

The limit process of the difference between the empirical distribution function and its concave majorant

Vladimir N. Kulikov[†]
Hendrik P. Lopuhaä[‡]

01.06.2004

Abstract: We consider the process $\hat{F}_n - F_n$, being the difference between the empirical distribution function F_n and its least concave majorant \hat{F}_n , corresponding to a sample from a decreasing density. We extend Wang's result on pointwise convergence of $\hat{F}_n - F_n$ and prove that this difference converges as a process in distribution to the corresponding process for two-sided Brownian motion with parabolic drift.

[†]EURANDOM, P.O. Box 513 - 5600 MB Eindhoven, the Netherlands, kulikov@eurandom.nl

[‡]Delft University of Technology, Mekelweg 4 2628 CD Delft, the Netherlands, h.p.lopuhaa@its.tudelft.nl

1 Introduction

Let X_1, X_2, \dots, X_n be a sample from a decreasing density f with corresponding concave distribution function F . Suppose that f has bounded support, which then without loss of generality may be taken to be the interval $[0, 1]$. Let \hat{F}_n be the least concave majorant of the empirical distribution function F_n on $[0, 1]$, by which we mean the smallest concave function that lies above F_n . In this paper we study the difference between the processes \hat{F}_n and F_n .

Define the process

$$A_n(t) = n^{2/3} \left\{ \hat{F}_n(t) - F_n(t) \right\}, \quad t \in [0, 1]. \quad (1.1)$$

In KIEFER AND WOLFOWITZ (1976), it was shown that $(\log n)^{-1} \sup_t |A_n(t)|$ converges to zero with probability one, but the precise rate of convergence or limiting distribution was not given. WANG (1994) investigated the asymptotic behavior of $A_n(t)$, for t being a fixed point in $(0, 1)$. The limiting distribution can be described in terms of the operator CM_I that maps a function $h : \mathbb{R} \rightarrow \mathbb{R}$ into the least concave majorant of h on the interval $I \subset \mathbb{R}$. If we define the process

$$Z(t) = W(t) - t^2, \quad (1.2)$$

where W denotes standard two-sided Brownian motion originating from zero, then it is shown in WANG (1994) that, for $t \in (0, 1)$ fixed, $A_n(t)$ converges in distribution to $c_1(t)\zeta(0)$, where $c_1(t)$ is some constant depending of f and t , and

$$\zeta(t) = (\text{CM}_{\mathbb{R}}Z)(t) - Z(t). \quad (1.3)$$

Recently, DUROT AND TOQUET (2002) obtained the same result in a regression setting. In the present paper we will extend the pointwise result of WANG (1994) by proving that for any $t \in (0, 1)$ fixed, a properly scaled version $\zeta_{nt}(s) = c_1(t)A_n(t + c_2(t)sn^{-1/3})$ converges in distribution to the process $\zeta(s)$ in the space $D(\mathbb{R})$ of cadlag functions on \mathbb{R} .

One of the main tools in proving process convergence is the continuous mapping theorem. Observe that A_n is the image of F_n under the mapping $h \mapsto \text{CM}_I h - h$, which is a continuous mapping from the space $D(I)$ into itself. This is one of basic properties of concave majorants that are described in Lemma 2.1. In Section 3, we use the Hungarian embedding and the representation $B(F(s)) = W(F(s)) - W(1)F(s)$ for the Brownian bridge, to approximate the empirical process by the process

$$s \mapsto W \left(n^{1/3} (F(t + n^{-1/3}s) - F(t)) \right). \quad (1.4)$$

As a consequence, a properly scaled version of the process F_n converges to Brownian motion with parabolic drift. After establishing some preliminary results for this process in Section 3, application of the continuous mapping theorem yields convergence of the process A_n . The limit process is obtained in Section 4.

2 Basic properties of concave majorants

We start by giving a number of properties of the operator CM . By h'_r and h'_l , we will denote the right- and left-derivative of a function h . The following lemma lists a number of general basic properties of the operator CM .

Lemma 2.1 *Let g and h be functions on an interval $B \subset \mathbb{R}$. Then the following properties hold.*

1. $[\text{CM}_B g](t) \leq \sup_B g$, for all $t \in B$.
2. For any linear function $l(t) = at + b$ on B , we have $[\text{CM}_B(g + l)](t) = [\text{CM}_B g](t) + l(t)$ for all $t \in B$.
3. For any interval A , such that $A \subset B$, we have $[\text{CM}_A g](t) \leq [\text{CM}_B g](t)$, for all $t \in A$.
4. If $g \leq h$ on B , then $\text{CM}_B g \leq \text{CM}_B h$ on B .
5. $(\text{CM}_B g) + \inf_B h \leq \text{CM}_B(g + h) \leq (\text{CM}_B g) + \sup_B h$ on B .
6. Let $a, b, t \in B$, such that $a < t < b$ and suppose that

$$[\text{CM}_B g]'_r(a) > [\text{CM}_B g]'_l(t) \geq [\text{CM}_B g]'_r(t) > [\text{CM}_B g]'_l(b).$$

Then $[\text{CM}_B g](t) = [\text{CM}_{[a,b]} g](t)$.

7. Suppose $[x - 1, x + 1] \subset B$. Then

$$\left| [\text{CM}_B g]'(x) \right| \leq \max \left\{ \sup_B g - g(x - 1), \sup_B g - g(x + 1) \right\}.$$

8. Let $[a, b] \subset B \subset \mathbb{R}$ and suppose that $[\text{CM}_{[a,b]} g](x_1) = [\text{CM}_B g](x_1)$ and $[\text{CM}_{[a,b]} g](x_2) = [\text{CM}_B g](x_2)$, for any $x_1 < x_2$ in $[a, b]$. Then $[\text{CM}_{[a,b]} g](t) = [\text{CM}_B g](t)$, for all $t \in [x_1, x_2]$.

Proof: The proofs of properties 1-6 are straightforward (see KULIKOV (2003)). To prove property 7, suppose $[\text{CM}_B g]'(x) > 0$. Since $[x - 1, x] \subset B$, $g(x - 1)$ is below the line through the points $(x, [\text{CM}_B g](x))$ and $(x - 1, [\text{CM}_B g](x - 1))$. This means that the line through the points $(x, [\text{CM}_B g](x))$ and $(x - 1, g(x - 1))$ has a slope that is greater than $[\text{CM}_B g]'(x)$. Hence by property 1 we find

$$0 < [\text{CM}_B g]'(x) \leq \frac{[\text{CM}_B g](x) - g(x - 1)}{x - (x - 1)} \leq \left(\sup_B g \right) - g(x - 1).$$

Similarly, if $[\text{CM}_B g]'(x) \leq 0$,

$$0 \geq [\text{CM}_B g]'(x) \geq \frac{[\text{CM}_B g](x) - g(x + 1)}{x - (x + 1)} \geq - \left(\sup_B g \right) + g(x + 1).$$

To prove property 8, according to property 3 it suffices to show that $[\text{CM}_{[a,b]} g](t) \geq [\text{CM}_B g](t)$, for all $t \in [x_1, x_2]$. Suppose that there exists a $t \in (x_1, x_2)$, such that $[\text{CM}_{[a,b]} g](t) < [\text{CM}_B g](t)$. Let $[\tau_1, \tau_2] \subset [x_1, x_2]$ be the segment of $\text{CM}_{[a,b]} g$ that contains t . Since $\text{CM}_{[a,b]} g = \text{CM}_B g$ at the endpoints of the interval $[x_1, x_2]$, there must be a vertex τ_0 of $\text{CM}_B g$ between τ_1 and τ_2 , for which $[\text{CM}_{[a,b]} g](\tau_0) < [\text{CM}_B g](\tau_0)$. But then $g(\tau_0) = [\text{CM}_B g](\tau_0) > [\text{CM}_{[a,b]} g](\tau_0)$, which is in contradiction with $\text{CM}_B g \geq g$ on B . \blacksquare

3 Brownian approximation

In this section we show that by means of a local scaling the empirical process can be approximated by a Brownian motion, and we prove some preliminary results for Brownian motion with drift. Let E_n denote the empirical process $\sqrt{n}(F_n - F)$ and let B_n be a Brownian bridge constructed on the same probability space as the uniform empirical process $E_n \circ F^{-1}$ via the Hungarian embedding of KÓMLOS, MAJOR AND TUSNÁDY (1975). Then

$$\sup_{t \in [0,1]} |E_n(t) - B_n(F(t))| = \mathcal{O}_p(n^{-1/2} \log n). \quad (3.1)$$

Let ξ_n be a $N(0, 1)$ distributed random variable independent of B_n . Define versions W_n of Brownian motion by

$$W_n(t) = B_n(t) + \xi_n t, \quad t \in [0, 1]. \quad (3.2)$$

For $t \in (0, 1)$ fixed, define the process

$$X_{nt}(s) = n^{2/3} \left(F_n(t + sn^{-1/3}) - F_n(t) - \left(F(t + sn^{-1/3}) - F(t) \right) \right). \quad (3.3)$$

Let F have a continuously differentiable density f on $[0, 1]$ that satisfies

$$(A) \quad 0 < \inf_{t \in [0,1]} |f'(t)| \leq \sup_{t \in [0,1]} |f'(t)| < \infty.$$

Then on compact sets the process X_{nt} converges to a time-scaled Brownian motion.

Lemma 3.1 *Let F satisfy condition (A). Fix $t \in (0, 1)$ and let X_{nt} be defined by (3.3). Then the process $\{X_{nt}(s) : s \in \mathbb{R}\}$ converges in distribution to the process $\{W(f(t)s) : s \in \mathbb{R}\}$ in $D(\mathbb{R})$, the space of cadlag functions on \mathbb{R} .*

Proof: All trajectories of the limiting process belong to $C(\mathbb{R})$, the separable subset of continuous functions on \mathbb{R} . This means that similar to Theorem V.23 in POLLARD (1984), it suffices to show that for any compact set $I \subset \mathbb{R}$ the process $\{X_{nt}(s) : s \in I\}$ converges in distribution to the process $\{W(f(t)s) : s \in I\}$ in $D(I)$, the space of cadlag functions on I . We will apply Theorem V.3 in POLLARD (1984), which is stated for $D[0, 1]$, but the same result holds for $D(I)$.

By applying (3.1), we can write

$$\begin{aligned} X_{nt}(s) &= n^{1/6} \left\{ E_n(t + sn^{-1/3}) - E_n(t) \right\} \\ &= n^{1/6} \left\{ B_n(F(t + sn^{-1/3})) - B_n(F(t)) \right\} + \mathcal{O}_p(n^{-1/3} \log n) \\ &= n^{1/6} \left\{ W_n(F(t + sn^{-1/3})) - W_n(F(t)) \right\} + \mathcal{O}_p(n^{-1/6} \log n), \end{aligned}$$

where the big \mathcal{O} -term is uniform for $s \in I$. By using Brownian scaling, a simple Taylor expansion, and the uniform continuity of Brownian motion on compacta, we find that

$$X_{nt}(s) \stackrel{d}{=} W(f(t)s) + R_n(s),$$

where $\sup_{s \in I} |R_n(s)| \rightarrow 0$ in probability. From this representation it follows immediately that the process $\{X_{nt}(s) : s \in I\}$ satisfies the conditions of Theorem V.3 in POLLARD (1984). This proves the lemma. ■

Let D_I be the operator that maps a function $h : \mathbb{R} \rightarrow \mathbb{R}$ into the difference between the least concave majorant of h on the interval I and h itself

$$D_I h = \text{CM}_I h - h.$$

Then the process A_n can be written as a functional of F_n only: $A_n = n^{2/3}[D_{[0,1]}F_n]$. This means that in order to obtain the limiting behaviour of A_n we must investigate the limiting behaviour of F_n itself. Note that

$$n^{2/3} \left(F(t + sn^{-1/3}) - F(t) \right) \approx n^{1/3} f(t)s + \frac{1}{2} f'(t)s^2.$$

By property 2 of Lemma 2.1 the operator $D_{[0,1]}$ will be invariant under addition of linear functions. Hence the term $n^{1/3} f(t)s$ will have no effect on the limiting behaviour of A_n . In view of Lemma 3.1 this means that limiting behaviour of A_n will be determined by the concave majorant of Brownian motion with a parabolic drift.

The following two lemmas are concerned with the tail behaviour of Brownian motion with polynomial drift. The first lemma ensures that the probability that the process $W(t) - K|t|^\alpha$ is still positive for $|t| > a$, decreases exponentially as $a \rightarrow \infty$. The second lemma states that the distribution function of $\sup_{t \in \mathbb{R}} (W(t) - K|t|^\alpha)$ has exponential tails.

Lemma 3.2 For all $K > 0$, $\alpha \geq 1$ and $a \geq 0$,

$$P \left\{ \sup_{|t| > a} \left(W(t) - K|t|^\alpha \right) \geq 0 \right\} \leq 2 \exp \left(-K^2 a^{2\alpha-1} / 2 \right).$$

Proof: Using that $\{tW(1/t) : t > 0\}$ has the same distribution as standard Brownian motion, we have

$$\begin{aligned} P \left\{ \sup_{|t| > a} \left(W(t) - K|t|^\alpha \right) \geq 0 \right\} &\leq 2P \left\{ \sup_{t > a} (W(t)/t) \geq Ka^{\alpha-1} \right\} \\ &= 2P \left\{ \sup_{t \in (0,1)} (tW(a/t)/a) \geq Ka^{\alpha-1} \right\} \\ &\leq 2P \left\{ \sup_{t \in [0,1]} W(t) \geq Ka^{\alpha-1/2} \right\}. \end{aligned}$$

Finally, note that

$$P \left\{ \sup_{t \in [0,1]} W(t) \geq Ka^{\alpha-1/2} \right\} \leq 2P \left\{ W(1) \geq Ka^{\alpha-1/2} \right\}, \quad (3.4)$$

and

$$2P\{W(1) \geq x\} \leq 2e^{-x^2/2} \int_0^\infty \phi(u) du = e^{-x^2/2}$$

which proves the lemma. ■

Lemma 3.3 For all $K > 0$, $\alpha \geq 1$, and $x \geq 0$,

$$P \left\{ \sup_{\mathbb{R}} \left(W(t) - K|t|^\alpha \right) > x \right\} \leq 4 \exp \left(-K^{1/\alpha} x^{2-1/\alpha} / 2 \right).$$

Proof: We can bound the tail probability on the left hand side by

$$P \left\{ \sup_{|t| \leq (x/K)^{1/\alpha}} W(t) > x \right\} + P \left\{ \sup_{|t| > (x/K)^{1/\alpha}} \left(W(t) - K|t|^\alpha \right) > 0 \right\}.$$

Similar to (3.4) the first probability is bounded by

$$P \left\{ \sup_{|t| \leq (x/K)^{1/\alpha}} W(t) > x \right\} \leq 2 \exp \left(-K^{1/\alpha} x^{2-1/\alpha} / 2 \right).$$

From Lemma 3.2, we find the same bound for the second probability. ■

With suitable standardization, the limiting Brownian motion with parabolic drift can be transformed to the process Z defined in (1.2). We will be dealing with the concave majorants of this process on large bounded intervals and on the whole real line. Property 8 of Lemma 2.1 guarantees that both concave majorants are the same on an interval as soon as their values coincide on the boundary of the interval. The next lemma states that for large intervals this happens with large probability. For $d > 0$, consider the event

$$N(d) = \left\{ [\text{CM}_{\mathbb{R}Z}](s) = [\text{CM}_{[-d,d]Z}](s), \text{ for } s = \pm d/2 \right\} \quad (3.5)$$

Lemma 3.4 *There exist constants $C_1 > 0$ and $C_2 > 0$, such that for all $d \geq 0$*

$$P(N(d)^c) \leq C_1(1 + d^{1/2}) \exp \left(-C_2 d^{3/2} \right).$$

Proof: Let \bar{Z} be the process

$$\bar{Z}(s) = Z(s + d/2).$$

Then, by symmetry and property 3 of Lemma 2.1, we have

$$\begin{aligned} P \{N(d)^c\} &\leq 2P \{ [\text{CM}_{\mathbb{R}Z}](d/2) > [\text{CM}_{[-d,d]Z}](d/2) \} \\ &= 2P \{ [\text{CM}_{\mathbb{R}\bar{Z}}](0) > [\text{CM}_{[-d/2,3d/2]\bar{Z}}](0) \} \\ &\leq 2P \{ [\text{CM}_{\mathbb{R}\bar{Z}}](0) > [\text{CM}_{[-d/2,d/2]\bar{Z}}](0) \}. \end{aligned}$$

Note that

$$\bar{Z}(s) = W(s + d/2) - (s + d/2)^2 \stackrel{d}{=} Z(s) + W(d/2) - d^2/4 - sd.$$

Hence by property 2 of Lemma 2.1, we find that

$$\begin{aligned} P \{N(d)^c\} &\leq 2P \{ [\text{D}_{\mathbb{R}\bar{Z}}](0) > [\text{D}_{[-d/2,d/2]\bar{Z}}](0) \} \\ &= 2P \{ [\text{D}_{\mathbb{R}Z}](0) > [\text{D}_{[-d/2,d/2]Z}](0) \} \\ &= 2P \{ [\text{CM}_{\mathbb{R}Z}](0) > [\text{CM}_{[-d/2,d/2]Z}](0) \}. \end{aligned}$$

Distinguish between

1. $Z(s) \leq -\frac{1}{2}s^2$, for all $|s| > d/2$,
2. $Z(s) > -\frac{1}{2}s^2$, for some $|s| > d/2$.

Consider the first case. Let $[\tau_1, \tau_2] \subset [-d/2, d/2]$ be the segment of $\text{CM}_{[-d/2, d/2]}$ that contains zero. If $[\text{CM}_{\mathbb{R}}Z](0) > [\text{CM}_{[-d/2, d/2]}Z](0)$, then the line through the points $(\tau_1, Z(\tau_1))$ and $(\tau_2, Z(\tau_2))$ must intersect the process Z outside the interval $[-d/2, d/2]$. In case 1, this can only happen if this line intersects the parabola $p(s) = -\frac{1}{2}s^2$ outside the interval $[-d/2, d/2]$. This is only possible if the slope of this line is greater than the tangent of $p(s)$ in the point $s = -d/2$ or smaller than the tangent of $p(s)$ in the point $s = d/2$:

$$\left| [\text{CM}_{[-d/2, d/2]}Z]'(0) \right| \geq d/2.$$

We find that

$$P\{N(d)^c\} \leq 2P\left\{ \sup_{|s|>d/2} \left(Z(s) + \frac{1}{2}s^2 \right) > 0 \right\} + 2P\left\{ \left| [\text{CM}_{[-d/2, d/2]}Z]'(0) \right| \geq d/2 \right\}. \quad (3.6)$$

By Lemma 3.2, the first term on the right hand side of (3.6) is equal to

$$2P\left\{ \sup_{|s|>d/2} \left(W(s) - \frac{1}{2}s^2 \right) > 0 \right\} \leq 4e^{-d^3/64}.$$

By property 7 of Lemma 2.1 the second term on the right hand side of (3.6) is bounded by

$$\begin{aligned} & 2P\left\{ \sup_{s \in \mathbb{R}} (W(s) - s^2) + 1 - \min\{W(1), W(-1)\} \geq d/2 \right\} \\ & \leq 2P\left\{ \sup_{s \in \mathbb{R}} (W(s) - s^2) \geq d/4 \right\} + 2P\{1 - \min\{W(1), W(-1)\} \geq d/4\} \end{aligned}$$

According to Lemma 3.3, the first term is bounded by $8e^{-d^3/2/16}$. Property (3.4) implies that the second term on the right hand side is bounded by

$$2P\{W(1) \leq 1 - d/4\} + 2P\{W(-1) \leq 1 - d/4\} \leq e^{-\frac{1}{2}(1-d/4)^2}.$$

This proves the lemma. ■

Likewise we will have to deal with the difference between concave majorants of F_n on intervals $[t - dn^{-1/3}, t + dn^{-1/3}]$ and on $[0, 1]$, as well as with the difference between concave majorants of the Brownian approximation of F_n . To this end define $F_n^E = F_n$ and let F_n^W be its Brownian approximation defined by

$$F_n^W(t) = F(t) + n^{-1/2}W_n(F(t)), \quad t \in [0, 1], \quad (3.7)$$

where W_n is defined in (3.2). For $t \in [0, 1]$ and $d > 0$ let $I_{nt}(d) = [t - dn^{-1/3}, t + dn^{-1/3}]$, and for $J = E, W$ define the event

$$N_{nt}^J(d) = \left\{ [\text{CM}_{[0,1]}F_n^J](t + sn^{-1/3}) = [\text{CM}_{I_{nt}(d)}F_n^J](t + sn^{-1/3}), \text{ for } s = \pm d/2 \right\}. \quad (3.8)$$

The following lemma ensures that the value of the two concave majorants of F_n^J coincide at $t \pm \frac{1}{2}dn^{-1/3}$ with high probability.

Lemma 3.5 Let $t \in (0, 1)$ and $d > 0$ be such that $0 < t - n^{-1/3}d < t + n^{-1/3}d < 1$. Moreover, suppose that for $\epsilon_n = \frac{1}{4} \inf |f'|n^{-1/3}d$,

$$f(1) < f(t + n^{-1/3}d) - \epsilon_n < f(t - n^{-1/3}d) + \epsilon_n < f(0).$$

Then, for any distribution function F satisfying condition (A), there exist a constant $C_2 > 0$, such that

$$P \{N_{nt}^J(d)^c\} \leq 8 \exp(-C_2 d^3),$$

where C_2 does not depend on d , t and n .

Proof: For $J = E, W$ define \hat{f}_n^J as the left-derivative of $\text{CM}_{[0,1]} F_n^J$. Define

$$U_n^J(a) = \operatorname{argmax}_{t \in [0,1]} \{F_n^J(t) - at\} \quad \text{and} \quad V_n^J(a) = n^{1/3} (U_n^J(a) - g(a)),$$

where g denotes the inverse of f . The process U_n^J is related to \hat{f}_n^J by the relation

$$\hat{f}_n^J(t) \leq a \iff U_n^J(a) \leq t \text{ with probability one.} \quad (3.9)$$

Suppose that the concave majorants of F_n^J on the intervals $[0, 1]$ and $[t - dn^{-1/3}, t + dn^{-1/3}]$ differ at $t - dn^{-1/3}/2$. A simple picture shows that in that case there cannot be point of jump of \hat{f}_n^J between $t - dn^{-1/3}$ and $t - dn^{-1/3}/2$, which implies that $\hat{f}_n^J(t - dn^{-1/3}) \leq \hat{f}_n^J(t - dn^{-1/3}/2)$. Similarly, if the concave majorants of F_n^J on the intervals $[0, 1]$ and $[t - dn^{-1/3}, t + dn^{-1/3}]$ differ at $t + dn^{-1/3}/2$, then $\hat{f}_n^J(t + dn^{-1/3}) \leq \hat{f}_n^J(t + dn^{-1/3}/2)$. Hence

$$\begin{aligned} P \{N_{nt}^J(d)^c\} &\leq P \left\{ \hat{f}_n^J(t - n^{-1/3}d) \leq \hat{f}_n^J(t - n^{-1/3}d/2) \right\} \\ &\quad + P \left\{ \hat{f}_n^J(t + n^{-1/3}d) \geq \hat{f}_n^J(t + n^{-1/3}d/2) \right\}. \end{aligned} \quad (3.10)$$

Consider the first probability on the right hand side of (3.10). Then with $s = t - n^{-1/3}d$ and $x = d/2$, we have

$$\begin{aligned} &P \left\{ \hat{f}_n^J(t - n^{-1/3}d) \leq \hat{f}_n^J(t - n^{-1/3}d/2) \right\} \\ &= P \left\{ \hat{f}_n^J(s + n^{-1/3}x) \geq \hat{f}_n^J(s) \right\} \\ &= P \left\{ \left(\hat{f}_n^J(s + n^{-1/3}x) - f(s + n^{-1/3}x) \right) - \left(\hat{f}_n^J(s) - f(s) \right) \geq f(s) - f(s + n^{-1/3}x) \right\} \\ &\leq P \left\{ \left(\hat{f}_n^J(s + n^{-1/3}x) - f(s + n^{-1/3}x) \right) - \left(\hat{f}_n^J(s) - f(s) \right) \geq n^{-1/3}x \inf |f'| \right\} \\ &\leq P \left\{ \hat{f}_n^J(s + n^{-1/3}x) - f(s + n^{-1/3}x) \geq \epsilon_n \right\} + P \left\{ \hat{f}_n^J(s) - f(s) \leq -\epsilon_n \right\}. \end{aligned} \quad (3.11)$$

By using (3.9), the first probability on the right hand side of (3.11) is equal to

$$\begin{aligned} &P \left\{ U_n^J \left(f(s + xn^{-1/3}) + \epsilon_n \right) \geq s + n^{-1/3}x \right\} \\ &= P \left\{ V_n^J \left(f(s + xn^{-1/3}) + \epsilon_n \right) \geq n^{1/3} \left(s + n^{-1/3}x - g(f(s + xn^{-1/3}) + \epsilon_n) \right) \right\} \\ &\geq P \left\{ V_n^J \left(f(s + xn^{-1/3}) + \epsilon_n \right) \geq n^{1/3} \epsilon_n \inf |g'| \right\} \\ &= P \left\{ V_n^J \left(f(s + xn^{-1/3}) + \epsilon_n \right) \geq \frac{\inf |f'|d}{4 \sup |f'|} \right\}. \end{aligned}$$

Since $f(s + xn^{-1/3}) + \epsilon_n = f(t - dn^{-1/3}/2) + \epsilon_n \in [f(1), f(0)]$, it follows from Theorems 2.1 and 3.1 in GROENEBOOM, HOOGHMSTRA AND LOPUHAÄ (1999) that

$$P \left\{ V_n^J \left(f(s + xn^{-1/3}) + \epsilon_n \right) \geq \frac{\inf |f'|d}{4 \sup |f'|} \right\} \leq 2e^{-C_2 d^3}, \quad (3.12)$$

for some constant $C_2 > 0$, not depending on n , t and d . The second probability on the right hand side of (3.11) can be bounded similarly,

$$P \left\{ \hat{f}_n^J(s) - f(s) \leq -\epsilon_n \right\} \leq 2e^{-C_2 d^3}.$$

Together with (3.12) we conclude that the probability of the first event on the right hand side of (3.10) can be bounded as follows

$$P \left\{ \hat{f}_n^J(t - n^{-1/3}d) \geq \hat{f}_n^J(t - n^{-1/3}d/2) \right\} \leq 4e^{-C_2 d^3}.$$

The probability of the second event on the right hand side of (3.10) can be bounded similarly, by taking $s = t + n^{-1/3}d/2$ and $x = d/2$ and using the same argument as above. This proves the lemma. \blacksquare

4 Process convergence

For $t \in (0, 1)$ fixed and $t + c_2(t)sn^{-1/3} \in (0, 1)$, define

$$\zeta_{nt}(s) = c_1(t)A_n(t + c_2(t)sn^{-1/3}), \quad (4.1)$$

where

$$c_1(t) = \left(\frac{|f'(t)|}{2f^2(t)} \right)^{1/3} \quad \text{and} \quad c_2(t) = \left(\frac{4f(t)}{|f'(t)|^2} \right)^{1/3}. \quad (4.2)$$

Define $\zeta_{nt}(s) = 0$ for $t + c_2(t)sn^{-1/3} \notin (0, 1)$. The following theorem states that the process ζ_{nt} converges to the process ζ on $D(\mathbb{R})$.

Theorem 4.1 *Let f be decreasing with support on $[0, 1]$. Suppose that f is continuous differentiable on $[0, 1]$ and satisfies*

$$(A) \quad 0 < \inf_{t \in [0, 1]} |f'(t)| < \sup_{t \in [0, 1]} |f'(t)| < \infty.$$

Let the processes ζ and ζ_{nt} be defined as in (1.3) and (4.1). Then the process $\{\zeta_{nt}(s) : s \in \mathbb{R}\}$ converges in distribution to the process $\{\zeta(s) : s \in \mathbb{R}\}$ in $D(\mathbb{R})$, the space of cadlag functions of \mathbb{R} .

Proof: Similar to the proof of Theorem 3.1 it is enough to show that for any compact set $K \subset \mathbb{R}$, the process $\{\zeta_{nt}(s) : s \in K\}$ converges in distribution to the process $\{\zeta(s) : s \in K\}$ on $D(K)$. Note that for this, it suffices to show that the process $\{A_n(t + sn^{-1/3}) : s \in K\}$ converges in distribution to the process $\{[D_{\mathbb{R}}Z_t](s) : s \in K\}$, where

$$Z_t(s) = W(f(t)s) + \frac{1}{2}f'(t)s^2. \quad (4.3)$$

This follows from the fact that by Brownian scaling $c_1(t)Z_t(c_2(t)s) \stackrel{d}{=} Z(s) = W(s) - s^2$.

Let $t \in (0, 1)$ fixed, and let $I_{nt} = [-tn^{1/3}, (1-t)n^{1/3}]$. Write $E_{nt}(s) = n^{2/3}F_n(t + sn^{-1/3})$, for $s \in I_{nt}$. Then by definition

$$A_n(t + sn^{-1/3}) = [D_{I_{nt}}E_{nt}](s) \quad \text{for } s \in I_{nt}.$$

Now take K fixed. For the processes $\{[D_{I_{nt}}E_{nt}](s) : s \in K\}$ and $\{[D_{\mathbb{R}}Z_t](s) : s \in K\}$, we must show that for any $g : D(K) \rightarrow \mathbb{R}$ bounded and continuous:

$$|Eg(D_{I_{nt}}E_{nt}) - Eg(D_{\mathbb{R}}Z_t)| \rightarrow 0.$$

Let $\epsilon > 0$ and let $I = [-d, d]$ be an interval, where according to Lemmas 3.4 and 3.5, $d > 0$ is chosen sufficiently large such that $K \subset [-d/2, d/2]$, and such that

$$P(N(d)^c) < \epsilon \quad \text{and} \quad P(N_{nt}^E(d)^c) < \epsilon, \quad (4.4)$$

where $N(d)$ and $N_{nt}^E(d)$ are defined in (3.5) and (3.8). Let n be sufficiently large, such that $K \subset [-d/2, d/2] \subset I \subset I_{nt}$. For $g : D(K) \rightarrow \mathbb{R}$ bounded and continuous, and processes $\{[D_{I_{nt}}E_{nt}](s) : s \in K\}$, $\{[D_{\mathbb{R}}Z_t](s) : s \in K\}$, and $\{[D_I Z_t](s) : s \in K\}$, we have

$$\begin{aligned} |Eg(D_{I_{nt}}E_{nt}) - Eg(D_{\mathbb{R}}Z_t)| &\leq |Eg(D_{I_{nt}}E_{nt}) - Eg(D_I E_{nt})| \\ &\quad + |Eg(D_I E_{nt}) - Eg(D_I Z_t)| \\ &\quad + |Eg(D_I Z_t) - Eg(D_{\mathbb{R}}Z_t)|. \end{aligned} \quad (4.5)$$

For the first term on the right hand side of (4.5) we have that

$$\begin{aligned} &|Eg(D_{I_{nt}}E_{nt}) - Eg(D_I E_{nt})| \\ &\leq 2 \sup |g| \cdot P\{D_{I_{nt}}E_{nt} \neq D_I E_{nt} \text{ on } [-d/2, d/2]\} \\ &\leq 2 \sup |g| \cdot P\{[D_{I_{nt}}E_{nt}](s) \neq [D_I E_{nt}](s) \text{ for } s = -d/2 \text{ or } s = d/2\} \\ &= 2 \sup |g| \cdot P\{[CM_{I_{nt}}E_{nt}](s) \neq [CM_I E_{nt}](s) \text{ for } s = -d/2 \text{ or } s = d/2\}. \end{aligned}$$

Suppose that $[CM_{I_{nt}}E_{nt}](s) \neq [CM_I E_{nt}](s)$. This means that the concave majorants of F_n itself on the intervals $[0, 1]$ and $I_{nt}(d) = [t - dn^{-1/3}, t + dn^{-1/3}]$ differ at $t + sn^{-1/3}$. Hence the probability on the right hand side above can be bounded as follows:

$$P\{[CM_{I_{nt}}E_{nt}](s) \neq [CM_I E_{nt}](s) \text{ for } s = -d/2 \text{ or } s = d/2\} \leq P(N_{nt}^E(d)^c).$$

According to (4.4), this yields

$$|Eg(D_{I_{nt}}E_{nt}) - Eg(D_I E_{nt})| \leq 2 \sup |g| \cdot \epsilon. \quad (4.6)$$

Similarly, application of property 8 of Lemma 2.1 gives

$$\begin{aligned} |Eg(D_I Z_t) - Eg(D_{\mathbb{R}}Z_t)| &\leq 2 \sup |g| \cdot P\{D_I Z_t \neq D_{\mathbb{R}}Z_t \text{ on } K\} \\ &\leq 2 \sup |g| \cdot P\{D_I Z_t \neq D_{\mathbb{R}}Z_t \text{ on } [-d/2, d/2]\} \\ &\leq 2 \sup |g| \cdot P(N(d)^c). \end{aligned}$$

Once more (4.4) yields

$$|Eg(D_I Z_t) - Eg(D_{\mathbb{R}}Z_t)| \leq 2 \sup |g| \cdot \epsilon. \quad (4.7)$$

In order to bound the second term on the right hand side of (4.5) define

$$Z_{nt}(s) = n^{2/3} \left(F_n(t + sn^{-1/3}) - F_n(t) - \left(F(t + sn^{-1/3}) - F(t) \right) \right) + \frac{1}{2} f'(t) s^2.$$

It follows from Lemma 3.1, that the process $\{Z_{nt}(s) : s \in I\}$ converges in distribution to the process $\{Z_t(s) : s \in I\}$. Because according to property 5 of Lemma 2.1, the mapping $D_I : D(I) \rightarrow D(I)$ is continuous, this means that

$$|Eh(D_I Z_{nt}) - Eh(D_I Z_t)| \rightarrow 0,$$

for any $h : D(I) \rightarrow \mathbb{R}$ bounded and continuous. Note that we can also write

$$E_{nt}(s) = Z_{nt}(s) + n^{2/3} F_n(t) + f(t) sn^{1/3} + R_{nt}(s),$$

where

$$R_{nt}(s) = n^{2/3} \left[F(t + sn^{-1/3}) - F(t) - f(t) sn^{-1/3} - \frac{1}{2} f'(t) s^2 n^{-2/3} \right].$$

Note that for some $|\theta - t| \leq n^{-1/3}|s|$, with $s \in I$, we have

$$R_{nt}(s) = \frac{1}{2} |f'(\theta) - f'(t)| s^2 \rightarrow 0,$$

uniformly for $s \in I$, using that f' is continuous. By continuity of the mapping D_I together with property 2 of Lemma 2.1, it then follows that on I :

$$D_I Z_{nt} = D_I (E_{nt} - R_{nt}) = D_I E_{nt} + o(1),$$

where the $o(1)$ -term is uniform for $s \in I$. We conclude that for any $h : D(I) \rightarrow \mathbb{R}$ bounded and continuous, and processes $\{[D_{I_{nt}} E_{nt}](s) : s \in I\}$ and $\{[D_I Z_t](s) : s \in I\}$,

$$|Eh(D_I E_{nt}) - Eh(D_I Z_t)| \rightarrow 0. \tag{4.8}$$

Now let $\pi_K : D(I) \rightarrow D(K)$ be defined as the restriction of an element of $D(I)$ to the set K . Since for any $g : D(K) \rightarrow \mathbb{R}$ bounded and continuous the composition $h = g \circ \pi_K$ is also bounded and continuous, (4.8) implies that for $g : D(K) \rightarrow \mathbb{R}$ bounded and continuous, and processes $\{[D_{I_{nt}} E_{nt}](s) : s \in K\}$ and $\{[D_I Z_t](s) : s \in K\}$,

$$|Eg(D_I E_{nt}) - Eg(D_I Z_t)| \rightarrow 0. \tag{4.9}$$

Putting together (4.6), (4.7), (4.9) and (4.5) proves the theorem. ■

References

- DUROT, C. AND TOQUET, A.S. (2002) On the distance between the empirical process and its concave majorant in a monotone regression framework. Submitted.
- GROENEBOOM, P., HOOGHIEMSTRA, G. AND LOPUHAÄ, H.P. (1999) Asymptotic normality of the L_1 error of the Grenander estimator. *Ann. Statist.* **27** 1316-1347.
- KIEFER, J. AND WOLFOWITZ, J. (1976) Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrsch. verw. Gebiete* **34** 73-85.

- KÓMLOS, J., MAJOR, P., AND TUSNÁDY, G. (1975) An approximation of partial sums of independent RV's and the sample DF. *Z. Wahrsch. verw. Gebiete* **32** 111-131.
- KULIKOV, V.N. (2003) Direct and Indirect use of Maximum Likelihood Methods in Testing. Phd. Thesis, TU Delft.
- POLLARD, D. (1984) *Convergence of stochastic processes*. Springer-Verlag, New York.
- WANG, Y. (1994) The limit distribution of the concave majorant of an empirical distribution function. *Statistics and Probability Letters* **20** 81-84.