

INCONSISTENCY OF BOOTSTRAP: THE GRENANDER ESTIMATOR

BY BODHISATTVA SEN,^{*} MOULINATH BANERJEE[†] AND MICHAEL WOODROOFE^{*}

Columbia University, University of Michigan and University of Michigan

In this paper we investigate the (in)-consistency of different bootstrap methods for constructing confidence intervals in the class of estimators that converge at rate $n^{\frac{1}{3}}$. The Grenander estimator, the nonparametric maximum likelihood estimator of an unknown non-increasing density function f on $[0, \infty)$, is a prototypical example. We focus on this example and explore different approaches to constructing bootstrap confidence intervals for $f(t_0)$, where $t_0 \in (0, \infty)$ is an interior point. We find that the bootstrap estimate, when generating bootstrap samples from the empirical distribution function \mathbb{F}_n or its least concave majorant \tilde{F}_n , does not have any weak limit in probability. We provide a set of sufficient conditions for the consistency of any bootstrap method in this example and show that bootstrapping from a smoothed version of \tilde{F}_n leads to strongly consistent estimators. The m out of n bootstrap method is also shown to be consistent while generating samples from \mathbb{F}_n and \tilde{F}_n .

1. Introduction. If $X_1, X_2, \dots, X_n \sim^{\text{ind}} f$ are a sample from a non-increasing density f on $[0, \infty)$, then the Grenander estimator, the nonparametric maximum likelihood estimator (NPMLE) \tilde{f}_n of f [obtained by maximizing the likelihood $\prod_{i=1}^n f(X_i)$ over all non-increasing densities], may be described as follows: Let \mathbb{F}_n denote the empirical distribution function (EDF) of the data, and \tilde{F}_n its least concave majorant. Then the NPMLE \tilde{f}_n is the left hand derivative of \tilde{F}_n . This result is due to Grenander (1956) and is described in detail by Robertson, Wright and Dykstra (1988), pp. 326-328. Prakasa Rao (1969) obtained the asymptotic distribution of \tilde{f}_n , properly normalized: Let \mathbb{W} be a two-sided standard Brownian motion on \mathbb{R} with $\mathbb{W}(0) = 0$ and

$$\mathbb{C} = \arg \max_{s \in \mathbb{R}} [\mathbb{W}(s) - s^2].$$

^{*}Supported by NSF Grant AST-05-07453

[†]Supported by NSF Grant DMS-07-05288

AMS 2000 subject classifications: Primary 62G09, 62G20; secondary 62G07

Keywords and phrases: decreasing density, empirical distribution function, least concave majorant, m out of n bootstrap, nonparametric maximum likelihood estimate, smoothed bootstrap

If $0 < t_0 < \infty$ and $f'(t_0) \neq 0$, then

$$(1.1) \quad n^{\frac{1}{3}} \{ \tilde{f}_n(t_0) - f(t_0) \} \Rightarrow 2 \left| \frac{1}{2} f(t_0) f'(t_0) \right|^{\frac{1}{3}} \mathbb{C},$$

where \Rightarrow denotes convergence in distribution. There are other estimators that exhibit similar asymptotic properties; for example, Chernoff's (1964) estimator of the mode, the monotone regression estimator [Brunk (1970)], Rousseeuw's (1984) least median of squares estimator, and the estimator of the shorth [Andrews et al. (1972) and Shorack and Wellner (1986)]. The seminal paper by Kim and Pollard (1990) unifies $n^{\frac{1}{3}}$ -rate of convergence problems in the more general M-estimation framework. Tables and a survey of statistical problems in which the distribution of \mathbb{C} arises are provided by Groeneboom and Wellner (2001).

The presence of nuisance parameters in the limiting distribution (1.1) complicates the construction of confidence intervals. Bootstrap intervals avoid the problem of estimating nuisance parameters and are generally reliable in problems with \sqrt{n} convergence rates. See Bickel and Freedman (1981), Singh (1981), Shao and Tu (1995) and its references. Our aim in this paper is to study the consistency of bootstrap methods for the Grenander estimator with the hope that the monotone density estimation problem will shed light on the behavior of bootstrap methods in similar cube-root convergence problems.

There has been considerable recent interest in this question. Kosorok (2007) show that bootstrapping from the EDF \mathbb{F}_n does not lead to a consistent estimator of the distribution of $n^{1/3} \{ \tilde{f}_n(t_0) - f(t_0) \}$. Lee and Pun (2006) explore m out of n bootstrapping from the empirical distribution function in similar non-standard problems and prove the consistency of the method. Léger and MacGibbon (2006) describe conditions for a resampling procedure to be consistent under cube root asymptotics and assert that these conditions are generally not met while bootstrapping from the EDF. They also propose a smoothed version of the bootstrap and show its consistency for Chernoff's estimator of the mode. Abrevaya and Huang (2005) show that bootstrapping from the EDF leads to inconsistent estimators in the setup of Kim and Pollard (1990) and propose corrections. Romano, Politis and Wolf (1999) show that subsampling based confidence intervals are consistent in this scenario.

Our work goes beyond that cited above as follows: We show that bootstrapping from the NPMLE \tilde{F}_n also leads to inconsistent estimators, a result that we found more surprising, since \tilde{F}_n has a density. Moreover, we find that *the bootstrap estimator, constructed from either the EDF or NPMLE, has no*

limit in probability. The finding is less than a mathematical proof, because one step in the argument relies on simulation; but the simulations make our point clearly. As described in Section 5 our findings are inconsistent with some claims of Abrevaya and Huang (2005). Also, our way of tackling the main issues differs from that of the existing literature: We consider conditional distributions in more detail than Kosorok (2007), who deduced inconsistency from properties of unconditional distributions; we directly appeal to the characterization of the estimators and use a continuous mapping principle to deduce the limiting distributions instead of using the “switching” argument [see Groeneboom (1985)] employed by Kosorok (2007) and Abrevaya and Huang (2005); and at a more technical level, we use the Hungarian Representation Theorem whereas most of the other authors use empirical process techniques similar to those described by van der Vaart and Wellner (2000).

Section 2 contains a uniform version of (1.1) that is used later on to study the consistency of different bootstrap methods and may be of independent interest. The main results on inconsistency are presented in Section 3. Sufficient conditions for the consistency of a bootstrap method are presented in Section 4 and applied to show that bootstrapping from smoothed versions of \tilde{F}_n does produce consistent estimators. The m out of n bootstrapping procedure is investigated, when generating bootstrap samples from \mathbb{F}_n and \tilde{F}_n . It is shown that both the methods lead to consistent estimators under mild conditions on m . In Section 5 we discuss our findings, especially the non-convergence and its implications. Section A, the appendix, provides the details of some arguments used in proving the main results.

2. Uniform Convergence. For the rest of the paper F denotes a distribution function with $F(0) = 0$ and a density f that is non-increasing on $[0, \infty)$ and continuously differentiable near $t_0 \in (0, \infty)$ with nonzero derivative $f'(t_0) < 0$. If $g : I \rightarrow \mathbb{R}$ is a bounded function, write $\|g\| := \sup_{x \in I} |g(x)|$. Next, let F_n be distribution functions with $F_n(0) = 0$, that converge weakly to F and, therefore,

$$(2.1) \quad \lim_{n \rightarrow \infty} \|F_n - F\| = 0.$$

Let $X_{n,1}, X_{n,2}, \dots, X_{n,m_n} \sim^{\text{ind}} F_n$, where $m_n \leq n$ is a non-decreasing sequence of integers for which $m_n \rightarrow \infty$; let \mathbb{F}_{n,m_n} denote the EDF of $X_{n,1}, X_{n,2}, \dots, X_{n,m_n}$; and let

$$\Delta_n := m_n^{\frac{1}{3}} \{ \tilde{f}_{n,m_n}(t_0) - f_n(t_0) \}$$

where $\tilde{f}_{n,m_n}(t_0)$ is the Grenander estimator computed from $X_{n,1}, X_{n,2}, \dots, X_{n,m_n}$ and $f_n(t_0)$ is the density of F_n at t_0 or a surrogate. Next, let $I_m = [-t_0 m^{\frac{1}{3}}, \infty)$ and

$$(2.2) \quad \mathbb{Z}_n(h) := m_n^{\frac{2}{3}} \left\{ \mathbb{F}_{n,m_n}(t_0 + m_n^{-\frac{1}{3}}h) - \mathbb{F}_{n,m_n}(t_0) - f_n(t_0)m_n^{-\frac{1}{3}}h \right\}$$

for $h \in I_{m_n}$ and observe that Δ_n is the left hand derivative at 0 of the least concave majorant of \mathbb{Z}_n . It is fairly easy to obtain the asymptotic distribution of \mathbb{Z}_n . The asymptotic distribution of Δ_n may then be obtained from the Continuous Mapping Theorem. Stochastic processes are regarded as random elements in $D(\mathbb{R})$, the space of right continuous functions on \mathbb{R} with left limits, equipped with the projection σ -field and the topology of uniform convergence on compacta. See Pollard (1984), Chapters IV and V for background.

2.1. *Convergence of \mathbb{Z}_n .* It is convenient to decompose \mathbb{Z}_n into the sum of $\mathbb{Z}_{n,1}$ and $\mathbb{Z}_{n,2}$ where

$$\begin{aligned} \mathbb{Z}_{n,1}(h) &:= m_n^{\frac{2}{3}} \left\{ (\mathbb{F}_{n,m_n} - F_n)(t_0 + m_n^{-\frac{1}{3}}h) - (\mathbb{F}_{n,m_n} - F_n)(t_0) \right\} \\ \mathbb{Z}_{n,2}(h) &:= m_n^{\frac{2}{3}} \left\{ F_n(t_0 + m_n^{-\frac{1}{3}}h) - F_n(t_0) - f_n(t_0)m_n^{-\frac{1}{3}}h \right\}. \end{aligned}$$

Observe that $\mathbb{Z}_{n,2}$ depends only on F_n and f_n ; only $\mathbb{Z}_{n,1}$ depends on $X_{n,1}, \dots, X_{n,m_n}$. Let \mathbb{W}_1 be a standard two-sided Brownian motion on \mathbb{R} with $\mathbb{W}_1(0) = 0$, and $\mathbb{Z}_1(h) = \mathbb{W}_1[f(t_0)h]$.

PROPOSITION 2.1. *If*

$$(2.3) \quad \lim_{n \rightarrow \infty} m_n^{\frac{1}{3}} \left| F_n(t_0 + m_n^{-\frac{1}{3}}h) - F_n(t_0) - f(t_0)m_n^{-\frac{1}{3}}h \right| = 0$$

uniformly on compacts (in h), then $\mathbb{Z}_{n,1} \Rightarrow \mathbb{Z}_1$; and if there is a continuous function \mathbb{Z}_2 for which

$$(2.4) \quad \lim_{n \rightarrow \infty} \mathbb{Z}_{n,2}(h) = \mathbb{Z}_2(h)$$

uniformly on compact intervals, then $\mathbb{Z}_n \Rightarrow \mathbb{Z} := \mathbb{Z}_1 + \mathbb{Z}_2$.

PROOF. The Hungarian Embedding Theorem of Kómlós, Major and Tusnády (1975) is used. We may suppose that $X_{n,i} = F_n^\#(U_i)$, where $F_n^\#(u) = \inf\{x : F_n(x) \geq u\}$ and U_1, U_2, \dots are i.i.d. Uniform(0,1) random variables. Let \mathbb{U}_n denote the EDF of U_1, \dots, U_n , $\mathbb{E}_n(t) = \sqrt{n}[\mathbb{U}_n(t) - t]$, and

$\mathbb{V}_n = \sqrt{m_n}(\mathbb{F}_{n,m_n} - F_n)$. Then $\mathbb{V}_n = \mathbb{E}_{m_n} \circ F_n$. By Hungarian Embedding, we may also suppose that the probability space supports a sequence of Brownian Bridges $\{\mathbb{B}_n^0\}_{n \geq 1}$ for which

$$(2.5) \quad \sup_{0 \leq t \leq 1} |\mathbb{E}_n(t) - \mathbb{B}_n^0(t)| = O\left[\frac{\log(n)}{\sqrt{n}}\right] \text{ a.s.}$$

and a standard normal random variable η that is independent of $\{\mathbb{B}_n^0\}_{n \geq 1}$. Define a version \mathbb{B}_n of Brownian motion by $\mathbb{B}_n(t) = \mathbb{B}_n^0(t) + \eta t$, for $t \in [0, 1]$. Then

$$(2.6) \quad \begin{aligned} \mathbb{Z}_{n,1}(h) &= m_n^{\frac{1}{6}} \left\{ \mathbb{E}_{m_n}[F_n(t_0 + m_n^{-\frac{1}{3}}h)] - \mathbb{E}_{m_n}[F_n(t_0)] \right\} \\ &= m_n^{\frac{1}{6}} \left\{ \mathbb{B}_{m_n}[F_n(t_0 + m_n^{-\frac{1}{3}}h)] - \mathbb{B}_{m_n}[F_n(t_0)] \right\} + \mathbb{R}_n(h) \end{aligned}$$

$$\begin{aligned} \text{where } |\mathbb{R}_n(h)| &\leq 2m_n^{\frac{1}{6}} \sup_{0 \leq t \leq 1} |\mathbb{E}_{m_n}(t) - \mathbb{B}_{m_n}^0(t)| \\ &+ m_n^{\frac{1}{6}} |\eta| |F_n(t_0 + m_n^{-\frac{1}{3}}h) - F_n(t_0)| \rightarrow 0 \end{aligned}$$

uniformly on compacta *w.p.1* using (2.3) and (2.5). Let

$$\mathbb{X}_n(h) := m_n^{\frac{1}{6}} \{ \mathbb{B}_{m_n}[F_n(t_0 + m_n^{-\frac{1}{3}}h)] - \mathbb{B}_{m_n}[F_n(t_0)] \}$$

and observe that \mathbb{X}_n is a mean zero Gaussian process defined on I_{m_n} with independent increments and covariance kernel

$$K_n(h_1, h_2) = m_n^{\frac{1}{3}} \left| F_n[t_0 + \text{sign}\{h_1\}m_n^{-\frac{1}{3}}(|h_1| \wedge |h_2|)] - F_n(t_0) \right| \mathbf{1}\{h_1 h_2 > 0\}.$$

It now follows from Theorem V.19 in Pollard (1984) and (2.3) that $\mathbb{X}_n(h)$ converges in distribution to $\mathbb{W}_1[f(t_0)h]$ in $D([-c, c])$ for every $c > 0$, establishing the first assertion of the Proposition. The second then follows from Slutsky's Theorem. \square

2.2. Convergence of Δ_n . Unfortunately, Δ_n is not quite a continuous functional of \mathbb{Z}_n . If $f : I \rightarrow \mathbb{R}$, write $f|_J$ to denote the restriction of f to $J \subseteq I$; and if I and J are intervals and f is bounded, write $L_J f$ for the least concave majorant of the restriction. Thus, $\tilde{F}_n = L_{[0, \infty)} \mathbb{F}_n$ in the Introduction.

LEMMA 2.2. *Let I be a closed interval; let $f : I \rightarrow \mathbb{R}$ be a bounded upper semi-continuous function on I ; and let $a_1, a_2, b_1, b_2 \in I$ with $b_1 < a_1 < a_2 < b_2$. If $2f[\frac{1}{2}(a_i + b_i)] > L_I f(a_i) + L_I f(b_i)$, $i = 1, 2$, then $L_I f(x) = L_{[b_1, b_2]} f(x)$ for $a_1 \leq x \leq a_2$.*

PROOF. This follows from the proof of Lemma 5.1 and Lemma 5.2 of Wang and Woodroffe (2007). In that lemma continuity was assumed, but only upper semi-continuity was used in the (short) proof. \square

Recall Marshall's Lemma: If I is an interval, $f : I \rightarrow \mathbb{R}$ is bounded, and $g : I \rightarrow \mathbb{R}$ is concave, then $\|L_I f - g\| \leq \|f - g\|$. See, for example, Robertson, Wright, and Dykstra (1988), pp. 329, for a proof. Write $\tilde{F}_{n,m_n} = L_{[0,\infty)} \mathbb{F}_{n,m_n}$.

LEMMA 2.3. *If $\delta > 0$ is so small that F is strictly concave on $[t_0 - 2\delta, t_0 + 2\delta]$ and (2.1) holds then $\tilde{F}_{n,m_n} = L_{[t_0-2\delta, t_0+2\delta]} \mathbb{F}_{n,m_n}$ on $[t_0 - \delta, t_0 + \delta]$ for all large n w.p.1.*

PROOF. Since F is strictly concave on $[t_0 - 2\delta, t_0 + 2\delta]$, $2F(t_0 \pm \frac{3}{2}\delta) > F(t_0 \pm \delta) + F(t_0 \pm 2\delta)$. Then,

$$\begin{aligned} \|\tilde{F}_{n,m_n} - F\| &\leq \|\mathbb{F}_{n,m_n} - F\| \\ &\leq \|\mathbb{F}_{n,m_n} - F_n\| + \|F_n - F\| \\ &\leq \frac{1}{\sqrt{m_n}} \|\mathbb{E}_{m_n}\| + \|F_n - F\| \rightarrow 0 \text{ w.p. } 1 \end{aligned}$$

by Marshall's Lemma, (2.1) and the Glivenko-Cantelli Theorem. Thus, $2\mathbb{F}_{n,m_n}(t_0 \pm \frac{3}{2}\delta) > \tilde{F}_{n,m_n}(t_0 \pm \delta) + \tilde{F}_{n,m_n}(t_0 \pm 2\delta)$, for all large n w.p.1, and Lemma 2.3 follows from Lemma 2.2. \square

PROPOSITION 2.4. *(i) Suppose that (2.1) and (2.3) hold and given $\gamma > 0$, there are $0 < \delta < 1$ and $C > 0$ for which*

$$(2.7) \quad \left| F_n(t_0 + h) - F_n(t_0) - f_n(t_0)h - \frac{1}{2}f'(t_0)h^2 \right| \leq \gamma h^2 + C m_n^{-\frac{2}{3}}$$

and

$$(2.8) \quad |F_n(t_0 + h) - F_n(t_0)| \leq C(|h| + m_n^{-\frac{1}{3}})$$

for $|h| \leq \delta$ and for all large n . If J is a compact interval and $\epsilon > 0$, then there is a compact $K \supseteq J$, depending only on ϵ , J , C , γ , and δ , for which

$$(2.9) \quad P[L_{I_{m_n}} \mathbb{Z}_n = L_K \mathbb{Z}_n \text{ on } J] \geq 1 - \epsilon$$

for all large n .

(ii) Let \mathbb{Y} be an a.s. continuous stochastic process on \mathbb{R} that is a.s. bounded above. If $\lim_{|h| \rightarrow \infty} \mathbb{Y}(h)/|h| = -\infty$ a.e., then the compact $K \supseteq J$ can be chosen so that

$$(2.10) \quad P[L_{\mathbb{R}} \mathbb{Y} = L_K \mathbb{Y} \text{ on } J] \geq 1 - \epsilon.$$

PROOF. For a fixed sequence $(F_n \equiv F)$ (2.9) would follow from the assertion in Example 6.5 of Kim and Pollard (1990), and it is possible to adapt their argument to a triangular array using (2.7) and (2.8) in place of Taylor series expansion. A different proof is presented in the Appendix. \square

We will use the following easily verified fact. In its statement the metric space \mathcal{X} is to be endowed with the projection σ -field. See Pollard (1984), page 70.

LEMMA 2.5. *Let $\{X_{n,c}\}, \{Y_n\}, \{W_c\}$ and Y be sets of random elements taking values in a metric space (\mathcal{X}, d) , $n = 0, 1, \dots$, and $c \in \mathbb{R}$. If for any $\delta > 0$,*

- (i) $\lim_{c \rightarrow \infty} \limsup_{n \rightarrow \infty} P\{d(X_{n,c}, Y_n) > \delta\} = 0$,
- (ii) $\lim_{c \rightarrow \infty} P\{d(W_c, Y) > \delta\} = 0$,
- (iii) $X_{n,c} \Rightarrow W_c$ as $n \rightarrow \infty$ for every $c \in \mathbb{R}$,

then $Y_n \Rightarrow Y$ as $n \rightarrow \infty$.

COROLLARY 2.6. *If (2.9) and (2.10) hold, and $Z_n \Rightarrow \mathbb{Y}$, then $L_{I_{m_n}} Z_n \Rightarrow L_{\mathbb{R}} \mathbb{Y}$ in $D(\mathbb{R})$ and $\Delta_n \Rightarrow (L_{\mathbb{R}} \mathbb{Y})'(0)$.*

PROOF. It suffices to show that $L_{I_{m_n}} Z_n|J \Rightarrow L_{\mathbb{R}} \mathbb{Y}|J$ in $D(J)$, for every compact interval $J \subseteq \mathbb{R}$. Given J and $\epsilon > 0$, there exists K_ϵ , a compact, $K_\epsilon \supseteq J$, such that (2.9) and (2.10) hold. This verifies (i) and (ii) of Lemma 2.5 with $c = 1/\epsilon$, $X_{n,c} = L_{K_\epsilon} Z_n$, $Y_n = L_{I_{m_n}} Z_n$, $W_c = L_{K_\epsilon} \mathbb{Y}$, $Y = L_{\mathbb{R}} \mathbb{Y}$ and $d(x, y) = \sup_{t \in J} |x(t) - y(t)|$. Clearly, $L_{K_\epsilon} Z_n|J \Rightarrow L_{K_\epsilon} \mathbb{Y}|J$ in $D(J)$, by the Continuous Mapping Theorem, verifying condition (iii). Thus $L_{I_{m_n}} Z_n \Rightarrow L_{\mathbb{R}} \mathbb{Y}$ in $D(\mathbb{R})$. Another application of the Continuous Mapping Theorem [via the lemma on page 330 of Robertson, Wright and Dykstra (1988)] in conjunction with (2.9), (2.10) and Lemma 2.5 then shows that $\Delta_n = (L_{I_{m_n}} Z_n)'(0) \Rightarrow (L_{\mathbb{R}} \mathbb{Y})'(0)$. \square

COROLLARY 2.7. *If (2.1), (2.3), (2.4), (2.7) and (2.8) hold and $\lim_{|h| \rightarrow \infty} Z(h)/|h| = -\infty$, then $L_{I_{m_n}} Z_n \Rightarrow L_{\mathbb{R}} Z$ in $D(\mathbb{R})$ and $\Delta_n \Rightarrow (L_{\mathbb{R}} Z)'(0)$; and if $Z_2(h) = f'(t_0)h^2/2$, then $\Delta_n \Rightarrow 2|\frac{1}{2}f(t_0)f'(t_0)|^{\frac{1}{3}}\mathbb{C}$, where \mathbb{C} has Chernoff's distribution.*

PROOF. The convergence follows directly from Proposition 2.4 and Corollary 2.6. Note that if $Z_2(h) = f'(t_0)h^2/2$, then (2.9) and (2.10) hold and Corollary 2.6 can be applied. That $(L_{\mathbb{R}} Z)'(0)$ is distributed as $2|\frac{1}{2}f(t_0)f'(t_0)|^{\frac{1}{3}}\mathbb{C}$ when $Z_2(h) = f'(t_0)h^2/2$ follows from elementary properties of Brownian motion via the “switching” argument of Groeneboom (1985). \square

2.3. *Remarks on the Conditions.* If $F_n \equiv F$ and $f_n \equiv f$, then clearly (2.1), (2.3), (2.4), (2.7) and (2.8) all hold with $\mathbb{Z}_2(h) = f'(t_0)h^2/2$ for some $0 < \delta < 1$ and $C \geq f(t_0 - \delta)$ by a Taylor expansion of F and the continuity of f and f' around t_0 .

COROLLARY 2.8. *If there is a $\delta > 0$ for which F_n has a continuously differentiable density f_n on $[t_0 - \delta, t_0 + \delta]$, and*

$$(2.11) \quad \lim_{n \rightarrow \infty} \left[\|F_n - F\| + \sup_{|t-t_0| < \delta} (|f_n(t) - f(t)| + |f'_n(t) - f'(t)|) \right] = 0$$

then (2.1), (2.3), (2.4), (2.7) and (2.8) hold with $\mathbb{Z}_2(h) = f'(t_0)h^2/2$, and $\Delta_n \Rightarrow 2|\frac{1}{2}f(t_0)f'(t_0)|^{\frac{1}{3}}\mathbb{C}$.

PROOF. The result can be immediately derived from Taylor expansion of F_n and the continuity of f and f' around t_0 . To illustrate the idea we show that (2.7) holds. Let $\gamma > 0$ be given. Clearly,

$$(2.12) \quad \begin{aligned} & \left| F_n(t_0 + h) - F_n(t_0) - f_n(t_0)h - \frac{1}{2}h^2f'(t_0) \right| \\ & \leq \frac{1}{2}h^2 \sup_{|s| \leq |h|} |f'_n(t_0 + s) - f'(t_0)|. \end{aligned}$$

Let $\delta > 0$ be so small that $|f'(t) - f'(t_0)| \leq \gamma$ for $|t - t_0| < \delta$, and let n_0 be so large that $\sup_{|t-t_0| \leq \delta} |f'_n(t) - f'(t)| \leq \gamma$ for $n \geq n_0$. Then the last line in (2.12) is at most γh^2 for $|h| \leq \delta$ and $n \geq n_0$. \square

Another useful remark, used below, is that if $\lim_{n \rightarrow \infty} m_n^{\frac{1}{3}} \|F_{m_n} - F\| = 0$, then (2.1), (2.3) and (2.8) hold.

In the next three sections, we apply Proposition 2.1 and Corollary 2.6 to bootstrap samples drawn from the EDF, its LCM, and smoothed versions thereof. Thus, let $X_1, X_2, \dots \sim^{\text{ind}} F$; let \mathbb{F}_n be the EDF of X_1, \dots, X_n ; and let \tilde{F}_n be its LCM. If $F_n = \mathbb{F}_n$, then (2.1), (2.3) and (2.8) hold almost surely by the above remark, since

$$(2.13) \quad \|\mathbb{F}_n - F\| = O \left[\sqrt{\frac{\log \log(n)}{n}} \right] a.s.$$

by the Law of the Iterated Logarithm for the EDF, which may be deduced from Hungarian Embedding; and the same is true if $F_n = \tilde{F}_n$ since $\|\tilde{F}_n - F\| \leq \|\mathbb{F}_n - F\|$, by Marshall's Lemma.

If $m_n = n$ and $f_n = \tilde{f}_n$, then (2.4) is not satisfied almost surely or in probability by either \mathbb{F}_n or \tilde{F}_n . For either choice (2.7) is satisfied in probability if $f_n = f$.

PROPOSITION 2.9. *Suppose that $m_n = n$ and that $f_n = f$. If F_n is either the EDF \mathbb{F}_n or its LCM \tilde{F}_n , then for any $\gamma, \epsilon > 0$, there are $C > 0$ and $0 < \delta < 1$ for which (2.7) holds with probability at least $1 - \epsilon$ for all large n .*

The proof is included in the Appendix.

3. Inconsistency and non-convergence of the bootstrap. We begin with a brief discussion of the bootstrap.

3.1. *Generalities.* Now, suppose that $X_1, X_2, \dots \sim^{\text{ind}} F$ are defined on a probability space (Ω, \mathcal{A}, P) . Write $\mathbf{X}_n = (X_1, \dots, X_n)$ and suppose that the distribution function, H_n say, of the random variable $R_n(\mathbf{X}_n, F)$ is of interest. The bootstrap methodology can be broken into three simple steps:

- i) Construct an estimator \hat{F}_n of F from \mathbf{X}_n ;
- ii) let $X_1^*, \dots, X_{m_n}^* \sim^{\text{ind}} \hat{F}_n$ be conditionally i.i.d. given \mathbf{X}_n ;
- iii) then let $\mathbf{X}_n^* = (X_1^*, \dots, X_{m_n}^*)$ and estimate H_n by the conditional distribution function of $R_n^* = R(\mathbf{X}_n^*, \hat{F}_n)$ given \mathbf{X}_n ; that is

$$H_n^*(x) = P^*\{R_n^* \leq x\},$$

where $P^*\{\cdot\}$ is the conditional probability given the data \mathbf{X}_n , or equivalently, the entire sequence $\mathbf{X} = (X_1, X_2, \dots)$.

Choices of \hat{F}_n considered below are the EDF \mathbb{F}_n , its least concave majorant \tilde{F}_n , and smoothed versions thereof.

Let d denote the Levy metric or any other metric metrizing weak convergence of distribution functions. We say that H_n^* is *weakly, respectively strongly, consistent* if $d(H_n, H_n^*) \xrightarrow{P} 0$, respectively $d(H_n, H_n^*) \rightarrow 0$ a.s. If H_n has a weak limit H , then consistency requires H_n^* to converge weakly to H , in probability; and if H is continuous, consistency requires

$$\sup_{x \in \mathbb{R}} |H_n^*(x) - H(x)| \xrightarrow{P} 0 \text{ as } n \rightarrow \infty.$$

There is also the apparent possibility that H_n^* could converge to a random limit; that is, that there is a $G : \Omega \times \mathbb{R} \rightarrow [0, 1]$ for which $G(\omega, \cdot)$ is a distribution function for each $\omega \in \Omega$, $G(\cdot, x)$ is measurable for each $x \in \mathbb{R}$, and $d(G, H_n^*) \xrightarrow{P} 0$. This possibility is only apparent, however, if

\hat{F}_n depends only on the order statistics. For if h is a bounded continuous function on \mathbb{R} , then any limit in probability of $\int_{\mathbb{R}} h(x) H_n^*(\omega; dx)$ must be invariant under finite permutations of X_1, X_2, \dots up to equivalence, and thus, must be almost surely constant by the Hewitt-Savage zero-one law [Breiman (1968)]. Let $\bar{G}(x) = \int_{\Omega} G(\omega; x) P(d\omega)$. Then \bar{G} is a distribution function and $\int_{\mathbb{R}} h(x) G(\omega; dx) = \int_{\mathbb{R}} h(x) \bar{G}(dx)$ a.s. for each bounded continuous h and therefore for any countable collection of bounded continuous h . It follows that $G(\omega; x) = \bar{G}(x)$ a.e. ω for all x by letting h approach indicator functions.

Now let

$$\Delta_n = n^{\frac{1}{3}} \left\{ \tilde{f}_n(t_0) - f(t_0) \right\} \text{ and } \Delta_n^* = m_n^{\frac{1}{3}} \left\{ \tilde{f}_{n,m_n}^*(t_0) - \hat{f}_n(t_0) \right\}$$

where $\hat{f}_n(t_0)$ is an estimate of $f(t_0)$, for example $\tilde{f}_n(t_0)$, and $\tilde{f}_{n,m_n}^*(t_0)$ is the Grenander estimator computed from the bootstrap sample $X_1^*, \dots, X_{m_n}^*$. Then weak (strong) consistency of the bootstrap means

$$(3.1) \quad \sup_{x \in \mathbb{R}} |P^*[\Delta_n^* \leq x] - P[\Delta_n \leq x]| \rightarrow 0$$

in probability (almost surely), since the limiting distribution (1.1) of Δ_n is continuous.

3.2. *Bootstrapping from the NPMLE \tilde{F}_n .* Consider now the case in which $m_n = n$, $\hat{F}_n = \tilde{F}_n$, and $\hat{f}_n(t_0) = \tilde{f}_n(t_0)$. Let

$$\mathbb{Z}_n^*(h) := n^{\frac{2}{3}} \left\{ \mathbb{F}_n^*(t_0 + n^{-\frac{1}{3}}h) - \mathbb{F}_n^*(t_0) - \tilde{f}_n(t_0)n^{-\frac{1}{3}}h \right\}$$

for $h \in I_n = [-n^{\frac{1}{3}}t_0, \infty)$, where \mathbb{F}_n^* is the EDF of the bootstrap sample $X_1^*, \dots, X_n^* \sim \tilde{F}_n$. Then $\mathbb{Z}_n^* = \mathbb{Z}_{n,1}^* + \mathbb{Z}_{n,2}$, where

$$\mathbb{Z}_{n,1}^*(h) = n^{\frac{2}{3}} \left\{ (\mathbb{F}_n^* - \tilde{F}_n)(t_0 + n^{-\frac{1}{3}}h) - (\mathbb{F}_n^* - \tilde{F}_n)(t_0) \right\},$$

$$\mathbb{Z}_{n,2}(h) = n^{\frac{2}{3}} \left\{ \tilde{F}_n(t_0 + hn^{-\frac{1}{3}}) - \tilde{F}_n(t_0) - \tilde{f}_n(t_0)n^{-\frac{1}{3}}h \right\}.$$

Further, let \mathbb{W}_1 and \mathbb{W}_2 be two independent two-sided standard Brownian motions on \mathbb{R} with $\mathbb{W}_1(0) = \mathbb{W}_2(0) = 0$,

$$\begin{aligned} \mathbb{Z}_1(h) &= \mathbb{W}_1[f(t_0)h], \\ \mathbb{Z}_2^0(h) &= \mathbb{W}_2[f(t_0)h] + \frac{1}{2}f'(t_0)h^2, \\ \mathbb{Z}_2(h) &= L_{\mathbb{R}}\mathbb{Z}_2^0(h) - L_{\mathbb{R}}\mathbb{Z}_2^0(0) - (L_{\mathbb{R}}\mathbb{Z}_2^0)'(0)h, \\ \mathbb{Z} &= \mathbb{Z}_1 + \mathbb{Z}_2. \end{aligned}$$

Then Δ_n^* equals the left derivative at $h = 0$ of the LCM of \mathbb{Z}_n^* . It is first shown that \mathbb{Z}_n^* converges in distribution to \mathbb{Z} but the conditional distributions of \mathbb{Z}_n^* do not have a limit. The following two lemmas are needed.

LEMMA 3.1. *Let W_n and W_n^* be random vectors in \mathbb{R}^l and \mathbb{R}^k respectively; let Q and Q^* denote distributions on the Borel sets of \mathbb{R}^l and \mathbb{R}^k ; and let \mathcal{F}_n be sigma-fields for which W_n is \mathcal{F}_n -measurable. If the distribution of W_n converges to Q and the conditional distribution of W_n^* given \mathcal{F}_n converges in probability to Q^* , then the joint distribution of (W_n, W_n^*) converges to the product measure $Q \times Q^*$.*

PROOF. The above lemma can be proved easily using characteristic functions. Kosorok (2007) includes a detailed proof. \square

The next lemma uses a special case of the Convergence of Types Theorem [Loève (1963), page 203]: Let V, W, V_n be random variables and b_n be constants; if V has a non-degenerate distribution, $V_n \Rightarrow V$ as $n \rightarrow \infty$, and $V_n + b_n \Rightarrow W$, then $b = \lim_{n \rightarrow \infty} b_n$ exists and W has the same distribution as $V + b$.

LEMMA 3.2. *Let \mathbf{X}_n^* be a bootstrap sample generated from the data \mathbf{X}_n . Let $Y_n := \psi_n(\mathbf{X}_n)$ and $Z_n := \phi_n(\mathbf{X}_n, \mathbf{X}_n^*)$ where $\psi_n : \mathbb{R}^n \rightarrow \mathbb{R}$ and $\phi_n : \mathbb{R}^{2n} \rightarrow \mathbb{R}$ are measurable functions; and let K_n and L_n be the conditional distribution functions of $Y_n + Z_n$ and Z_n given \mathbf{X}_n , respectively. If there are distribution functions K and L for which L is non-degenerate, $d(K_n, K) \xrightarrow{P} 0$ and $d(L_n, L) \xrightarrow{P} 0$ then there is a random variable Y for which $Y_n \xrightarrow{P} Y$.*

PROOF. If $\{n_k\}$ is any subsequence then there exists a further subsequence $\{n_{k_l}\}$ for which $d(K_{n_{k_l}}, K) \rightarrow 0$ a.s. and $d(L_{n_{k_l}}, L) \rightarrow 0$ a.s. Then $Y := \lim_{l \rightarrow \infty} Y_{n_{k_l}}$ exists a.s. by the Convergence of Types Theorem, applied conditionally given $\mathbf{X} := (X_1, X_2, \dots)$ with $b_l = Y_{n_{k_l}}$. Note that Y does not depend on the sub-sequence n_{k_l} , since two such sub-sequences can be joined to form another sub-sequence using which we can argue the uniqueness. \square

- THEOREM 3.1. *i) The conditional distribution of $\mathbb{Z}_{n,1}^*$ given $\mathbf{X} = (X_1, X_2, \dots)$ converges a.s. to the distribution of \mathbb{Z}_1 .*
ii) The unconditional distribution of $\mathbb{Z}_{n,2}$ converges to that of \mathbb{Z}_2 and the unconditional distributions of $(\mathbb{Z}_{n,1}^, \mathbb{Z}_{n,2})$, and \mathbb{Z}_n^* converge to those of $(\mathbb{Z}_1, \mathbb{Z}_2)$ and \mathbb{Z} .*
iii) The unconditional distribution of Δ_n^ converges to that of $(L_{\mathbb{R}}\mathbb{Z})'(0)$, and (3.1) fails.*

- iv) Conditional on \mathbf{X} , the distribution of \mathbb{Z}_n^* does not have a weak limit in probability.
- v) If the conditional distribution function of Δ_n^* converges in probability, then $(L_{\mathbb{R}}\mathbb{Z})'(0)$ and \mathbb{Z}_2 must be independent.

PROOF. i): The conditional convergence of $\mathbb{Z}_{n,1}^*$ follows from Proposition 2.1 with $m_n = n$, $F_n = \tilde{F}_n$, $\mathbb{F}_{n,m_n} = \mathbb{F}_n^*$, applied conditionally given \mathbf{X} . It is only necessary to show that (2.3) holds a.s., and this follows from the Law of the Iterated Logarithm for \mathbb{F}_n and Marshall's Lemma, as explained in Sub-section 2.3. The unconditional limiting distribution of $\mathbb{Z}_{n,1}^*$ must also be that of \mathbb{Z}_1 .

ii): Let

$$\mathbb{Z}_{n,2}^0(h) = n^{\frac{2}{3}} [\mathbb{F}_n(t_0 + n^{-\frac{1}{3}}h) - \mathbb{F}_n(t_0) - f(t_0)n^{-\frac{1}{3}}h],$$

and observe that

$$\mathbb{Z}_{n,2}(h) = L_{I_n}\mathbb{Z}_{n,2}^0(h) - [L_{I_n}\mathbb{Z}_{n,2}^0(0) + (L_{I_n}\mathbb{Z}_{n,2}^0)'(0)h].$$

The unconditional convergence of $\mathbb{Z}_{n,2}^0$ and $L_{I_n}\mathbb{Z}_{n,2}^0$ follow from Corollary 2.7 applied with $F_n \equiv F$, as explained in Sub-section 2.3. The convergence in distribution of $\mathbb{Z}_{n,2}$ now follows from the Continuous Mapping Theorem, using Lemma 2.5 and arguments similar to those in the proof of Corollary 2.6.

It remains to show that $\mathbb{Z}_{n,1}^*$ and $\mathbb{Z}_{n,2}^0$ are asymptotically independent, i.e., the joint limit distribution of $\mathbb{Z}_{n,1}^*$ and $\mathbb{Z}_{n,2}^0$ is the product of their marginal limit distributions. For this it suffices to show that $(\mathbb{Z}_{n,1}^*(t_1), \dots, \mathbb{Z}_{n,1}^*(t_k))$ and $(\mathbb{Z}_{n,2}^0(s_1), \dots, \mathbb{Z}_{n,2}^0(s_l))$ are asymptotically independent, for all choices $-\infty < t_1 < \dots < t_k < \infty$ and $-\infty < s_1 < \dots < s_l < \infty$. This is an easy consequence of Lemma 3.1 applied with $W_n^* = (\mathbb{Z}_{n,1}^*(t_1), \dots, \mathbb{Z}_{n,1}^*(t_k))$ and $W_n = (\mathbb{Z}_{n,2}^0(s_1), \dots, \mathbb{Z}_{n,2}^0(s_l))$, and $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$.

iii): We will appeal to Corollary 2.6 to find the unconditional distribution of Δ_n^* . We already know that \mathbb{Z}_n^* converges in distribution to \mathbb{Z} . That (2.10) holds for the limit \mathbb{Z} can be directly verified from the definition of the process. We only have to show that (2.9) holds unconditionally with $\mathbb{Z}_n = \mathbb{Z}_n^*$.

Let $\epsilon > 0$ and $\gamma > 0$ be given. By Proposition 2.9, there exists $\delta > 0$ and $C > 0$ such that $P(A_n) \geq 1 - \epsilon$ for all $n > N_0$, where

$$A_n := \left\{ \left| \tilde{F}_n(t_0 + h) + \tilde{F}_n(t_0) - f(t_0)h - \frac{1}{2}f'(t_0)h^2 \right| \leq \gamma h^2 + Cn^{-\frac{2}{3}}, \forall |h| \leq \delta \right\}.$$

We can also assume that $|F(t_0 + h) + F(t_0) - f(t_0)h - (1/2)f'(t_0)h^2| \leq \gamma h^2$ for $|h| \leq \delta$. Let $\mathbb{Y}_n^*(h) = n^{\frac{2}{3}}[\mathbb{F}_n^*(t_0 + n^{-\frac{1}{3}}h) - \mathbb{F}_n^*(t_0) - f(t_0)n^{-\frac{1}{3}}h]$, so that

$\mathbb{Z}_n^*(h) = \mathbb{Y}_n^*(h) - \Delta_n h$ for all $h \in I_n$, and

$$L_K \mathbb{Z}_n^* = L_K \mathbb{Y}_n^* - \Delta_n h$$

for all $h \in K$ for any interval $K \subseteq I_n$.

Let $G_n = \tilde{F}_n \mathbf{1}_{A_n} + F \mathbf{1}_{A_n^c}$ and let $P_{G_n}^\infty$ denote the probability when generating the bootstrap samples from G_n . Then G_n satisfies (2.1), (2.3), (2.7) and (2.8) a.s. with $m_n = n$, $F_n = G_n$, $\mathbb{F}_{n,m_n} = \mathbb{F}_n^* \mathbf{1}_{A_n} + \mathbb{F}_n \mathbf{1}_{A_n^c}$ and $f_n = f$. Let J be a compact interval. By Proposition 2.4, applied conditionally, there exists a compact interval K (not depending on ω , by the *remark* near the end of the proof of Proposition 2.4) such that $K \supseteq J$ and

$$P_{G_n}^\infty[L_{I_n} \mathbb{Y}_n^* = L_K \mathbb{Y}_n^* \text{ on } J](\omega) \geq 1 - \epsilon$$

for $n \geq N(\omega)$ for a.e. ω . As $N(\cdot)$ is bounded in probability, there exists $N_1 > 0$ such that $P(B) \geq 1 - \epsilon$, where $B := \{\omega : N(\omega) \leq N_1\}$. By increasing N_1 if necessary, let us also suppose that $N_1 \geq N_0$. Then,

$$\begin{aligned} P[L_{I_{m_n}} \mathbb{Z}_n^* = L_K \mathbb{Z}_n^* \text{ on } J] &= P[L_{I_{m_n}} \mathbb{Y}_n^* = L_K \mathbb{Y}_n^* \text{ on } J] \\ &\geq \int_{A_n} P^*[L_{I_{m_n}} \mathbb{Y}_n^* = L_K \mathbb{Y}_n^* \text{ on } J](\omega) dP(\omega) \\ &= \int_{A_n} P_{G_n}^\infty[L_{I_{m_n}} \mathbb{Y}_n^* = L_K \mathbb{Y}_n^* \text{ on } J](\omega) dP(\omega) \\ &\geq \int_{A_n \cap B} P_{G_n}^\infty[L_{I_{m_n}} \mathbb{Y}_n^* = L_K \mathbb{Y}_n^* \text{ on } J](\omega) dP(\omega) \\ &\geq \int_{A_n \cap B} (1 - \epsilon) dP(\omega) \geq 1 - 3\epsilon, \text{ for all } n \geq N_1 \end{aligned}$$

as $P(A_n \cap B) \geq 1 - 2\epsilon$ for $n \geq N_1$. Thus, (2.9) holds and Corollary 2.6 gives $\Delta_n^* \Rightarrow (L_{\mathbb{R}} \mathbb{Z})'(0)$.

If (3.1) holds in probability, then the unconditional limit distribution of Δ_n^* would be that of $2|\frac{1}{2}f(t_0)f'(t_0)|^{\frac{1}{3}}\mathbb{C}$, which is different from the distribution of $(L_{\mathbb{R}} \mathbb{Z})'(0)$, giving rise to a contradiction.

iv): We use the method of contradiction. Let $Z_n := \mathbb{Z}_{n,1}^*(h_0)$ and $Y_n := \mathbb{Z}_{n,2}(h_0)$ for some fixed $h_0 > 0$ (say $h_0 = 1$) and suppose that the conditional distribution function of $Z_n + Y_n = \mathbb{Z}_n^*(h_0)$ converges in probability to the distribution function G . By Proposition 2.1, the conditional distribution of Z_n converges in probability to a normal distribution, which is obviously non-degenerate. Thus the assumptions of Lemma 3.2 are satisfied and we conclude that $Y_n \xrightarrow{P} Y$, for some random variable Y . It then follows from the Hewitt-Savage zero-one law that Y is a constant, say $Y = c_0$ w.p.1. The

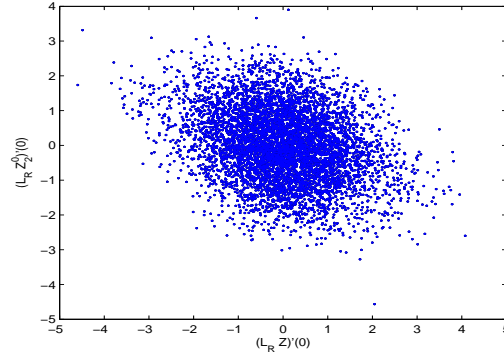


FIG 1. Scatter plot of 10000 random draws of $((L_{\mathbb{R}}Z)'(0), L_{\mathbb{R}}Z_2^0)'(0))$ when $f(t_0) = 1$ and $f'(t_0) = -2$.

contradiction arises since Y_n converges in distribution to $Z_2(h_0)$ which is not a constant a.s.

v): We can show that the (unconditional) joint distribution of $(\Delta_n^*, Z_{n,2}^0)$ converges to that of $((L_{\mathbb{R}}Z)'(0), Z_2^0)$. But Δ_n^* and $Z_{n,2}^0$ are asymptotically independent by Lemma 3.1 applied to $W_n = (Z_{n,2}^0(t_1), Z_{n,2}^0(t_2), \dots, Z_{n,2}^0(t_l))$, where $t_i \in \mathbb{R}$, $W_n^* = \Delta_n^*$ and $\mathcal{F}_n = \sigma(X_1, X_2, \dots, X_n)$. Therefore, $(L_{\mathbb{R}}Z)'(0)$ and Z_2^0 are independent. The proposition follows directly since Z_2 is a measurable function of Z_2^0 . \square

If the conditional distribution of Δ_n^* converges in probability, as a consequence of v) of Theorem 3.1, $(L_{\mathbb{R}}Z)'(0)$ and $(L_{\mathbb{R}}Z_2^0)'(0)$ must also be independent. Figure 1 shows the scatter plot of $(L_{\mathbb{R}}Z)'(0)$ and $(L_{\mathbb{R}}Z_2^0)'(0)$ obtained from a simulation study with 10000 samples, $f(t_0) = 1$ and $f'(t_0) = -2$. The correlation coefficient obtained -0.2999 is highly significant (p-value < 0.0001). Thus, when combined with simulations, v) of Theorem 3.1 strongly suggests that the conditional distribution of Δ_n^* does not converge in probability.

3.3. Bootstrapping from the EDF. A similar, slightly simpler pattern arises if the bootstrap sample is drawn from $\hat{F}_n = F_n$. Define Z_n^* as before, and let $Z_{n,1}^*(h) = n^{\frac{2}{3}}\{(\mathbb{F}_n^* - F_n)(t_0 + n^{-\frac{1}{3}}h) - (\mathbb{F}_n^* - F_n)(t_0)\}$ and $Z_{n,2}(h) = n^{\frac{2}{3}}\{F_n(t_0 + hn^{-\frac{1}{3}}) - F_n(t_0) - \tilde{f}_n(t_0)n^{-\frac{1}{3}}h\}$. Then $Z_n^* = Z_{n,1}^* + Z_{n,2}$. Recall the definition of the processes $\mathbb{W}_1, \mathbb{W}_2, Z_1, Z_2^0$ in Section 3.2. Define

$$Z_2(h) = Z_2^0(h) - (L_{\mathbb{R}}Z_2^0)'(0)h.$$

- THEOREM 3.2. *i) The conditional distribution of $\mathbb{Z}_{n,1}^*$ given $\mathbf{X} = (X_1, X_2, \dots)$ converges a.s. to the distribution of \mathbb{Z}_1 .*
ii) The unconditional distribution of $\mathbb{Z}_{n,2}$ converges to that of \mathbb{Z}_2 and the unconditional distributions of $(\mathbb{Z}_{n,1}^, \mathbb{Z}_{n,2})$, and \mathbb{Z}_n^* converge to those of $(\mathbb{Z}_1, \mathbb{Z}_2)$ and \mathbb{Z} .*
iii) The unconditional distribution of Δ_n^ converges to that of $(L_{\mathbb{R}}\mathbb{Z})'(0)$, and (3.1) fails.*
iv) Conditional on \mathbf{X} , the distribution of \mathbb{Z}_n^ does not have a weak limit in probability.*
v) If the conditional distribution function of Δ_n^ converges in probability, then $(L_{\mathbb{R}}\mathbb{Z})'(0)$ and \mathbb{Z}_2 must be independent.*

REMARK. The proof of this theorem runs along similar lines to that of Theorem 3.1. We briefly highlight the differences.

i): The conditional convergence of $\mathbb{Z}_{n,1}^*$ follows from Proposition 2.1 with $m_n = n$, $F_n = \mathbb{F}_n$, $\mathbb{F}_{n,m_n} = \mathbb{F}_n^*$, applied conditionally given \mathbf{X} . It is only necessary to show that (2.3) is satisfied almost surely, and this follows from the Law of the Iterated Logarithm for \mathbb{F}_n , as explained in sub-section 2.3. Then the unconditional limiting distribution of $\mathbb{Z}_{n,1}^*$ must also be that of \mathbb{Z}_1 .

ii): The proof is similar to that of *ii)* of Theorem 3.1, except that now $\mathbb{Z}_{n,2}(h) = \mathbb{Z}_{n,2}^0(h) - (L_{\mathbb{F}_n}\mathbb{Z}_{n,2}^0)'(0)h$.

The proofs of *iii)* - *v)* are very similar to that of *iii)* - *v)* of Theorem 3.1.

4. Consistent bootstrap methods. The main reason for the inconsistency of bootstrap methods discussed in the previous section is the lack of smoothness of the distribution function from which the bootstrap samples are generated. The EDF \mathbb{F}_n does not have a density, and \tilde{F}_n does not have a differentiable density, whereas F is assumed to have a nonzero differentiable density at t_0 . At a more technical level, the lack of smoothness manifests itself through the failure of (2.4).

The results from Section 2 may be directly applied to derive sufficient conditions on the smoothness of the distribution from which the bootstrap samples are generated. Let $X_1, X_2, \dots \sim^{\text{ind}} F$; let \hat{F}_n be an estimate of F computed from X_1, \dots, X_n ; and let \hat{f}_n be the density of \hat{F}_n or a surrogate, as in Section 3.

THEOREM 4.1. *If (2.1), (2.3), (2.4), (2.7) and (2.8) hold a.s. with $F_n = \hat{F}_n$ and $f_n = \hat{f}_n$, then the bootstrap estimate is strongly consistent, i.e., (3.1) holds w.p.1. In particular, the bootstrap estimate is strongly consistent if there is a $\delta > 0$ for which \hat{F}_n has a continuously differentiable density \hat{f}_n on $[t_0 - \delta, t_0 + \delta]$, and (2.11) holds a.s. with $F_n = \hat{F}_n$ and $f_n = \hat{f}_n$.*

PROOF. That Δ_n^* converges weakly to the distribution on the right side of (1.1) a.s. follows from Corollary 2.7 applied conditionally given \mathbf{X} with $F_n = \hat{F}_n$ and $f_n = \hat{f}_n$. The second assertion follows similarly from Corollary 2.8. \square

4.1. *Smoothing \tilde{F}_n .* We show that generating bootstrap samples from a suitably smoothed version of \tilde{F}_n leads to a consistent bootstrap procedure. To avoid boundary effects and ensure that the smoothed version has a decreasing density on $(0, \infty)$, we use a logarithmic transformation. Let K be a twice continuously differentiable symmetric density for which

$$(4.1) \quad \int_{-\infty}^{\infty} [K(z) + |K'(z)| + |K''(z)|] e^{\eta|z|} dz < \infty$$

for some $\eta > 0$. Let

$$(4.2) \quad \begin{aligned} K_h(x, u) &= \frac{1}{hx} K \left[\frac{1}{h} \log \left(\frac{u}{x} \right) \right], \text{ and} \\ \check{f}_n(x) &= \int_0^{\infty} K_h(x, u) \tilde{f}_n(u) du = \int_0^{\infty} K_h(1, u) \tilde{f}_n(xu) du. \end{aligned}$$

Thus, $e^y \check{f}_n(e^y) = \int_{-\infty}^{\infty} h^{-1} K[h^{-1}(y - z)] \tilde{f}_n(e^z) e^z dz$. Integrating and using capital letters to denote distribution functions,

$$\begin{aligned} \check{F}_n(e^y) &= \int_{-\infty}^y \check{f}_n(e^s) e^s ds \\ &= \int_{-\infty}^y \int_{-\infty}^{\infty} \frac{1}{h} K \left(\frac{s - v}{h} \right) \tilde{f}_n(e^v) e^v dv ds \\ &= \int_{-\infty}^{\infty} K(z) \tilde{F}_n(e^{y-hz}) dz. \end{aligned}$$

Alternatively, integrating (4.2) by parts yields

$$\check{f}_n(x) = - \int_0^{\infty} \frac{\partial}{\partial u} K_h(x, u) \tilde{F}_n(u) du.$$

The proof of (3.1) requires showing that \check{F}_n and its derivatives are sufficiently close to those of F , and it is convenient to separate the estimation error $\check{F}_n - F$ into sampling and approximation error. Thus, let

$$(4.3) \quad \bar{F}_h(e^y) = \int_{-\infty}^{\infty} K(z) F(e^{y-hz}) dz.$$

We denote the first and second derivatives of \bar{F}_h by \bar{f}_h and \bar{f}_h' respectively. Recall that F is assumed to have a non-increasing density on $(0, \infty)$ that is continuously differentiable near t_0 .

LEMMA 4.1. $\lim_{h \rightarrow 0} \|\bar{F}_h - F\| = 0$, and there is a $\delta > 0$ for which

$$(4.4) \quad \lim_{h \rightarrow 0} \sup_{|x-t_0| \leq \delta} [|\bar{f}_h(x) - f(x)| + |\bar{f}'_h(x) - f'(x)|] = 0.$$

PROOF. First observe that

$$\bar{F}_h(e^y) - F(e^y) = \int_{-\infty}^{\infty} K(z)[F(e^{y-hz}) - F(e^y)]dz$$

by (4.3). That $\lim_{h \rightarrow 0} \bar{F}_h(x) = F(x)$ for all $x \geq 0$ follows easily from the Dominated Convergence Theorem, and uniform convergence then follows from Polya's Theorem. This establishes the first assertion of the lemma. Next consider (4.4). Given $t_0 > 0$, let $y_0 = \log(t_0)$ and let $\delta > 0$ be so small that $e^y f(e^y)$ is continuously differentiable (in y) on $[y_0 - 2\delta, y_0 + 2\delta]$. Then,

$$\begin{aligned} \bar{f}_h(x) - f(x) &= \int_{-\infty}^{\infty} K(z)[f(xe^{hz}) - f(x)]e^{hz} dz \\ &+ f(x) \int_{-\infty}^{\infty} (e^{hz} - 1)K(z)dz \end{aligned}$$

and thus

$$\sup_{|x-t_0| \leq \delta} |\bar{f}_h(x) - f(x)| \leq \int_{-\infty}^{\infty} \sup_{|x-t_0| \leq \delta} |f(xe^{hz}) - f(x)|e^{hz} K(z)dz + O(h^2)$$

for any $0 < \delta < t_0$. For sufficiently small δ , the integrand approach zero as $h \rightarrow 0$; and it is bounded by $\sup_{|x-t_0| \leq \delta} (e^{-hz}/x + f(x))e^{hz} K(z)$, since $f(x) \leq 1/x$ for all $x > 0$. So the right side approaches zero as $h \rightarrow 0$ by the Dominated Convergence Theorem. That $\sup_{|x-t_0| \leq \delta} |\bar{f}'_h(x) - f'(x)| \rightarrow 0$ may be established similarly. \square

THEOREM 4.2. *Let K be a twice continuously differentiable, symmetric density for which (4.1) holds. If*

$$h = h_n \rightarrow 0 \quad \text{and} \quad h_n^2 \sqrt{\frac{n}{\log \log(n)}} \rightarrow \infty,$$

then the bootstrap estimator is strongly consistent; that is, (3.1) holds a.s.

PROOF. By Theorem 4.1, it suffices to show that (2.11) holds a.s. with $\hat{F}_n = \check{F}_n$ and $\hat{f}_n = \check{f}_n$; and this would follow from

$$\|\check{F}_n - \bar{F}_h\| + \sup_{|x-t_0| \leq \delta} [|\check{f}_n(x) - \bar{f}_h(x)| + |\check{f}'_n(x) - \bar{f}'_h(x)|] \rightarrow 0 \text{ a.s.}$$

for some $\delta > 0$ and Lemma 4.1. Clearly, using (4.3),

$$(4.5) \quad \check{F}_n(e^y) - \bar{F}_h(e^y) = \frac{1}{h} \int_{-\infty}^{\infty} [\tilde{F}_n(e^t) - F(e^t)] K\left(\frac{y-t}{h}\right) dt$$

for all y , so that

$$\|\check{F}_n - \bar{F}_h\| \leq \|\tilde{F}_n - F\| \leq \|\mathbb{F}_n - F\| = O\left[\sqrt{\log \log(n)/n}\right] \text{ a.s.}$$

by Marshall's Lemma and the Law of the Iterated Logarithm. Differentiating (4.5) gives

$$\check{f}_n(e^y) - \bar{f}_h(e^y) = \frac{e^{-y}}{h^2} \int_{-\infty}^{\infty} [\tilde{F}_n(e^t) - F(e^t)] K'\left(\frac{y-t}{h}\right) dt.$$

Differentiating (4.5) again and then taking absolute values and considering $0 < h \leq 1$, we get

$$\begin{aligned} & \sup_{|x-t_0| \leq \delta} \{|\check{f}_n(x) - \bar{f}_h(x)| + |\check{f}'_n(x) - \bar{f}'_h(x)|\} \\ & \leq \frac{M}{h^3} \sup_{|x-t_0| \leq \delta} \int_{-\infty}^{\infty} |\tilde{F}_n(e^t) - F(e^t)| \left[\left| K'\left(\frac{\log x - t}{h}\right) \right| + \left| K''\left(\frac{\log x - t}{h}\right) \right| \right] dt \\ & \leq \frac{M}{h^2} \|\mathbb{F}_n - F\| \int_{-\infty}^{\infty} [|K'(z)| + |K''(z)|] dz \rightarrow 0 \text{ a.s.} \end{aligned}$$

for a constant $M > 0$, as $h_n^2 \sqrt{n/\log \log(n)} \rightarrow \infty$, where Marshall's Lemma and the Law of Iterated Logarithm have been used again. \square

4.2. m out of n Bootstrap. In Section 3 we showed that the two most intuitive methods of bootstrapping are inconsistent. In this section we show that the corresponding m out of n bootstrap procedures are weakly consistent.

THEOREM 4.3. *If $\hat{F}_n = \mathbb{F}_n$, $\hat{f}_n = \tilde{f}_n$, and $m_n = o(n)$ then the bootstrap procedure is weakly consistent, i.e., (3.1) holds in probability.*

PROOF. Conditions (2.1), (2.3) and (2.8) hold a.s. from (2.13), as explained in Sub-section 2.3. To verify (2.7), let $\gamma > 0$ be given. From the proof of Proposition 2.4 [also see Kim and Pollard (1990), pp 218] there exists $\delta > 0$ such that $|\mathbb{F}_n(t_0 + h) - \mathbb{F}_n(t_0) - F(t_0 + h) + F(t_0)| \leq \gamma h^2 + \mathcal{C}_n n^{-2/3}$, for $|h| \leq \delta$, where \mathcal{C}_n 's are random variables of order $O_P(1)$. We can also

assume that $|F(t_0 + h) + F(t_0) - f(t_0)h - (1/2)f'(t_0)h^2| \leq (1/2)\gamma h^2$ for $|h| \leq \delta$. Then, using the inequality $2|ab| \leq \gamma a^2 + b^2/\gamma$,

$$\begin{aligned}
& \left| \mathbb{F}_n(t_0 + h) - \mathbb{F}_n(t_0) - h\tilde{f}_n(t_0) - \frac{1}{2}h^2f'(t_0) \right| \\
& \leq \left| \mathbb{F}_n(t_0 + h) - \mathbb{F}_n(t_0) - hf(t_0) - \frac{1}{2}h^2f'(t_0) \right| + |h| |\tilde{f}_n(t_0) - f(t_0)| \\
& \leq \left\{ \gamma h^2 + \mathcal{C}_n n^{-\frac{2}{3}} + \frac{1}{2}\gamma h^2 \right\} + \left\{ \frac{1}{2}\gamma h^2 + \frac{1}{2\gamma} |\tilde{f}_n(t_0) - f(t_0)|^2 \right\} \\
(4.6) \quad & \leq 2\gamma h^2 + \mathcal{C}_n n^{-\frac{2}{3}} + O_P(n^{-2/3}) \leq 2\gamma h^2 + o_P(m_n^{-\frac{2}{3}}).
\end{aligned}$$

For (2.4), write

$$\begin{aligned}
m_n^{\frac{2}{3}} \{ \mathbb{F}_n(t_0 + m_n^{-\frac{1}{3}}h) & - \mathbb{F}_n(t_0) - m_n^{-\frac{1}{3}}\tilde{f}_n(t_0)h \} \\
& = m_n^{\frac{2}{3}} \left\{ (\mathbb{F}_n - F)(t_0 + m_n^{-\frac{1}{3}}h) - (\mathbb{F}_n - F)(t_0) \right\} \\
& + m_n^{\frac{1}{3}} [f(t_0) - \tilde{f}_n(t_0)]h + \frac{1}{2}f'(t_0)h^2 + o(1) \\
(4.7) \quad & \xrightarrow{P} \frac{1}{2}f'(t_0)h^2
\end{aligned}$$

uniformly on compacts using Hungarian Embedding to bound the second line and (1.1) (and a two term Taylor expansion) in the third.

Given any subsequence $\{n_k\} \subset \mathbb{N}$, there exists a further subsequence $\{n_{k_l}\}$ such that (4.6) and (4.7) hold a.s. and Theorem 4.1 is applicable. Thus (3.1) holds for the subsequence $\{n_{k_l}\}$, thereby showing that (3.1) holds in probability. \square

Next consider bootstrapping from \tilde{F}_n . We will assume slightly stronger conditions on F , namely, conditions (a)-(d) mentioned in Theorem 7.2.3 of Robertson, Wright and Dykstra (1988):

- (a) $\alpha_1(F) = \inf\{x : F(x) = 1\} < \infty$,
- (b) F is twice continuously differentiable on $(0, \alpha_1(F))$,
- (c) $\gamma(F) = \frac{\sup_{0 < x < \alpha_1(F)} |f'(x)|}{\inf_{0 < x < \alpha_1(F)} f^2(x)} < \infty$,
- (d) $\beta(F) = \inf_{0 < x < \alpha_1(F)} \left| \frac{-f'(x)}{f^2(x)} \right| > 0$.

THEOREM 4.4. *Suppose that (a)-(d) hold. If $\hat{F}_n = \tilde{F}_n$, $\hat{f}_n = \tilde{f}_n$, and $m_n = o[n(\log n)^{-\frac{3}{2}}]$ then (3.1) holds in probability.*

PROOF. Conditions (2.1), (2.3) and (2.8) again follow from (2.13), as explained in Sub-section 2.3. The verification of (2.7) is similar to the argument in the proof of Theorem 4.3. We show that (2.4) holds. Adding and subtracting $m_n^{\frac{2}{3}}[\mathbb{F}_n(t_0 + m_n^{-\frac{1}{3}}h) - \mathbb{F}_n(t_0)]$ from $\mathbb{Z}_{n,2}(h)$ and using (4.7) and the result of Kiefer and Wolfowitz (1976)

$$\begin{aligned} \sup_{|h| \leq c} \left| \mathbb{Z}_{n,2}(h) - \frac{1}{2}f'(t_0)h^2 \right| &\leq 2m_n^{\frac{2}{3}}\|\tilde{F}_n - \mathbb{F}_n\| + o_P(1) \\ &\leq 2m_n^{\frac{2}{3}}\|\tilde{F}_n - \mathbb{F}_n\| + o_P(1) \\ &= O_P[m_n^{\frac{2}{3}}n^{-\frac{2}{3}}\log(n)] + o_P(1) \end{aligned}$$

for any $c > 0$ from which (2.4) follows easily. \square

5. Discussion. We have shown that bootstrap estimators are inconsistent when bootstrap samples are drawn from either the EDF \mathbb{F}_n or its least concave majorant \tilde{F}_n but consistent when the bootstrap samples are drawn from a smoothed version of \tilde{F}_n or an m out of n bootstrap is used. We have also derived necessary conditions for the bootstrap estimator to have a conditional weak limit, when bootstrapping from either \mathbb{F}_n or \tilde{F}_n and presented compelling numerical evidence that these conditions are not satisfied. While these results have been obtained for the Grenander estimator, our results and findings have broader implications for the (in)-consistency of the bootstrap methods in problems with an $n^{\frac{1}{3}}$ convergence rate.

To illustrate the broader implications, we contrast our finding with those of Abrevaya and Huang (2005), who considered a more general framework, as in Kim and Pollard (1990). For simplicity, we use the same notation as in Abrevaya and Huang (2005). Let $W_n := r_n(\theta_n - \theta_0)$ and $\hat{W}_n := r_n(\hat{\theta}_n - \theta_n)$ be the sample and bootstrap statistics of interest. In our case $r_n = n^{\frac{1}{3}}$, $\theta_0 = f(t_0)$, $\theta_n = \tilde{f}_n(t_0)$ and $\hat{\theta}_n = \tilde{f}_n^*(t_0)$. When specialized to the Grenander estimator, Theorem 2 of Abrevaya and Huang (2005) would imply [by calculations similar to those in their Theorem 5 for the NPMLE in a binary choice model] that

$$\hat{W}_n \Rightarrow \arg \max \hat{Z}(t) - \arg \max Z(t)$$

conditional on the original sample, in P^∞ -probability, where $Z(t) = W(t) - ct^2$ and $\hat{Z}(t) = W(t) + \hat{W}(t) - ct^2$, W and \hat{W} are two independent two sided Brownian motions on \mathbb{R} with $W(0) = \hat{W}(0) = 0$ and c is a positive constant depending on F . We also know that $W_n \Rightarrow \arg \max Z(t)$ unconditionally. By v) of Theorem 3.1, this would force the independence of $\arg \max Z(t)$ and

$\arg \max \hat{Z}(t) - \arg \max Z(t)$; but, there is overwhelming numerical evidence that these random variables are correlated.

APPENDIX A: APPENDIX SECTION

LEMMA A.1. *Let $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ be a function such that $\Psi(h) \leq M$ for all $h \in \mathbb{R}$, for some $M > 0$, and*

$$(A.1) \quad \lim_{|h| \rightarrow \infty} \frac{\Psi(h)}{|h|} = -\infty.$$

Then for any $b > 0$, there exists $c_0 > b$ such that for any $c \geq c_0$, $L_{\mathbb{R}}\Psi(h) = L_{[-c,c]}\Psi(h)$ for all $|h| \leq b$.

PROOF. Note that for any $c > 0$, $L_{\mathbb{R}}\Psi(h) \geq L_{[-c,c]}\Psi(h)$ for all $h \in [-c, c]$. Given $b > 0$, consider $c > b$ and $\Phi_c(h) = L_{[-c,c]}\Psi(h)$ for $h \in [-b, b]$, and let Φ_c be the linear extension of $L_{[-c,c]}\Psi|_{[-b,b]}$ outside $[-b, b]$. We will show that there exists $c_0 > b + 1$ such that $\Phi_{c_0} \geq \Psi$. Then Φ_{c_0} will be a concave function everywhere greater than Ψ , and thus $\Phi_{c_0} \geq L_{\mathbb{R}}\Psi$. Hence, $L_{\mathbb{R}}\Psi(h) \leq \Phi_{c_0}(h) = L_{[-c_0,c_0]}\Psi(h)$ for $h \in [-b, b]$, yielding the desired result.

For any $c > b + 1$, $\Phi_c(h) = \Phi_c(b) - \Phi'_c(b) + \Phi'_c(b)(h - b + 1)$ for $h \geq b$. Using the min-max formula

$$\begin{aligned} \Phi'_c(b) &= \min_{-c \leq s \leq b} \max_{b \leq t \leq c} \frac{\Psi(t) - \Psi(s)}{t - s} \\ &\geq \min_{-c \leq s \leq b} \frac{\Psi(b + 1) - \Psi(s)}{(b + 1) - s} \geq \Psi(b + 1) - M =: B_0 \leq 0. \end{aligned}$$

Thus,

$$\begin{aligned} \Phi_c(h) &= \Phi_c(b) - \Phi'_c(b) + \Phi'_c(b)(h - b + 1) \\ &\geq \{\Psi(b) - \Phi'_c(b)\} + \Phi'_c(b)(h - b + 1) \\ &\geq \Psi(b) + (h - b)B_0 \end{aligned}$$

for $h \geq b + 1$. Observe that B_0 does not depend on c . Combining this with a similar calculation for $h < -(b + 1)$, there are $K_0 \geq 0$ and $K_1 \geq 0$, depending only on b , for which $\Phi_c(h) \geq K_0 - K_1|h|$ for $|h| \geq b + 1$. From (A.1), there is $c_0 > b + 1$ for which $\Psi(h) \leq K_0 - K_1|h|$ for all $|h| \geq c_0$ in which case $\Psi(h) \leq \Phi_{c_0}(h)$ for all h . It follows that $L_{\mathbb{R}}\Psi \leq \Phi_{c_0}(h)$ for $|h| \leq b$. \square

LEMMA A.2. *Let \mathbb{B} be a standard Brownian motion. If $a, b, c > 0$, $a^3b = 1$, then*

$$(A.2) \quad P \left[\sup_{t \in \mathbb{R}} \frac{|\mathbb{B}(t)|}{a + bt^2} > c \right] = P \left[\sup_{s \in \mathbb{R}} \frac{|\mathbb{B}(s)|}{1 + s^2} > c \right].$$

PROOF. This follows directly from rescaling properties of Brownian motion by letting $t = a^2 s$. \square

PROOF OF PROPOSITION 2.4. Let $J = [a_1, a_2]$ and $\epsilon > 0$ be as in the statement of the proposition; let $\gamma = |f'(t_0)|/16$; and recall Equations (2.5) and (2.6) from the proof of Proposition 2.1. Then there exists $0 < \delta < 1$, $C \geq 1$, and $n_0 \geq 1$ for which (2.7) and (2.8) hold for all $n \geq n_0$. Let $I_{m_n}^* := [-\delta m_n^{\frac{1}{3}}, \delta m_n^{\frac{1}{3}}]$. By making δ smaller, if necessary, and using Lemma 2.3, $L_{I_{m_n}} \mathbb{Z}_n(h) = L_{I_{m_n}^*} \mathbb{Z}_n(h)$ for $|h| \leq \delta m_n^{\frac{1}{3}}/2$ for all but a finite number of n w.p.1. By increasing the values of C and n_0 , if necessary, we may suppose that the right side of (A.2) (with $c = C$) is less than $\epsilon/3$, that $P[|\eta| > C] + P[\sup_{0 \leq t \leq 1} m_n^{\frac{1}{6}} |\mathbb{E}_{m_n}(t) - \mathbb{B}_{m_n}^0(t)| > C] \leq \epsilon/3$, and that $L_{I_{m_n}} \mathbb{Z}_n = L_{I_{m_n}^*} \mathbb{Z}_n$ on $[-\frac{1}{2}\delta m_n^{\frac{1}{3}}, \frac{1}{2}\delta m_n^{\frac{1}{3}}]$ with probability at least $1 - \epsilon/3$ for all $n \geq n_0$. We can also assume that $\alpha := 8C^3/\gamma > 1$. Then, using Lemma A.2 with $a = \alpha m_n^{-\frac{1}{6}}$ and $b = a^{-3}$, the following relations hold simultaneously with probability at least $1 - \epsilon$ for $n \geq n_0$:

$$\begin{aligned} |\mathbb{B}_{m_n}[F_n(t_0) + s] - \mathbb{B}_{m_n}[F_n(t_0)]| &\leq C(\alpha m_n^{-\frac{1}{6}} + \alpha^{-3} \sqrt{m_n} s^2), \text{ for all } s \\ L_{I_{m_n}} \mathbb{Z}_n &= L_{I_{m_n}^*} \mathbb{Z}_n \text{ on } [-\frac{\delta}{2} m_n^{\frac{1}{3}}, \frac{\delta}{2} m_n^{\frac{1}{3}}], \\ |\eta| \leq C, \text{ and } \sup_{0 \leq t \leq 1} m_n^{\frac{1}{6}} |\mathbb{E}_{m_n}(t) - \mathbb{B}_{m_n}^0(t)| &\leq C. \end{aligned}$$

Let B_n be the event that these four conditions hold. Then $P(B_n) \geq 1 - \epsilon$ for $n \geq n_0$, and from (2.6), B_n implies

$$\begin{aligned} |\mathbb{Z}_{n,1}(h)| &\leq C \left\{ \alpha + \alpha^{-3} m_n^{\frac{2}{3}} [F_n(t_0 + m_n^{-\frac{1}{3}} h) - F_n(t_0)]^2 \right\} + 2C \\ &\quad + C m_n^{\frac{1}{6}} |F_n(t_0 + m_n^{-\frac{1}{3}} h) - F_n(t_0)| \\ (A.3) \quad &\leq 4C \left\{ \alpha + \alpha^{-1} m_n^{\frac{2}{3}} [F_n(t_0 + m_n^{-\frac{1}{3}} h) - F_n(t_0)]^2 \right\} \end{aligned}$$

using the inequalities $|F_n(t_0 + m_n^{-\frac{1}{3}} h) - F_n(t_0)| \leq \alpha m_n^{-\frac{1}{6}} + \alpha^{-1} m_n^{\frac{1}{6}} [F_n(t_0 + m_n^{-\frac{1}{3}} h) - F_n(t_0)]^2$ and $\alpha > 1$. For sufficiently large n , using (2.8), we have

$$\begin{aligned} |\mathbb{Z}_{n,1}(h)| &\leq 4C[\alpha + \alpha^{-1} C^2 m_n^{\frac{2}{3}} (m_n^{-\frac{1}{3}} |h| + m_n^{-\frac{1}{3}})^2] \\ &\leq 4C[\alpha + 2\alpha^{-1} C^2 (h^2 + 1)] \\ (A.4) \quad &= \gamma h^2 + C \end{aligned}$$

for $|h| \leq \delta m_n^{\frac{1}{3}}$ with $\mathcal{C} = 4C\alpha + 8C^3\alpha^{-1}$. Also, we can show that $|\mathbb{Z}_{n,2}(h) - f'(t_0)h^2/2| \leq \gamma h^2 + \mathcal{C}$ for all $|h| \leq \delta m_n^{\frac{1}{3}}$ by (2.7). Let $b_2 > a_2$ be such that $-5\gamma(a_2 + b_2)^2 + 6\gamma(a_2^2 + b_2^2) - 8\mathcal{C} > 0$.

Recalling that $\gamma = -f'(t_0)/16$, B_n implies

$$-10\gamma h^2 - 2\mathcal{C} \leq \mathbb{Z}_n(h) = \mathbb{Z}_{n,1}(h) + \mathbb{Z}_{n,2}(h) \leq -6\gamma h^2 + 2\mathcal{C}$$

for $|h| \leq \delta m_n^{\frac{1}{3}}$ and sufficiently large n . Since the right side is concave, B_n also implies $L_{I_{m_n}^*} \mathbb{Z}_n(h) \leq -6\gamma h^2 + 2\mathcal{C}$ for $|h| \leq \delta m_n^{\frac{1}{3}}$. Therefore, for sufficiently large n , using the upper bound on $L_{I_{m_n}^*} \mathbb{Z}_n$, the lower bound on \mathbb{Z}_n obtained above, and $L_{I_{m_n}} \mathbb{Z}_n(h) = L_{I_{m_n}^*} \mathbb{Z}_n(h)$ for $|h| \leq \delta m_n^{\frac{1}{3}}/2$ on B_n , and $[a_2, b_2] \subset I_{m_n}^*$, we have

$$\begin{aligned} 2\mathbb{Z}_n\left(\frac{a_2 + b_2}{2}\right) - [L_{I_{m_n}} \mathbb{Z}_n(a_2) + L_{I_{m_n}} \mathbb{Z}_n(b_2)] \\ \geq -5\gamma(a_2 + b_2)^2 + 6\gamma(a_2^2 + b_2^2) - 8\mathcal{C} > 0 \end{aligned}$$

with probability at least $1 - \epsilon$. Thus, B_n implies $2\mathbb{Z}_n[\frac{1}{2}(a_2 + b_2)] > L_{I_{m_n}} \mathbb{Z}_n(a_2) + L_{I_{m_n}} \mathbb{Z}_n(b_2)$ with probability at least $1 - \epsilon$. Similarly, B_n implies that there is a $b_1 < a_1$ for which $2\mathbb{Z}_n[\frac{1}{2}(a_1 + b_1)] > L_{I_{m_n}} \mathbb{Z}_n(a_1) + L_{I_{m_n}} \mathbb{Z}_n(b_1)$ with probability at least $1 - \epsilon$. Relation (2.9) then follows from Lemma 2.2. It is worth noting as a *remark* that b_1, b_2 do not depend on the sequence F_n .

Next consider (2.10). Given a compact $J = [-b, b]$, let $c_0(\omega)$ be the smallest positive integer such that for any $c \geq c_0$, $L_{\mathbb{R}} \mathbb{Z}(h) = L_{[-c, c]} \mathbb{Z}(h)$ for $h \in J$. That c_0 exists and is finite w.p.1 follows from Lemma A.1. Defining $W_c := L_{[-c, c]} \mathbb{Z}$ and $Y = L_{\mathbb{R}} \mathbb{Z}$, the event $\{W_c \neq Y \text{ on } J\} \subset \{c_o > c\}$. Now given any $\epsilon > 0$, there exist c such that $P[c_o \leq c] > 1 - \epsilon$. Therefore,

$$P[L_{\mathbb{R}} \mathbb{Z} = L_{[-c, c]} \mathbb{Z} \text{ on } J] \geq P[c_o \leq c] > 1 - \epsilon.$$

□

PROOF OF PROPOSITION 2.9. First consider \mathbb{F}_n . Let $0 < \gamma < |f'(t_0)|/2$ be given. There is a $0 < \delta < \frac{1}{2}t_0$ such that

$$(A.5) \quad \left| F(t_0 + h) - F(t_0) - f(t_0)h - \frac{1}{2}f'(t_0)h^2 \right| \leq \frac{1}{2}\gamma h^2$$

for $|h| \leq 2\delta$. From the proof of Proposition 2.4, using arguments similar to deriving (A.3) and (A.4), we can show that

$$|(\mathbb{F}_n - F)(t_0 + h) - (\mathbb{F}_n - F)(t_0)| < \frac{1}{2}\gamma h^2 + Cn^{-\frac{2}{3}}$$

for $|h| \leq 2\delta$ with probability at least $1 - \epsilon$ for sufficiently large n . Therefore, by adding and subtracting $F(t_0 + h) - F(t_0)$ and using (A.5),

$$(A.6) \quad \left| \mathbb{F}_n(t_0 + h) - \mathbb{F}_n(t_0) - f(t_0)h - \frac{1}{2}f'(t_0)h^2 \right| \leq \gamma h^2 + Cn^{-\frac{2}{3}}$$

for $|h| \leq 2\delta$ with probability at least $1 - \epsilon$ for large n .

Next, consider \tilde{F}_n . Let B_n denote the event that (A.6) holds. Then, $P(B_n)$ is eventually larger than $1 - \epsilon$ and on B_n , we have

$$\mathbb{F}_n(t_0 + h) - \mathbb{F}_n(t_0) - f(t_0)h \leq \left\{ \gamma - \frac{1}{2}|f'(t_0)| \right\} h^2 + Cn^{-\frac{2}{3}},$$

for $|h| \leq 2\delta$. Let E_n be the event that $\tilde{F}_n(h) = L_{[t_0-2\delta, t_0+2\delta]} \mathbb{F}_n(h)$ for $h \in [t_0 - \delta, t_0 + \delta]$. Then by Lemma 2.3, $P(E_n) \geq 1 - \epsilon$, for all sufficiently large n . Taking concave majorants on either side of the above display for $|h| \leq 2\delta$ and noting that the right side of the display is already concave, we have: $\tilde{F}_n(t_0 + h) - \mathbb{F}_n(t_0) - f(t_0)h \leq \left\{ \gamma - \frac{1}{2}|f'(t_0)| \right\} h^2 + Cn^{-\frac{2}{3}}$, for $|h| \leq \delta$ on $B_n \cap E_n$. Setting $h = 0$ shows that on $E_n \cap B_n$, $\tilde{F}_n(t_0) - \mathbb{F}_n(t_0) \leq Cn^{-\frac{2}{3}}$. Now, as $\mathbb{F}_n(t_0) \leq \tilde{F}_n(t_0)$, it is also the case that on $E_n \cap B_n$, for $|h| \leq \delta$,

$$(A.7) \quad \tilde{F}_n(t_0 + h) - \tilde{F}_n(t_0) - f(t_0)h \leq \left\{ \gamma - \frac{1}{2}|f'(t_0)| \right\} h^2 + Cn^{-\frac{2}{3}},$$

Furthermore on $E_n \cap B_n$,

$$\begin{aligned} & \tilde{F}_n(t_0 + h) - \tilde{F}_n(t_0) - f(t_0)h - \frac{1}{2}f'(t_0)h^2 \\ & \geq \mathbb{F}_n(t_0 + h) - \{\mathbb{F}_n(t_0) + Cn^{-\frac{2}{3}}\} - f(t_0)h - \frac{1}{2}f'(t_0)h^2 \\ (A.8) \quad & \geq -\gamma h^2 - 2Cn^{-\frac{2}{3}}. \end{aligned}$$

Therefore, combining (A.7) and (A.8),

$$\left| \tilde{F}_n(t_0 + h) - \tilde{F}_n(t_0) - f(t_0)h - \frac{1}{2}f'(t_0)h^2 \right| \leq \gamma h^2 + 2Cn^{-\frac{2}{3}}$$

for $|h| \leq \delta$ with probability at least $1 - 2\epsilon$ for large n . \square

REFERENCES

- [1] ABREYAYA, J. and HUANG, J. (2005). On the Bootstrap of the Maximum Score Estimator. *Econometrica*, **73** 1175–1204.
- [2] ANDREWS, D. F., BICKEL, P. J., HAMPEL, F. R., HUBER, P. J., ROGERS, W. H. and TUKEY, J. W. (1972). *Robust Estimates of Location*. Princeton Univ. Press, Princeton, N.J.

- [3] BICKEL, P. and FREEDMAN, D. (1981). Some Asymptotic Theory for the Bootstrap. *Ann. Statist.*, **9** 1196–1217.
- [4] BREIMAN, L. (1968). *Probability*. Addison-Wesley Series in Statistics.
- [5] BRUNK, H. D. (1968). Estimation of Isotonic Regression. *Nonparametric Techniques in Statistical Inference*, 177–195. Cambridge Univ. Press.
- [6] CHERNOFF, H. (1964). Estimation of the mode. *Ann. Inst. Statist. Math.*, **16** 31–41.
- [7] CSÖRGÖ, M. and RÉVÉSZ, P. (1981). *Strong Approximations in Probability and Statistics*. Academic Press, New York-Akadémiai Kiadó, Budapest.
- [8] DEVROYE, L. (1987). *A course in Density Estimation*. Birkhäuser, Boston.
- [9] GRENANDER, U. (1956). On the theory of mortality measurement, Part II. *Skand. Akt.*, **39** 125–153.
- [10] GROENEBOOM, P. (1985). Estimating a monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. M. Le Cam and R. A. Olshen, eds.), **2** 539–554. IMS, Hayward, CA.
- [11] GROENEBOOM, P. and WELLNER, J. A. (2001). Computing Chernoff’s distribution. *J. Comput. Graph. Statist.*, **10** 388–400.
- [12] KIEFER, J. and WOLFOWITZ, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrsch. Verw. Gebiete.*, **34** 73–85.
- [13] KIM, J. and POLLARD, D. (1990). Cube-root Asymptotics. *Ann. Statist.*, **18** 191–219.
- [14] KÓMLOS, 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.
- [15] KOSOROK, M. (2007). Bootstrapping the Grenander estimator. *Beyond Parametrics in Interdisciplinary Research: Festschrift in honour of Professor Pranab K. Sen*. IMS Lecture Notes and Monograph Series. Eds.: N. Balakrishnan, E. Pena and M. Silvapulle.
- [16] LEE, S. M. S. and PUN, M. C. (2006). On m out of n Bootstrapping for Nonstandard M-Estimation With Nuisance Parameters. *J. Amer. Statist. Assoc.*, **101** 1185–1197.
- [17] LÉGER, C. and MACGIBBON, B. (2006). On the bootstrap in cube root asymptotics. *Can. J. of Statist.*, **34** 29–44.
- [18] LOÈVE, M. (1963). *Probability Theory*. Van Nostrand, Princeton.
- [19] POLITIS, D. N., ROMANO, J. P. and WOLF, M. (1999). *Subsampling*. Springer-Verlag, New York.
- [20] POLLARD, D. (1984). *Convergence of Stochastic Processes*. Springer-Verlag, New York. Available at <http://www.stat.yale.edu/~pollard/1984book/pollard1984.pdf>
- [21] PRAKASA RAO, B. L. S. (1969). Estimation of a unimodal density. *Sankhyā Ser. A*, **31** 23–36.
- [22] ROBERTSON, T., WRIGHT, F. T. and DYKSTRA, R. L. (1988). *Order restricted statistical inference*. Wiley, New York.
- [23] ROUSSEEUW, P. J. (1984). Least median of squares regression. *J. Amer. Statist. Assoc.*, **79** 871–880.
- [24] SHAO, J. and TU, D. (1995). *The Jackknife and Bootstrap*. Springer-Verlag, New York.
- [25] SHORACK, G. R. and WELLNER, J. A. (1986). *Empirical Processes with Applications to Statistics*. Wiley, New York.
- [26] SINGH, K. (1981). On asymptotic accuracy of Efron’s bootstrap. *Ann. Statist.*, **9** 1187–1195.
- [27] VAN DER VAART, A. W. and WELLNER, J. A. (2000). *Weak Convergence and Empirical Processes*. Springer, New York.
- [28] WANG, X. and WOODROOFE, M. (2007). A Kiefer Wolfowitz Comparison Theorem for Wicksell’s Problem, *Ann. Statist.*, **35** 1559–1575.

- [29] WANG, Y. (1994). The limit distribution of the concave majorant of an empirical distribution function. *Statist. Probab. Lett.*, **20** 81-84.

B. SEN
DEPARTMENT OF STATISTICS
COLUMBIA UNIVERSITY
1255 AMSTERDAM AVENUE
NEW YORK, NY 10027
USA
E-MAIL: bs2528@columbia.edu
URL: <http://www.stat.columbia.edu/~bodhi>

M. BANERJEE
M. WOODROOFE
DEPARTMENT OF STATISTICS
THE UNIVERSITY OF MICHIGAN
1085 SOUTH UNIVERSITY
ANN ARBOR, MI 48109-1107
USA
E-MAIL: moulib@umich.edu
michaelw@umich.edu
URL: <http://www.stat.lsa.umich.edu/~moulib>
<http://www.stat.lsa.umich.edu/~michaelw>