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

Vladimir N. Kulikova, Hendrik P. Lopuhaä

Abstract

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

MSC: primary: 60F05 62E20; secondary: 62G30

Keywords: Least concave majorant; Empirical distribution function; Brownian motion with parabolic drift; Isotonic estimation; Monotone density

1. Introduction and main result

Let \( X_1, X_2, \ldots, X_n \) be a sample from a decreasing density \( f \). Suppose that \( f \) has bounded support, which then without loss of generality may be taken to be the interval \([0, 1]\). Let \( \tilde{F}_n \) 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 \( F_n \). The process

\[
A_n(t) = n^{2/3} (\tilde{F}_n(t) - F_n(t)), \quad t \in [0, 1],
\]

has been of interest to several authors. 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 \( t > 0 \) being fixed. The limiting distribution can be described in terms of the operator \( CM_I \) that maps a function \( h : \mathbb{R} \to \mathbb{R} \) into the least concave majorant of \( h \) on the interval \( I \subset \mathbb{R} \). If we define the process

\[
Z(t) = W(t) - t^2,
\]

\[\approx\]
where \( W \) denotes standard two-sided Brownian motion originating from zero, then it is shown in Wang (1994) that, for \( t > 0 \) fixed, \( A_n(t) \) converges in distribution to \( c_1(t) \zeta(0) \), where \( c_1(t) \) is defined in (1.5), and
\[
\zeta(t) = [CM_{\mathbb{R}}Z](t) - Z(t).
\]

Recently, Durot and Toquet (2002) obtained the same result in a regression setting.

In the present paper, we extend the pointwise result of Wang (1994) to process convergence of a suitably scaled version of \( A_n(t) \). For \( t \in (0, 1) \) fixed and \( t + c_2(t)sn^{-1/3} \in (0, 1) \), define
\[
\zeta_m(s) = c_1(t)A_n(t + c_2(t)sn^{-1/3}),
\]
where
\[
c_1(t) = \left( \frac{|f'(t)|}{2f^2(t)} \right)^{1/3} \quad \text{and} \quad c_2(t) = \left( \frac{4f(t)}{|f'(t)|^2} \right)^{1/3}.
\]

Define \( \zeta_m(s) = 0 \) for \( t + c_2(t)sn^{-1/3} \notin (0, 1) \). Our main result is the following theorem.

**Theorem 1.1.** Suppose that \( f \) satisfies

\( \begin{align*}
\text{(A1)} & \quad f \text{ is decreasing and continuously differentiable;} \\
\text{(A2)} & \quad 0 < f'(t) \leq f(t) \leq f(0) < \infty \text{ for } 0 \leq s \leq t \leq 1; \\
\text{(A3)} & \quad 0 < \inf_{t \in [0,1]} |f'(t)|.
\end{align*} \)

Let \( \zeta \) and \( \zeta_m \) be defined as in (1.3) and (1.4). Then the process \( \{\zeta_m(s) : s \in \mathbb{R}\} \) converges in distribution to the process \( \{\zeta(s) : s \in \mathbb{R}\} \) in \( D(\mathbb{R}) \), the space of cadlag functions on \( \mathbb{R} \).

In the remainder of this section we sketch the line of argument used to prove Theorem 1.1. All proofs are transferred to Section 2.

Let \( D_I \) be the operator that maps a function \( h : \mathbb{R} \to \mathbb{R} \) into the difference between the least concave majorant of \( h \) on the interval \( I \) and \( h \) itself:
\[
D_I h = CM_I h - h.
\]

Then \( D_I h \) is a continuous mapping from the space \( D(I) \) into itself. This is a consequence of the basic property for concave majorants, that for any interval \( I \subset \mathbb{R}, \)
\[
\inf_{t \in I} h(t) \leq CM_I (g + h) - CM_I g \leq \sup_{t \in I} h(t)
\]
(see for instance Kulikov, 2003). The key observation for proving process convergence is the fact that the process \( A_n \) is the image of \( F_n \) under the mapping \( D_{[0,1]} : A_n = n^{2/3}D_{[0,1]}F_n \). This means that in order to obtain the limiting behaviour of \( A_n \), we must investigate the limiting behaviour of \( F_n \) itself. This is described in the following lemma.

**Lemma 1.1.** Let \( f \) satisfy conditions (A1)–(A3). Then for \( t \in (0, 1) \) fixed, the process
\[
n^{2/3}(F_n(t + sn^{-1/3}) - F_n(t)) - (F(t + sn^{-1/3}) - F(t)) \quad \text{for } s \in \mathbb{R},
\]
converges in distribution to the process \( \{W(f(t)s) : s \in \mathbb{R}\} \) in \( D(\mathbb{R}) \).

Note that \( n^{2/3}(F(t + sn^{-1/3}) - F(t)) \approx n^{1/3}f(t)s + f'(t)s^2/2 \) and that the operator \( D_{[0,1]} \) is invariant under addition of linear functions. For this reason, the term \( n^{1/3}f(t)s \) will have no effect on the limiting behaviour of \( A_n \), which will therefore be determined by the concave majorant of Brownian motion with a parabolic drift.

In order to apply Lemma 1.1 together with the continuity of the mapping \( h \to CM_I h - h \), we must consider concave majorants on fixed intervals. However, up to scaling constants, the processes \( \zeta_m(s) \) correspond to concave majorants of \( F_n(t + sn^{-1/3}) \) with \( s \) in intervals \( [-tn^{1/3}, (1 - t)n^{1/3}] \) increasing to \( \mathbb{R} \), whereas the process \( \zeta(s) \) corresponds to the concave majorant of \( Z \) on \( \mathbb{R} \). Hence, in order to establish Theorem 1.1, we must show that with high probability, the concave majorants of \( Z \) on the whole real line and on large bounded intervals
Proof of Lemma 1.1. Let \( X_{n}(s) \) denote the process in (1.6). 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 \( \{X_{n}(s) : s \in I\} \) 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 \( D(I) \).

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 Kőmlos et al. (1975). Let \( \xi_{n} \) be a \( N(0,1) \) distributed random variable independent of \( B_{n} \). Define versions \( W_{n} \) of Brownian motion by \( W_{n}(t) = B_{n}(t) + \xi_{n}t, t \in [0,1] \). Since

\[
\sup_{t \in [0,1]} |E_{n}(t) - B_{n}(F(t))| = C_{p}(n^{-1/2} \log n),
\]

we can write

\[
X_{n}(s) = n^{-1/6} \{W_{n}(F(t) + sn^{-1/3}) - W_{n}(F(t))\} + C_{p}(n^{-1/6} \log n),
\]

where the big \( C \)-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

\[
X_{n}(s) \overset{d}{=} W(f(t)s) + R_{n}(s),
\]

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

Proof of Lemma 1.2. Note that the concave majorants of \( Z \) on \([-d,d]\) and on \( \mathbb{R} \) are the same on the interval \([-d/2,d/2]\) as soon as their values coincide at the boundary points \( \pm d/2 \). Hence, by symmetry

\[
P(N(d)^{c}) \leq 2P([CM_{R}Z](d/2) > [CM_{[-d,d]}Z](d/2))
\]

\[
\leq 2P([CM_{R}Z](d/2) > [CM_{[0,d]}Z](d/2)).
\]

From Lemma 3.2 in Durot and Toquet (2002), with \( c = d/4 \) and \( t = d/2 \) the latter probability is bounded by \( 8 \exp(-d^{3}/2^{7}) \).
Proof of Lemma 1.3. Define $\hat{f}_n$ as the left-derivative of $\hat{F}_n$. Define

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

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

$$\hat{f}_n(t) \leq a \iff U_n(a) \leq t,$$

with probability one. First suppose that $0 < t - dn^{-1/3} < t + dn^{-1/3} < 1$, so that $I_n(d) = [t - dn^{-1/3}, t + dn^{-1/3}]$. On the event $N_n(d)^c$, the concave majorants of $F_n$ on the intervals $[0, 1]$ and $[t - dn^{-1/3}, t + dn^{-1/3}]$ differ either at $s = t - dn^{-1/3}/2$ or at $s = t + dn^{-1/3}/2$. A simple picture shows that in that case $\hat{f}_n$ cannot have a point of jump both on the intervals $[t - dn^{-1/3}, t - dn^{-1/3}/2]$ and $[t + dn^{-1/3}/2, t + dn^{-1/3}]$. This implies

$$P[N_n(d)^c] \leq P[\hat{f}_n(t - n^{-1/3}d) = \hat{f}_n(t - n^{-1/3}d/2)] + P[\hat{f}_n(t + n^{-1/3}d) = \hat{f}_n(t + n^{-1/3}d/2)].$$

Consider the first probability on the right-hand side of (2.2). Define $\epsilon_n = \frac{1}{3}\inf |f'|dn^{-1/3}$. Then with $s = t - n^{-1/3}d$ and $x = d/2$, we have

$$P[\hat{f}_n(t - n^{-1/3}d) = \hat{f}_n(t - n^{-1/3}d/2)]
= P[\hat{f}_n(s + n^{-1/3}x) = \hat{f}_n(s)]
\leq P[\hat{f}_n(s + n^{-1/3}x) - f(s + n^{-1/3}x) < n^{-1/3}x\inf |f'|]
\leq P[\hat{f}_n(s + n^{-1/3}x) - f(s + n^{-1/3}x) > \epsilon_n] + P[\hat{f}_n(s) - f(s) < -\epsilon_n].$$

By using (2.1), the first probability on the right-hand side of (2.3) is equal to

$$P[U_n(f(s + xn^{-1/3}) + \epsilon_n) > s + n^{-1/3}x]
= P[V_n(f(s + xn^{-1/3}) + \epsilon_n) > n^{1/3}(s + n^{-1/3}x - g(f(s + xn^{-1/3}) + \epsilon_n))]
\leq P \left\{ V_n(f(s + xn^{-1/3}) + \epsilon_n) > \frac{\inf |f'|d}{8 \sup |f'|} \right\}.$$ 

Clearly, $f(s + xn^{-1/3}) + \epsilon_n \geq f(s) + \epsilon_n$, and since $t > dn^{-1/3}$, it follows that $f(s + xn^{-1/3}) + \epsilon_n = f(t - dn^{-1/3}/2) + \epsilon_n < f(0)$. According to Theorem 2.1 in Groeneboom et al. (1999) (note that the proof of this theorem does not use that $f''$ exists), for $a \in (f(1), f(0))$ and $x > 0$, $P[V_n(a) \geq x] \leq 2e^{-Mx^2}$, with $M > 0$ only depending on $f$. This means that

$$P \left\{ V_n(f(s + xn^{-1/3}) + \epsilon_n) > \frac{\inf |f'|d}{8 \sup |f'|} \right\} \leq 2e^{-Cd^3},$$

for some constant $C > 0$ not depending on $n, t$ and $d$. The second probability on the right-hand side of (2.3) can be bounded similarly,

$$P[\hat{f}_n(s) - f(s) < -\epsilon_n] \leq 2e^{-Cd^3}.$$ 

Together with (2.4) we conclude that the probability of the first event on the right-hand side of (2.2) can be bounded as follows

$$P[\hat{f}_n(t - n^{-1/3}d) = \hat{f}_n(t - n^{-1/3}d/2)] \leq 4e^{-Cd^3}.$$ 

The probability of the second event on the right-hand side of (2.2) 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 for the case $0 < t - dn^{-1/3} < t + dn^{-1/3} < 1$.

When $0 < t - dn^{-1/3} < t + dn^{-1/3} < 1$, then $I_n(d) = [t - dn^{-1/3}, 1]$ and on $N_n(d)^c$ the concave majorants of $F_n$ on the intervals $[0, 1]$ and $[t - dn^{-1/3}, 1]$ differ at $s = t - dn^{-1/3}/2$. In that case $\hat{f}_n$ cannot have a point of jump on the interval $[t - dn^{-1/3}, t - dn^{-1/3}/2]$, so that

$$P[N_n(d)^c] \leq P[\hat{f}_n(t - n^{-1/3}d) = \hat{f}_n(t - n^{-1/3}d/2)] \leq 4e^{-Cd^3}.$$
Finally, when $t - dn^{-1/3} \leq 0 < t + dn^{-1/3} < 1$, then $I_n(d) = [0, t + dn^{-1/3}]$ and on $N_n(d)$ the concave majorants of $F_n$ on the intervals $[0, 1]$ and $[0, t + dn^{-1/3}]$ differ at $s = t + dn^{-1/3}/2$. In that case $\hat{f}_n$ cannot have a point of jump on the interval $[t + dn^{-1/3}/2, t + dn^{-1/3}]$, so that

$$P(N_n(d)) \leq P(\hat{f}_n(t + n^{-1/3} - d) = \hat{f}_n(t + n^{-1/3}/2)) \leq 4e^{-Cd}. \quad \Box$$

**Proof of Theorem 1.1.** Similar to the proof of Lemma 1.1 it is enough to show that for any compact set $K \subseteq \mathbb{R}$, the process $\{\zeta_n(s) : s \in K\}$ 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 $\{D_n Z_s(s) : s \in K\}$, where

$$Z_t(s) = W(f(t)s) + \frac{1}{2}f'(t)s^2. \quad (2.5)$$

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

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

$$A_n(t + sn^{-1/3}) = [D_n E_{nt}(s) : s \in K]. \quad \text{for } s \in I_{nt}. \quad (2.6)$$

Now consider $K$ is fixed. For the processes $\{D_n E_{nt}(s) : s \in K\}$ and $\{D_n Z_s(s) : s \in K\}$, we must show that for any $g : D(K) \rightarrow \mathbb{R}$ bounded and continuous:

$$|E g(D_n E_{nt}(s)) - E g(D_n Z_s(s))| \rightarrow 0. \quad \text{for } t \rightarrow 0.$$ 

Let $\varepsilon > 0$ and let $I = [-d/d, d/d]$ be an interval, where $d > 0$ is chosen sufficiently large such that $K \subseteq [-d/2, d/2]$, and such that according to Lemmas 1.2 and 1.3

$$P(N(d/c_2(t))) < \varepsilon \quad \text{and} \quad P(N_n(d)) < \varepsilon, \quad (2.7)$$

where $N(d)$ and $N_n(d)$ are defined in (1.7) and (1.8). Let $n$ be sufficiently large, such that $K \subseteq [-d/2, d/2] \subseteq I \subseteq I_{nt}$. For $g : D(K) \rightarrow \mathbb{R}$ bounded and continuous, and processes $\{D_n E_{nt}(s) : s \in K\}$, $\{D_n Z_s(s) : s \in K\}$, and $\{D_n Z_s(s) : s \in K\}$, we have

$$|E g(D_n E_{nt}(s)) - E g(D_n Z_s(s))| \leq |E g(D_n E_{nt}(s)) - E g(D_n Z_s(s))|$$

$$+ |E g(D_n Z_s(s)) - E g(D_n Z_s(s))|$$

$$\leq |E g(D_n Z_s(s)) - E g(D_n Z_s(s))|. \quad (2.8)$$

For the last term on the right-hand side of (2.7) we have that

$$|E g(D_n Z_s(s)) - E g(D_n Z_s(s))| \leq 2 \sup |g| \cdot P(D_n Z_s(s) \neq D_n Z_s(s) \text{ on } [t])$$

$$\leq 2 \sup |g| \cdot P(D_n Z_s(s) \neq D_n Z_s(s) \text{ on } [-d/2, d/2]).$$

Since $c_1(t) Z_t(c_2(t)s) \overset{d}{=} Z(s)$, the latter probability is bounded by $P(N(d/c_2(t)))$, so that (2.6) yields

$$|E g(D_n Z_s(s)) - E g(D_n Z_s(s))| \leq 2 \sup |g| \cdot \varepsilon. \quad (2.9)$$

The first term can be bounded similarly:

$$|E g(D_n E_{nt}(s)) - E g(D_n E_{nt}(s))| \leq 2 \sup |g| \cdot P(CM_n E_{nt} \neq CM_n E_{nt} \text{ on } [-d/2, d/2])$$

$$\leq 2 \sup |g| \cdot P(N_n(d)) \leq 2 \sup |g| \cdot \varepsilon. \quad (2.10)$$

In order to bound the second term on the right-hand side of (2.7), define

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

It follows from Lemma 1.1, that the process $\{Z_n(s) : s \in I\}$ converges in distribution to the process $\{Z_t(s) : s \in I\}$. Because the mapping $D_t : D(I) \rightarrow D(I)$ is continuous, this means that

$$|E h(D_t Z_n) - E h(D_t Z_n)| \rightarrow 0,$$

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
\[
R_{nt}(s) = n^{2/3} [F(t + sn^{-1/3}) - F(t) - f(t)sn^{-1/3} - \frac{1}{2}f'(t)s^2 n^{-2/3}].
\]
Note that for some \(|y - t| \leq n^{-1/3}|s|\), with \(s \in I\), we have
\[
R_{nt}(s) = \frac{1}{2}f'(y) - f'(t)|s^2 | \to 0,
\]
uniformly for \(s \in I\), using that \(f'\) is continuous. By continuity of the mapping \(D_I\) together with the property that \(D_I\) is invariant under addition of linear functions, 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 \(h : D(I) \to \mathbb{R}\) bounded and continuous, and processes \(\{[D_I E_{nt}](s) : s \in I\}\) and \(\{[D_I Z_{t}](s) : s \in I\}\),
\[
|E h(D_I E_{nt}) - E h(D_I Z_{t})| \to 0. \tag{2.10}
\]
Now let \(\pi_K : D(I) \to D(K)\) be defined as the restriction of an element of \(D(I)\) to the set \(K\). Since for any \(g : D(K) \to \mathbb{R}\) bounded and continuous the composition \(h = g \circ \pi_K\) is also bounded and continuous, \(2.10\) implies that for \(g : D(K) \to \mathbb{R}\) bounded and continuous, and processes \(\{[D_I E_{nt}](s) : s \in K\}\) and \(\{[D_I Z_{t}](s) : s \in K\}\),
\[
|E g(D_I E_{nt}) - E g(D_I Z_{t})| \to 0. \tag{2.11}
\]
Putting together \(2.8\), \(2.9\), \(2.11\) and \(2.7\) proves the theorem. \(\square\)

Acknowledgements

We thank the referee for his/her suggestions, which helped us to improve the paper. Especially, we thank the referee for providing a simpler proof of Lemma 1.2.

References