n\le s\le 1-1/n}\frac{n^{1/2-1/p}|\alpha_n(s)-B_n(s)|}{(s(1-s))^{1/p}}\right\},$$

which by (7), Proposition 2.2, (44) through (48) and $\gamma_n = n^{-1/2}$ is $o(1)$. Next observe that, with $\gamma_n = n^{-1/2}$,

$$\sup_{f\in\mathcal{F}_{0,n}} n^{1/2}E|\int_0^{1/n}\{\alpha_n(s)-B_n(s)\}d\phi_{0,f}(s)|$$

$$\le 2\sup_{f\in\mathcal{F}_{0,n}} n^{1/2}\gamma_n\int_0^{1-1/n}s^{1/2}d[h_1(s)+h_2(s)]$$

$$\le 2n^{1/p}\gamma_n\int_0^{1-1/n}s^{1/p}dK(s) = o(1).$$

Similarly

$$\sup_{f\in\mathcal{F}_{0,n}} n^{1/2}E\left|\int_{1-1/n}^{1}\{\alpha_n(s)-B_n(s)\}d\phi_{0,f}(s)\right| = o(1).$$

Thus we have established (86).

Set for any $f \in \mathcal{F}_{0,n}$

$$D_n(f) = \{\log\Lambda_{0,n}(f,f_0) - \log\Lambda_{1,n}(f,f_0)\}/2. \tag{87}$$

Choose any $0<\varepsilon<1$. From Inequality 4.1 we get

$$1-\int\int_{[0,1]^n\times\mathcal{C}[0,1]}\sqrt{\Lambda_{0,n}(f,f_0)}\sqrt{\Lambda_{1,n}(f,f_0)}dP_n$$

$$\le 1-e^{-\varepsilon}+\int\int_{[0,1]^n\times\mathcal{C}[0,1]}\Lambda_{1,n}(f,f_0)1\{D_n(f)\le-\varepsilon\}dP_n, \tag{88}$$

where P_n is the probability measure of Theorem 1.1. Applying the Cauchy-Schwarz inequality we get that

$$\int\int_{[0,1]^n\times\mathcal{C}[0,1]}\Lambda_{1,n}(f,f_0)1\{D_n(f)\le-\varepsilon\}dP_n$$

$$\le\left[E\Lambda_{1,n}^2(f,f_0)\right]^{1/2}\left[P_n\{|D_n(f)|\ge\varepsilon\}\right]^{1/2}.$$

Notice that by (51)

$$\left[E\Lambda_{1,n}^2(f,f_0)\right]^{1/2}=\exp\left(\frac{n\mathrm{Var}\phi_{0,f}(U)}{2}\right)\le\exp(\eta/2).$$

Using (86) we get as $n\to\infty$

$$\sup_{f\in\mathcal{F}_{0,n}} P\{|D_n(f)|\ge\varepsilon\}\to 0,$$

which implies that as $n\to\infty$

$$\sup_{f\in\mathcal{F}_{0,n}}\int\int_{[0,1]^n\times\mathcal{C}[0,1]}\Lambda_{1,n}(f,f_0)1\{D_n(f)\le-\varepsilon\}dP_n\to 0.$$

Thus by (88) and the arbitrary choice of ε we infer that $n\to\infty$

$$\inf_{f\in\mathcal{F}_{0,n}}\int\int_{[0,1]^n\times\mathcal{C}[0,1]}\sqrt{\Lambda_{0,n}(f,f_0)}\sqrt{\Lambda_{1,n}(f,f_0)}dP_n\to 1,$$

which by Lemma 4.1 implies (54).□

4.2 Proof of Proposition 2.5

From now on, $\gamma_n = o(n^{-1/3})$ is as in (56), and $D_n(f)$ is as in (87). First notice that by (55) and (50) for any choice of $\varepsilon > 0$ and all large n

$$\sup_{f\in\mathcal{F}_{0,n}} |D_n(f)| \le \kappa\gamma_n \sup_{0\le s\le 1} \sqrt{n}|\alpha_n(s) - B_n(s)| + \varepsilon/2.$$

Thus

$$P\left\{\sup_{f\in\mathcal{F}_{0,n}} |D_n(f)| \ge \varepsilon\right\} \le P\left\{\sup_{0\le s\le 1} \sqrt{n}|\alpha_n(s) - B_n(s)| \ge \gamma_n^{-1}\kappa^{-1}\varepsilon/2\right\},$$

which by Theorem 1.1 applied with $d = n$ is for all large n

$$\le b\exp(-c\gamma_n^{-1}\kappa^{-1}\varepsilon/4) =: b\exp(-\varepsilon d\gamma_n^{-1}).$$

Moveover as before

$$E\Lambda_{1,n}^2(f, f_0) = \exp\left(n\mathrm{Var}\phi_{0,f}(U)\right),$$

which by (51) is for all large n

$$\le \exp(\eta n\gamma_n^2).$$

Thus as in the proof of Proposition 2.4

$$\int\int_{[0,1]^n\times C[0,1]} \Lambda_{1,n}(f, f_0)1\{D_n(f) \le -\varepsilon\}dP_n \le \sqrt{b\exp(-\gamma_n^{-1}(\varepsilon d - \eta n\gamma_n^3))}.$$

This last bound converges to 0 as $n \to 0$ since we assume that $\gamma_n = o(n^{-1/3})$. Hence we conclude as in the proof of Proposition 2.4 that (54) holds.□

Acknowledgements. The author is grateful to Paul Deheuvels for carefully reading the manuscript and making a number of important suggestions. He also thanks Sándor Csörgő, John Einmahl and Chris Klaassen for useful comments.

REFERENCES

Barrio, del, E., Giné, E. and Matrán, C. (1999). Central limit theorems for the Wasserstein distance between the empirical and true distribution. *Ann. Probab.* **27** 1009-1071.

Castelle, N. and Laurent-Bonvalot, F (1998). Strong approximations of bivariate uniform empirical processes. *Ann. Inst. Henri Poincaré* **34** 425-480.

Csörgő, M. and Horváth, L. (1993). *Weighted Approximations in Probability and Statistics*, John Wiley & Sons, Chichester etc.

Csörgő, M., Csörgő, S., Horváth, L. and Mason, D.M. (1986). Weighted empirical and quantile processes. *Ann. Probab.* **14** 31-85.

Csörgő, S., Haeusler, E. and Mason, D. M. (1988). A probabilistic approach to the asymptotic distribution of sums of independent, identically distributed random variable. *Adv. in Appl. Math.* **9** 259-333.

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution functions and of the classical multinomial estimator. *Ann. Math. Statist.* **27** 642-669.

Einmahl, J. H. J. and Mason, D. M. (1988). Strong limit theorems for weighted quantile processes. *Ann. Probab.* **16** 1623-1643.

Haeusler, E. and Mason, D. M. (1987). Laws of the iterated logarithm for sums of the middle portion of the sample. *Math. Proc. Cambridge Phil.* Soc. **101** 301–312.

Hahn, M., Mason, D. M. and Weiner, D. (1991). Eds. *Sums, Trimmed Sums and Extremes.* Birkhäuser, Boston

Hájek, J. (1961) On a simple linear model in Gaussian processes. *Trans. 2nd. Prague Conf. Information Theory* pp. 185-197 Publ.House Czechoslovak Acad. Sci., Prague, Academic Press, New York

Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent rv's and the sample df I. *Z. Wahrsch. verw. Gebiete.* **32** 111-131.

Le Cam, L. and Yang, G. (1990). *Asymptotics in Statistics: Some Basic Concepts.* Springer-Verlag, New York

Mason, D. M. (1991). A note on weighted approximations to the uniform empirical and quantile processes. *Sums, Trimmed Sums and Extremes.* pp. 269-284 Birkhäuser, Boston

Mason, D. M., Shorack, G. and Wellner, J. A. (1983). Strong limit theorems for oscillation moduli of the uniform empirical process. *Z. Wahrsch. verw. Gebiete.* **65** 83-97.

Mason, D. M. and Van Zwet, W.R. (1987). A refinement of the KMT inequality for the uniform empirical process. *Ann. Probab.* **15** 871-884.

Massart, P. (1990). The tight constant in the Dvoretzky–Kiefer–Wolfowitz inequality. *Ann. Probab.* **18** 1269-1283.

Nussbaum, M. (1996). Asymptotic equivalence of density estimation and Gaussian white noise. *Ann. Statist.* **24** 2399-2430.

Pollard, D. (1984). *Convergence of Stochastic Processes.* Springer–Verlag, New York–Berlin.

Rio, E. (1994). Local invariance principles and their application to density estimation. *Probab. Theory. Related Fields* **98** 21-45.

Shorack, G. (1997). Inequalities for quantile functions with a uniform studentized CLT that includes trimming. *Nonpar. Statist.* **8** 307-335.

Shorack, G. and Wellner, J. (1986). *Empirical Processes with Applications to Statistics*, Wiley, New York.

DEPARTMENT OF MATHEMATICAL SCIENCES
501 EWING HALL
UNIVERSITY OF DELAWARE
NEWARK, DELAWARE
USA
davidm@math.udel.edu